Dynamic Processes on Complex Networks: from Disease Spreading to Neural Activity

by

Christos Nicolaides

B.S., Aristotle University of Thessaloniki (2008)
M.Sc., Imperial College London (2009)
S.M., Massachusetts Institute of Technology (2011)

Submitted to the Department of Civil and Environmental Engineering
in partial fulfillment of the requirements for the degree of

Doctor of Philosophy in Civil and Environmental Engineering

at the

MASSACHUSETTS INSTITUTE OF TECHNOLOGY

June 2014

Signature redacted

Author .....
Department of Civil and Environmental Engineering
May 6, 2014

Certified by ..... Signature redacted
Ruben Juanes
Associate Professor of Civil and Environmental Engineering
Thesis Supervisor

Signature redacted

Accepted by .....
Heidi M. Nepf
Chair, Departmental Committee for Graduate Students

Dynamic Processes on Complex Networks: from Disease Spreading to Neural Activity

by

Christos Nicolaides

Submitted to the Department of Civil and Environmental Engineering
on May 6, 2014, in partial fulfillment of the
requirements for the degree of
Doctor of Philosophy in the field of Civil and Environmental Engineering

Abstract

The study of dynamic processes that take place on heterogeneous networks is essential to better understand, forecast, and manage human activities in an increasingly connected world. In this Thesis, we elucidate the role of the network topology as well as the nature of the underlying processes in a variety of phenomena rooted on highly connected network systems. We use real world applications as the motivation to address three distinct questions.

The first question is: how is the spread of infectious diseases at the global scale mediated by long-range human travel? We show that network topology, geography, traffic structure and individual mobility patterns are all essential for accurate predictions of disease spreading. Specifically, we study contagion dynamics through the air transportation network by means of a stochastic agent-tracking model that accounts for the spatial distribution of airports, detailed air traffic and the correlated nature of mobility patterns and waiting-time distributions of individual agents. We formulate a metric of influential spreading–the geographic spreading centrality–which provides an accurate measure of the early-time spreading power of individual nodes.

The second question is: what is the effect of human behavioral changes in their mobility patterns on the dynamics of contagion through transportation networks? We address this question by developing a model of awareness coupled to disease spreading through mobility networks, where we implement two kinds of behavioral changes: selfish and policy-driven. In analogy with the concept of price of anarchy in transportation networks subject to congestion, we show that maximizing individual utility leads to a loss of welfare for the social group, measured here by the size of the outbreak.

The third question is: what are the mechanisms behind the formation of cell assemblies in neural activity networks? From a neuroscience perspective: How can one explain functional compartmentalization in a globally-connected brain? Here we show that simple mechanisms of neural interaction allow for the emergence of robust cell assemblies through self-organization. We demonstrate the properties of such neural network processes with a minimal-ingredients model of excitation and inhibition between neurons that leads to self-organization of neural activity into local quantized states, even though the underlying network system is globally connected.

Thesis Supervisor: Ruben Juanes
Title: Associate Professor of Civil and Environmental Engineering

Acknowledgments

I am deeply grateful to the family, friends, colleagues, and mentors who have contributed, in one way or another, to the completion of this Thesis. To name only a few, I offer my thanks

to my advisor, Ruben Juanes, for the insight, wisdom, and friendship he has shared with me over the years;

to my co-advisor, Luis Cueto-Felgueroso, for his support and enthusiastic motivation;

to my thesis committee, Marta C. González, and César A. Hidalgo, for their guidance and encouragement;

to all members of the Juanes Research Group for their enthusiastic and collaborative approach to science, their kindness and friendship, which I will miss but hope to continue;

to the members of the Parsons Laboratory and the Civil and Environmental Engineering community that I am sad to leave;

to my parents, Andreas and Maria, and my sister, Ioanna, for their unwavering confidence that I would always meet their very high expectations;

and, now and always, to my wife, Elli Loizidou, for her love and support.

Contents

List of Figures

Chapter 1

Introduction

We live in the age of an increasingly connected world (Lazer et al., 2009). The Internet, the world wide web, and social media are networks that we navigate and explore on a daily basis (Albert et al., 1999; Faloutsos et al., 1999). Mobility, ecological, and epidemiological models rely on networks that consist of entire populations interlinked by the exchange of individuals (Montoya et al., 2006; González et al., 2008; Brockmann et al., 2006; Hancock et al., 2009). Life is based on biological networks like the "connectome" of neural interactions in the brain (Bullmore and Sporns, 2009; Sporns, 2011) and the network of molecular interactions in the body (Jeong et al., 2000; Guimera and Amaral, 2005; Barabási et al., 2007). Network science, therefore, is where we can expect answers to many problems and challenges of our modern world, from controlling traffic flow and flu pandemics to unlocking the mysteries of the human mind (Barabási, 2009).

Over the last decade, the study of complex systems has dramatically expanded across diverse scientific fields, ranging from social sciences and physics to biology and medicine (Albert and Barabási, 2002; Barabási, 2009; Girvan and Newman, 2002; Eagle et al., 2009). This expansion reflects modern trends and currents that have changed the way scientific questions are formulated and research is carried out. In our days science is increasingly concerned with the structure, behavior, and evolution of complex systems from the micro scale–like cells and brains–to the macro scale such as ecosystems, societies or the global economy. To understand these systems, we

require not only knowledge of the elementary system components but also knowledge of the ways in which these components interact and the emergent properties of they interactions. The recent evolution in big data availability and computing power makes it easier than ever before, to record, analyze and model the behavior of complex systems composed of thousands or millions of interacting element components (Lohr, 2012).

All such complex systems display characteristic diverse and organized patters. These patterns emerge as a manifestation of collective behavior between the individual elements, achieved through an intricate web of connectivity. Connectivity comes in many forms–for example, molecular interactions, metabolic pathways, synaptic connections, semantic associations, ecological and food webs, social networks, web hyperlinks, human mobility and transportation, economic exchanges between countries or citation pattern (Jeong et al., 2000; Spirin and Mirny, 2003; Sporns, 2011; Steyvers and Tenenbaum, 2005; DallAsta et al., 2006; Montoya et al., 2006; Dune et al., 2002; González et al., 2007; Apicella et al., 2012; Rutherford et al., 2013; Aral and Walker, 2012; Schweitzer et al., 2009; Wang et al., 2013). In all of these cases, the quantitative analysis of connectivity and other structural properties requires sophisticated mathematical and statistical tools (Albert and Barabási, 2002).

The study of complex systems began with the effort to identify their structure and develop models that can reproduce their statistical properties. The first model was proposed by Erdős and Rényi at the end of the 1950s (Erdos and Rényi, 1960) and was at the basis of most studies until recently. They assumed that nodes in complex systems are wired randomly together, a hypothesis that was adopted by sociology, biology, and computer science at the second half of the 20th century. It had considerable predictive power, explaining for example why everybody is only six handshakes from anybody else , a phenomenon observed as early as 1929 and is well known as *'the six degrees of separation'* (Milgram, 1967) . However, this model failed to explain a common property of social networks where cliques form, representing circles of friends or acquaintances in which every member knows every other member (Jin et al., 2001). This latter property is characteristic of ordered

regular lattices.

In 1998, the interest in networks was however renewed when Watts and Strogatz extracted stylized facts about the properties of real-world networks. They show that a large variety of socio-technical and biological networks exhibit the so-called *small-world* property of being both highly clustered and having a short path-length and they proposed a new model of random networks that is a simple interpolation between an ordered finite-dimensional lattice and a random graph (Watts and Strogatz, 1999).

In the above models, the number of nodes a node is connected with (degree or connectivity) is similar for all the nodes. In detail, the degree distribution of a random graph follows a a Bionomial distribution for small system sizes and Poisson distribution in the large system limit (Newman et al., 2001). One of the most interesting developments in our understanding of complex networks was the recent discovery that for most large real-world networks the degree distribution significantly deviates from a Poisson distribution. In particular, for a large number of networks, including the World Wide Web (Albert et al., 1999), the Internet (Faloutsos et al., 1999), or metabolic networks (Jeong et al., 2000), the degree distribution has a power-law tail,

$$P(k) \sim k^{-\gamma}. \tag{1.1}$$

indicate the lack of scale. Such systems are usually called *scale free networks* (Barabási and Albert, 1999). While some networks display an exponential tail, often the functional form of $P(k)$ still deviates significantly from the Poisson distribution expected for a random graph.

The origin of the power-law degree distribution observed in networks was first addressed by Barabási and Albert (1999) (Barabási and Albert, 1999), who argued that the scale-free nature of real networks is rooted in two generic mechanisms shared by many real networks: Growth and Preferential Attachment: (i) *Growth.* Starting with a small number of nodes, at every time step, a new node with open edges is introduced in the system . (ii) *Preferential Attachment.* The probability that the recently introduced node is connected with an already existing node is proportional

to the the degree of the latter. Numerical simulations indicate that this network evolves into a scale-invariant state with the degree of a node following a scale-free distribution with power law exponent close to 3 (Barabási and Albert, 1999).

While the initial research interest focus on characterizing the structure of real complex systems, shortly came the realization that the complexity in structure affects a variety of real-world phenomena. A prototypical example is that of contagion processes. Epidemiologists, computer scientists and social scientists share a common interest in studying contagion phenomena and rely on very similar spreading models for the description of the diffusion of viruses, knowledge and innovations (Lloyd and May, 2001; Goffman and Newill, 1964). Questions concerning how pathogens spread in population networks, how blackouts can spread on a nationwide scale, or how efficiently we can search and retrieve data on large information structures are generally related to the dynamics of spreading and diffusion processes on underlying heterogeneous topologies.

Recent work has shed light in our understanding of how dynamical systems behave on complex systems. The adoption of ideas through social networks (Toole et al., 2012; Centola, 2010), the spreading of diseases through structured populations via human mobility (Balcan and Vespignani, 2011; Belik et al., 2011; Nicolaides et al., 2012), the diffusion of viruses through computer systems (Pastor-Satorras and Vespignani, 2001, 2002) and the neuronal activity that leads to perception in human brain (Belykh et al., 2005; Bullmore and Sporns, 2009; Sporns, 2011) are only a small number of dynamical models that have been studying extensively on network topologies. These models offer a number of interesting and sometimes unexpected insights, whose theoretical understanding represents a new challenge that has considerably transformed the mathematical and conceptual framework for the study of dynamical processes in complex systems (Vespignani, 2012).

In this thesis, we present dynamic models on heterogeneous network topologies in the context of mobility driven epidemic spreading and neuronal activity in human brain networks. We develop analytical, semi-analytical and numerical solutions to the models in several limiting cases and we use these solutions to get insights into real

world processes. We then support our results using data driven simulations. Finally, we discuss the applicability and the limitations of the models and we draw future research directions.

1.1 Influential Spreading During Contagion Dynamics Through the Air Transportation Network

Public health crises of the past decade – such as the 2003 SARS outbreak, which spread to almost forty countries and caused about a thousand deaths (Consortium et al., 2004; Anderson et al., 2004), and the 2009 H1N1 flu pandemic that killed about 300,000 people worldwide (Fraser et al., 2009; Hancock et al., 2009) – have heightened awareness that new viruses or bacteria could spread quickly across the globe, aided by long range travel through the global transportation network (Guimera and Amaral, 2005; Colizza et al., 2006). While epidemiologists and scientists who study complex network systems – such as contagion patterns and information spread in social networks – are working to create mathematical models that describe the worldwide spread of disease, to date these models reflect an emphasis on the asymptotic late-time behavior of contagion processes, typically characterized by infection thresholds and the number of infected cases (Colizza et al., 2007; Meloni et al., 2009; Balcan and Vespignani, 2011; Belik et al., 2011), but leave open the question of what the early-time behavior of an outbreak is (Balcan et al., 2009).

In the second chapter of this thesis, we study contagion dynamics through the air transportation network by means of a stochastic agent-tracking model that accounts for the spatial distribution of airports, detailed air traffic and the correlated nature of mobility patterns and waiting-time distributions of individual agents. From the simulation results and the empirical air-travel data, we formulate a metric of influential spreading–the geographic spreading centrality–which accounts for spatial organization and the hierarchical structure of the network traffic, and provides an accurate

measure of the early-time spreading power of individual nodes [Fig 1-3]. We finally study intervention scenario during an outbreak emergency and we discuss potential policy implications.

Figure 1-1: The map shows flight the U.S. – centric air transportation network. The size of the airport indicates how influential the airport is to globally spread a contagious disease shortly after the outbreak starts (Nicolaides et al., 2012). New York's JFK airport taking the top spot, followed closely by Los Angeles' LAX and Honolulu airport.

1.2 The Price of Anarchy in Mobility Driven Contagion Dynamics

In an epidemic or a bioterrorist attack, the response of government officials could range from a drastic restriction of mobility imposed isolation or total lockdown of a city to moderate travel restrictions in some areas or simple suggestions that people remain at home. Deciding to institute any measure would require officials to weigh the costs and benefits of action, but at present theres little data to guide them on the question of how disease spreads through transportation networks (Ferguson et al., 2006; Hollingsworth et al., 2006; Epstein et al., 2007). However, official policy recommendations by themselves, cannot determine the patterns of human mobility through transportation networks during epidemics. Instead, individual incentives catalyze the behavioral changes of individuals.

In the event of a health emergency, the pursuit of maximum social or individual utility may lead to conflicting objectives in the routing strategies of network users. Individuals tend to avoid exposure so as to minimize the risk of contagion, whereas policymakers aim at coordinated behavior that maximizes the social welfare. In the third chapter of this thesis, we study agent-driven contagion dynamics through transportation networks, coupled to the adoption of either selfish- or policy-driven rerouting strategies. In analogy with the concept of price of anarchy in transportation networks subject to congestion (Youn et al., 2008), we show that maximizing individual utility leads to a loss of welfare for the social group, measured here by the total population infected after an epidemic outbreak (Nicolaides et al., 2013). We test our hypothesis, and discuss its policy implications, through mean-field theories and Monte Carlo simulations on synthetic and data-driven network models [Fig 1-2].

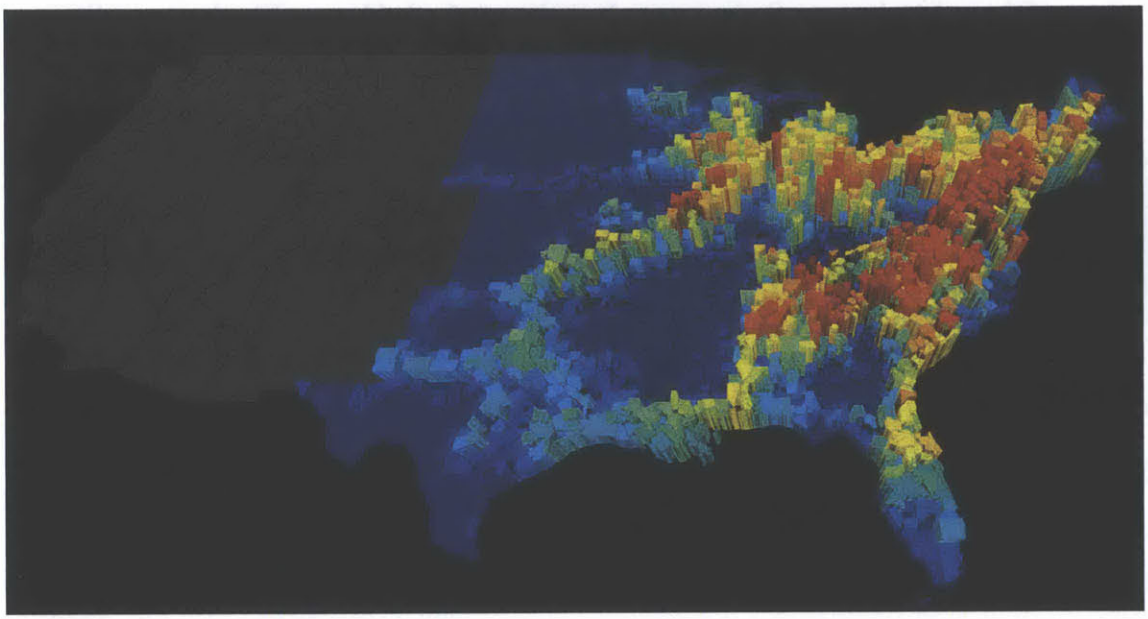

Figure 1-2: The price of anarchy during an epidemic spreading scenario through the US commuting network, calculated two weeks after the outbreak starts from each county in the eastern contiguous US calculated as the difference of the size of the outbreak in the presence of *selfish* and *coordinated* awareness.

1.3 Self-Organization and Quantized States in Neural Activity

The functional activity of neurons in human brain is often organized in finite areas of the cerebral cortex. Recent experiments have shown that distinct concepts and memories are mapped into a small fraction out of the billions of neurons that form the medial temporal lobe of a normal brain (Quiroga et al., 2005). However, what are the mechanisms that allow quantized and localized pattern formation in a globally connected network are still poorly understood (Bear, 1996; Buzsáki, 2010).

Strong nonlinear feedbacks in dynamical systems out of equilibrium lead to the emergence of complex spatiotemporal patterns. Reaction-diffusion systems exhibit a rich variety of self-organized patterns, from stationary dissipative structures and traveling waves, to rotating spirals and chemical turbulence (Smoller, 1983; Cross and Hohenberg, 1993; Vanag and Epstein, 2001; Kim et al., 2001; Kondo and Miura, 2010). In network-organized systems, pattern formation often appears in the form

of synchronization and Turing patterns (Turing, 1952; Nakao and Mikhailov, 2010). However, the mechanisms that allow for the formation of localized patterns of activity in globally interconnected systems is still unknown.

In the fourth chapter of this thesis, we propose a minimal ingredients model of neuron dynamics and synaptic interaction that reproduces both global and local self-organized patterns of activation observed in the brain's functional activity. We relate the characteristics of the pattern formation to both the topological properties of the network and to the nonlinear structure of the underlying process. We finally discuss the implications of our findings in learning, perception and brain computation theories (Nicolaides et al., 2014).

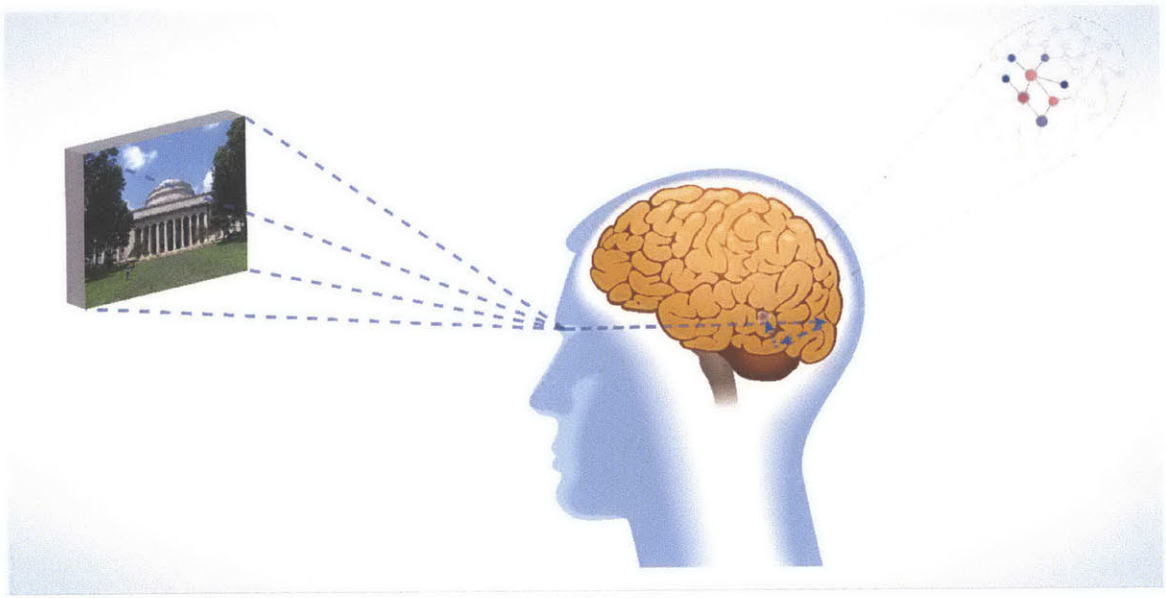

Figure 1-3: The sight of a familiar concept triggers a cascade of brain processes that creates a representation leading to the recognition of the concept through the firing of a finite number of neurons in the temporal lobe of the brain. The big question is: what are the mechanisms that can lead to these kind of localized patterns (cell assemblies) in a globally connected network? In the third chapter of this thesis, we propose a minimal ingredients model of firing in neuronal networks which is able to trigger self-organized "quantized" patterns of activity.

Chapter 2

A metric of influential spreading during a contagion dynamics through the air transportation network

In this chapter, we present a new metric to identify and rank influential spreaders of infectious diseases in human transportation networks. Our metapopulation model of contagion dynamics is based on a time-resolved stochastic description of individual agent mobility through the air transportation system. The model is traffic-driven, and agents traverse the network following empirical stochastic rules that reflect the patterns of individual human mobility (González et al., 2008; Song et al., 2010). These rules include exploration and preferential visit (Song et al., 2010), and distributions of waiting times between successive flights that depend on demography. We show that the late-time spreading, as measured by the global attack, depends strongly on traffic and heterogeneity of transition times. We are interested in characterizing, *a priori*, the early-time spreading potential of individual nodes, as measured by the total square displacement of infected agents. We find that existing metrics of influential spreading—including connectivity (Barabási and Albert, 1999), betweenness

centrality (Guimerá et al., 2005) and k-shell rank (Kitsak et al., 2010; Kempe et al., 2005)—do not successfully capture the spreading ability of individual nodes, as revealed by Monte Carlo simulations. We show that the origin of this disparity lies on the role of geography and traffic on the network (Onnela et al., 2011), and we propose a new metric—the geographic spreading centrality—tailored to early-time spreading in complex networks with spatial imbedding and heterogeneous traffic structure. The results are published in *PLoS ONE* (Nicolaides et al., 2012).

2.1 Motivation

The spreading of infectious diseases is an important example that illustrates the societal impact of global connectivity in man-made transportation systems (Hufnagel et al., 2004; Balcan et al., 2009). Outbreaks expose the vulnerability of current human mobility systems, and challenge our ability to predict the likelihood of a global pandemic, and to mitigate its consequences (Bajardi et al., 2011).

Network models of epidemic spreading have rationalized our understanding of how diseases propagate through a mobile interactome like the human population. "Fermionic" models regard each node as an individual, or a perfectly homogeneous community. In these models, the epidemic threshold for disease spreading vanishes in (infinite-size) scale-free networks, owing to the broad degree distribution (Pastor-Satorras and Vespignani, 2001; Castellano and Pastor-Satorras, 2010). "Bosonic", or metapopulation, models conceptualize nodes as subpopulations that can be occupied by a collection of individuals (Colizza et al., 2007; Colizza and Vespignani, 2007). Metapopulation network models thus recognize that spreading of a disease within a node is not instantaneous. Here we adopt a metapopulation-network approach, precisely because of the interacting timescales for traffic-driven transport between nodes and contagion kinetics within nodes.

It has been shown recently that advection-driven transport, or bias, in complex networks exerts a fundamental control on agent spreading (Nicolaides et al., 2010), leading to anomalous growth of the mean square displacement, in contrast with purely

diffusive processes. The crucial role of traffic-driven transport has also been pointed out in the context of epidemic spreading (Meloni et al., 2009), where it has been shown to directly affect epidemic thresholds.

Given that epidemic spreading is mediated by human travel, and that individual human mobility is far from being random (Brockmann et al., 2006; González et al., 2008; Song et al., 2010), it is natural to ask how the non-Markovian nature of individual mobility affects contagion dynamics. A model of recurrent mobility patterns characterized by a return rate to the individual's origin has recently been incorporated into an otherwise diffusive random-walk metapopulation network model (Balcan and Vespignani, 2011; Belik et al., 2011). A mean-field approximation, as well as Monte Carlo agent-based simulations of the process, reveal a transition separating global invasion from extinction, and show that this transition is heavily influenced by the exponent of the network's degree distribution (Balcan and Vespignani, 2011).

The impact of behavioral changes on the invasion threshold and global attack have recently been analyzed in the context of an SIR infection model (Meloni et al., 2011). In that study it is shown how individual re-routing strategies, where individuals modify their travel paths to avoid infected nodes, influence the invasion threshold and global levels of infection. It is found that selfish individual behavior can have a detrimental effect on society as a whole by inducing a larger fraction of infected nodes, suggesting that the concept of *price of anarchy* in transportation networks (Youn et al., 2008) operates also during disease spreading at the system level.

Taken together, these previous results reflect an emphasis on the asymptotic late-time behavior of contagion processes, typically characterized by infection thresholds and the fraction of infected nodes for both "fermionic" (Meloni et al., 2009; Gómez et al., 2010) and "bosonic" networks (Colizza et al., 2007; Colizza and Vespignani, 2007; Balcan and Vespignani, 2011; Meloni et al., 2011), but leave open the question of what the early-time behavior is (Balcan et al., 2009). Here, we address this question by developing a framework for contagion dynamics on a metapopulation network that incorporates geographic and traffic information, as well as the time-resolved collective transport behavior of individual stochastic agents that carry the disease. Resolving

the temporal dynamics is critical to capture the nontrivial interplay between the transport and reaction timescales.

2.2 Stochastic model of agent mobility

2.2.1 Air transportation data

We develop a stochastic model of human mobility through a US-centric air transportation network. We use air-travel data provided by the Federal Aviation Administration (www.faa.gov) that includes all flights from all domestic and international airlines with at least one origin or destination inside the US (including Alaska and Hawaii), for the period between January 2007 and July 2010. Note that we do not have traffic information about flights whose origin and destination is outside the US. The air transportation network is a space-embedded network with 1833 airports, or nodes, and approximately 50,000 connections, or directed links (Fig. 2-1a). It is a highly heterogeneous network with respect to the degree k (or connectivity) of each node, the population associated with each node, as well as the traffic volume through the links of the network (Guimerá et al., 2005; Meloni et al., 2009). The traffic data is organized in two datasets: "Market" and "Segment". The Market dataset counts trips as origin-to-final-destination, independently of the number of intermediate connecting fights. The Segment dataset counts passengers between pairs of airports, without consideration of the origin and final destination of the whole trip. For example, a passenger that travels from Boston (BOS) to Anchorage (ANC), with connecting flight at Seattle (SEA), would be counted only once in the Market dataset as a passenger from BOS to ANC. In the Segment dataset, however, the passenger would be counted both in the segment BOS-SEA, and in the segment SEA-ANC. From these datasets we extract two weighted matrices that characterize the network traffic: a traffic flux matrix $\mathbf{W}_f = [w_{ij}^f]$ where w_{ij}^f is the yearly passenger traffic from origin i to destination j; and a traffic transport matrix $\mathbf{W}_t = [w_{ij}^t]$ where w_{ij}^t is the yearly passenger traffic in the segment from airport i to airport j.

In addition to the aggregate traffic data, we use information of individual itineraries, provided by a major US airline for domestic trips (Barnhart et al., 2010). This dataset extends over a period of four months in 2004 and includes 3.2 million tickets. We use it to extract the waiting time distribution at final destinations and at connecting airports (Fig. 2-1*b*).

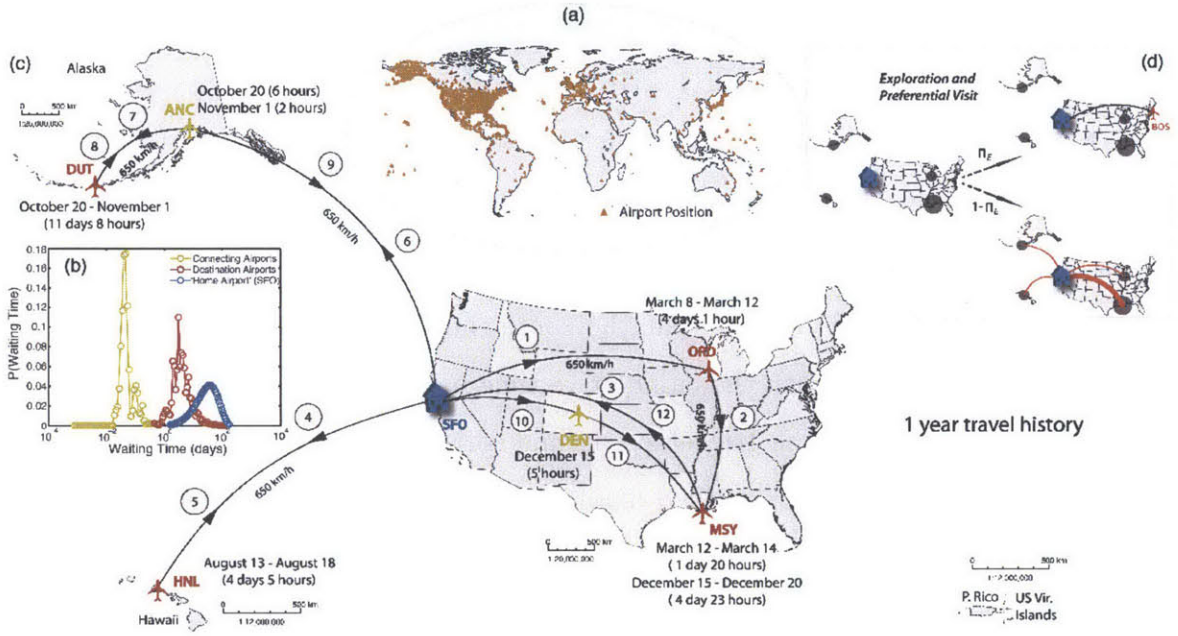

Figure 2-1: **Pictorial view of the key elements of our empirical model of human mobility through the air transportation network.** (*a*) World map with the location of the 1833 airports in the US database from the Federal Aviation Administration (www.faa.gov). (*b*) Waiting time distributions at connecting and destination airports (from (Barnhart et al., 2010)), and at the "home" airport. (*c*) Illustration of a 1-year travel history of an individual with "home" at San Francisco International Airport (SFO). (*d*) Graphical representation of the probabilities for exploration and preferential visit of the same individual, after the 1-year "training period." During exploration the agent visits a new airport while during preferential visit the agent visits a previously-visited place with probability proportional to the frequency of previous visits to that location.

2.2.2 Empirical model

We use the data to build an empirical model of human mobility through the air transportation network. To each airport i, we assign a population P_i by an empirical relation (Colizza et al., 2006), $P_i \sim \sqrt{T_i}$, which reflects a correlation between popu-

lation and yearly total outgoing traffic at that airport, $T_i = \sum_j w_{ij}^f$. Therefore, each individual agent in the model has a "home airport" (Balcan and Vespignani, 2011; Meloni et al., 2011).

Individual agents traverse the network following empirical stochastic rules. Initially, before individuals build up a travel history, each individual positioned at their "home airport" chooses a destination airport with probability proportional to the traffic flux (Meloni et al., 2009, 2011), $\Pi_{ij} \sim w_{ij}^f$. Since the flux matrix accounts for trips in which the individual remains under the same flight number, we allow for an agent choosing some other destination with a small probability, $\Pi_{ik} \sim \min_j w_{ij}^f$.

The agent then establishes an itinerary, or space-time trajectory, to reach the destination. We make the ansatz that the route chosen minimizes a cost function, which generally increases with the cumulative time-in-transit and the monetary cost of the ticket. Given that the trip elapsed time correlates well with the number of connections and the physical travelled distance, and that ticket price decreases with route traffic, we use the following empirical cost function associated with origin i and destination j:

$$C_{ij} = \sum_{\text{all segments}} \frac{d_{kl}^{\delta}}{(w_{kl}^t)^{\varepsilon}}, \tag{2.1}$$

where d_{kl} is the physical distance of the segment $k \to l$ (accounting for the sphericity of the Earth), and the exponents δ and ε lie on the value ranges $0.1 < \delta < 0.3$ and $0.1 < \varepsilon < 0.5$. Which trip route is selected depends on the particular values of δ and ε. The ranges of values for these two parameters are chosen on the basis of producing itineraries that closely match those from real itinerary data (Barnhart et al., 2010). To incorporate in our model the uniqueness of each passenger's needs, we choose a unique combination of these two exponents for each individual. This reflects the current endemic heterogeneity in route selection from the wide range of connections, airline and price choices.

When an agent is off ground, we assume he moves between airports with a constant velocity of 650 km/h. When not flying, an agent can be at one of three distinct places: at their home node, at a connecting airport, or at a destination. The wait-

ing times of an individual at each of these locations is clearly very different. We obtain waiting time distributions for connecting airports and final destinations from the individual mobility dataset (Barnhart et al., 2010), which indeed reflect a very different mean waiting time: in the order of a few hours at connecting airports, and a few days at destinations (Fig. 2-1*b*). Since the dataset lacks individual travel history, we cannot extract waiting times at the home airport, and we assume they are normally distributed (Colizza et al., 2007; Balcan and Vespignani, 2011) with mean $\bar{\tau}_i^h \sim P_i/T_i \sim T_i^{-1/2}$ and standard deviation $\sigma_{\tau_i^h} \sim \bar{\tau}_i^h$, which recognizes that the average person in densely populated areas travels more often. This is based on the empirical relation between total traffic and population of an area (Colizza et al., 2006). For simplicity, we truncate the home waiting time distribution from below at $\tau^h = 1$ day.

An important aspect of our empirical model is the stochastic pattern of individual mobility that we implement. Initially, during a "training period" of ~1 year, we let all agents choose destinations according to a traffic-weighted probability, as explained earlier (Fig. 2-1*c*). However, it is by now well established that individual mobility patterns are far from random (González et al., 2008) and that their statistics can be reproduced with two rules, exploration and preferential visit (Song et al., 2010), which we introduce after the training period, once individuals have built some travel history (Fig. 2-1*d*). During exploration, an agent visits a new airport with probability $\Pi_E = \rho S^{-\gamma}$, where S is the number of airports an agent has visited in the past. We use $\gamma = 0.21 \pm 0.02$ and ρ ($\rho \geq 0$) from a Gaussian distribution with mean $\overline{\rho} = 0.6$ and standard deviation $\sigma_\rho = 0.09$, values that fit human mobility patterns from real mobile phone data (Song et al., 2010). In the absence of comprehensive data for individual long-range travel history, we make the assumption that the parameters used to reproduce local human mobility can be applied for long range travel. The new airport is chosen according to traffic from node i. During preferential visit, the agent selects a previously-visited airport with complementary probability $\Pi_R = 1 - \Pi_E$. For an agent with home at airport i, the probability Π_{ij} of visiting an airport j is proportional to the frequency f_j of previous visits to that location, $\Pi_{ij} \sim f_j$. Because

the travel history built by individuals is mediated by traffic, the mobility model with exploration and preferential visit honors the initial traffic flux matrix.

2.2.3 Monte Carlo simulations of disease spreading

For a single 'mobility' realization, we run our empirical model of human mobility through the air transportation network with 5×10^5 agents that are initially distributed in different "home" subpopulations. During an initial period of one year (training period), the agents are forced to choose destinations according to the traffic flux matrix. During this training period each individual develops a history of mobility patterns. Collectively, the mobility patterns honor the aggregate traffic structure from the dataset. During the second year, we incorporate the exploration and preferential-visit rules to assign destinations to individual agents. We use a time step of 0.5 hours, which we have confirmed is sufficient to resolve the temporal dynamics of the traffic-driven contagion process. For a given 'mobility' realization, we simulate the 'reaction' process as follows: we apply the SIR compartmental model at a randomly chosen time during the first half of the second year by infecting 10 individuals. In the study of late-time global attack, those 10 individuals are selected randomly across the entire network. For the study of early-time spreading, they are selected from the same subpopulation. For the Monte Carlo study, we average the results (global attack and TSD) over 20 mobility and 200 reaction realizations.

2.2.4 Reference models

Our empirical model of human mobility through the air transportation network incorporates a number of dependencies that reflect the complex spatiotemporal structure of collective human dynamics. To understand which of these dependencies are essential, and which affect the modeling results to a lesser degree, we consider four different models of increasing complexity.

In *Model 1*, we consider the US air transportation network but retain only information about the topology of the network. We model mobility as a simplified diffusion

process, in which all individuals perform a synchronous random walk, moving from one node to another, all at the same rate (Colizza and Vespignani, 2007; Colizza et al., 2007). We choose this rate to be the average rate at which individuals travel in our empirical model. Under these assumptions, all nodes with the same degree k have the same behavior. We assign to each node a population corresponding to the stationary state, predicted by the mean-field theory (Colizza and Vespignani, 2007): for a node of degree k, $N_k = \bar{N}k/\langle k \rangle$, where $\langle k \rangle$ denotes the mean of the degree distribution $P_k(k)$, and $\bar{N} = \sum_k N_k P_k(k)$ is the average nodal population.

In *Model 2*, we extend Model 1 by incorporating heterogeneity in the transition rates, as evidenced by the traffic data. To each node i we assign a transition rate $\sigma_i \sim T_i^{1/2}$, but individuals still select a destination randomly, with probability $1/k_i$.

In *Model 3*, we extend Model 2 by enforcing that destination selection by individuals is done according to traffic: the probability of an individual at node i selecting destination j is proportional to w_{ij}^f.

In *Model 4*, we extend Model 3 by considering a simplified model of recurrent mobility patterns (Balcan and Vespignani, 2011; Meloni et al., 2011). Each individual is initially assigned to a "home" node. Individuals perform a random walk through the network of quenched transition rates and heterogeneous traffic, but return to their original subpopulation with a single recurrent rate τ^{-1} (Balcan and Vespignani, 2011). We select $\tau = 7$ days, corresponding to the mean waiting time at destination airports obtained from actual data (Barnhart et al., 2010).

Several important differences exist between the reference models described above and our empirical model of human mobility. For instance, the reference models all discard geographic information. They also all assume that agent displacements are instantaneous and synchronous, taking place at discrete time integers (e.g. one day), and neglect the large heterogeneity in waiting times. We will see that resolving these spatio-temporal processes, while not critical for late-time measures of disease spreading, is essential in the early-time contagion dynamics.

2.3 Global attack

To study the dynamics of disease spreading through the air transportation network, we use the Susceptible–Infected–Recovered (SIR) contagion model. This model divides each subpopulation into a number of healthy (or susceptible, S), infected (I) and recovered (R) individuals, and it is characterized by a contagion reaction,

$$S + I \xrightarrow{\beta} 2I, \tag{2.2}$$

and a recovery reaction,

$$I \xrightarrow{\mu} R, \tag{2.3}$$

where β and μ are the infection and recovery reaction rates, respectively, defined as the number of newly infected (resp. recovered) individuals per unit time for each initial infectious individual in a fully-susceptible subpopulation. Let $(S_i(t), I_i(t), R_i(t))$ be the number of individuals in each class in node i at time t, which satisfy

$$S_i(t) + I_i(t) + R_i(t) \equiv N_i \tag{2.4}$$

at all times. Under the assumption of homogeneous mixing within a city, the probabilities for a susceptible individual to become infected is $\Pi_{S\to I} = 1 - (1 - \beta\Delta t/N_i)^{I_i}$, and for an infected individual to recover is $\Pi_{I\to R} = \mu\Delta t$, which reflect the dependence on the time step Δt. According to these rules, the expected increment in the infected and recovered populations at time $t + \Delta t$ are

$$\Delta I_i = \beta\Delta t I_i(t) S_i(t)/N_i \tag{2.5}$$

and

$$\Delta R_i = \mu\Delta t I_i(t), \tag{2.6}$$

respectively, assuming that during the reaction step Δt the subpopulation does not experience inflow or outflow of individuals. In our model, however, we track the state

of each individual in the network. The reproductive number $R_0 = \beta/\mu$ determines the ratio of newly infected to newly recovered individuals in a homogeneous, well-mixed and fully-susceptible population. From this observation follows the classic result on the epidemic threshold in a single population, $R_0 > 1$. Much work has been devoted to the study of epidemic thresholds in metapopulation networks (Colizza and Vespignani, 2007; Colizza et al., 2007; Balcan and Vespignani, 2011), which generally shows that the reproductive number must be greater than 1 for global spreading of an outbreak.

We apply the SIR contagion model to the four reference models described above and to our empirical mobility model. We employ the *global attack*, defined as the asymptotic (late-time) fraction of the population affected by the outbreak, as our measure of the incidence of the epidemic. We initialize the disease with a small number of infected individuals randomly chosen from the whole population. We obtain representative statistics by performing a Monte Carlo study and averaging over many realizations.

We find that the global attack is quite sensitive to the degree of fidelity of the metapopulation mobility model, especially in the range of low reproductive numbers (Fig. 2-2). Naturally, the global attack increases with R_0 for all models. There is a dramatic difference in the global attack between Models 1 and 2, highlighting the critical influence of quenched disorder in the transition rates σ_i out of individual sub-populations. The global attack increases also from Model 2 to Model 3, reflecting the super-diffusive anomalous nature of spreading when agent displacements are driven by traffic, as opposed to a diffusive random walk (Nicolaides et al., 2010; Meloni et al., 2009). In comparison with these two effects—quenched disorder in transition rates and traffic-driven spreading—recurrent individual mobility patterns (Balcan and Vespignani, 2011; Meloni et al., 2011) have a relatively mild influence on the global attack, as evidenced by the differences between Models 3 and 4. We observe that the additional complexity included in our empirical model—geographic information, high-fidelity individual mobility, and time-resolved agent displacements—induces a slight *delay* in the epidemic threshold with respect to Models 3 and 4, indicating the

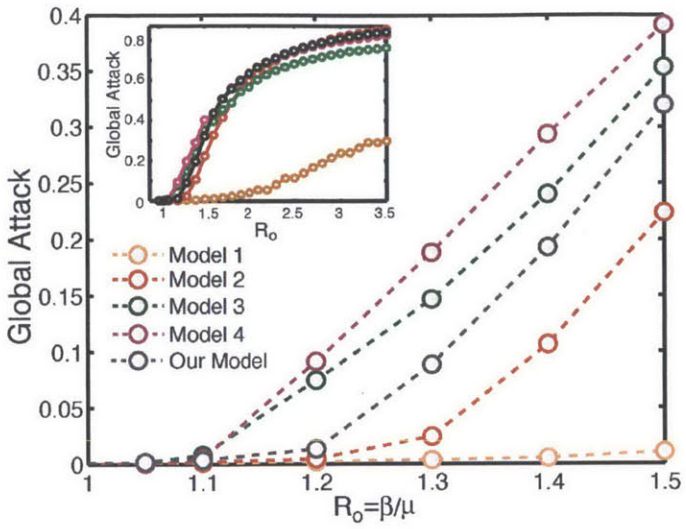

Figure 2-2: **Monte Carlo study of the global attack of an epidemic as a function of the reproductive number R_0, for the different models explained in the text.** We used a value of the recovery rate $\mu^{-1} = 4$ days. We initialized the epidemic with 10 infected individuals chosen randomly across the network. We used a population of 5×10^5 individuals, and average our results over 200 realizations. (Inset) The global attack for larger values of R_0 exhibits smaller differences among models, except for those between annealed and quenched transition rates at the nodes, as evidenced by the simulation results of Model 1 vs. the other models.

nontrivial dependence of contagion dynamics on human mobility.

2.4 Influential spreaders

Finding measures of power and centrality of individuals has been a primary interest of network science (Freeman, 1979; Bonacich, 1987). The very mechanism of preferential attachment shapes the growth and topology of real-world networks (Barabási and Albert, 1999), indicating that the degree of a node is a natural measure of its influence on the network dynamics. Another traditional measure of a node's influence is the betweenness centrality, defined as the number of shortest paths that cross through this node (Freeman, 1979). Betweenness centrality does not always correlate strongly with the degree, the air transportation network being precisely an example of poor correlation between the two (Guimerá et al., 2005). It has been shown, however, that certain dynamic processes such as SIS or SIR epidemic spreading in complex networks

appear to be controlled by a subset of nodes that do not necessarily have the highest degree or the largest betweenness (Kitsak et al., 2010).

Here we revisit what is meant by spreading, and make a crucial distinction between the asymptotic *late-time* behavior—which has been studied more extensively—and the *early-time* dynamics, for which much less is known. We show that the two behaviors are controlled by different mechanisms and, as a result, require different measures of spreading.

2.4.1 Influential spreaders at late times

We perform numerical simulations of epidemic spreading in our model by initializing the SIR compartmental model with infectious individuals at one single subpopulation. We compare the asymptotic, late-time spreading ability of different subpopulations by means of the global attack of the SIR epidemic (Fig. 2-3*a*). We study low values of the reproductive number R_0, between 1 and 1.5, because the relative differences among different sources of infection are largest in this limit. Recent outbreaks of influenza A are estimated to lie within this range (Fraser et al., 2009). We rank the 40 major airports in the United States in terms of their asymptotic global attack, after aggregating the ranking over the range of reproductive numbers studied (Fig. 3*b*). The ability of a node to spread an epidemic depends on fast dispersal of agents to many other nodes, thereby increasing the probability of infectious individuals contacting a large population before they recover. Thus, intuitively, the asymptotic spreading ability of a node increases with its traffic and connectivity. In fact, we find that both degree and traffic provide fair rankings of influential late-time spreaders because in the air transportation network both quantities are strongly correlated (Fig. 2-3*b*, inset).

2.4.2 Influential spreaders at early times

Late-time measures of spreading, such as the asymptotic global attack, cannot capture the details of early-time contagion dynamics. The vigor of initial spreading, however, is likely the crucial aspect in the assessment and implementation of remedial action

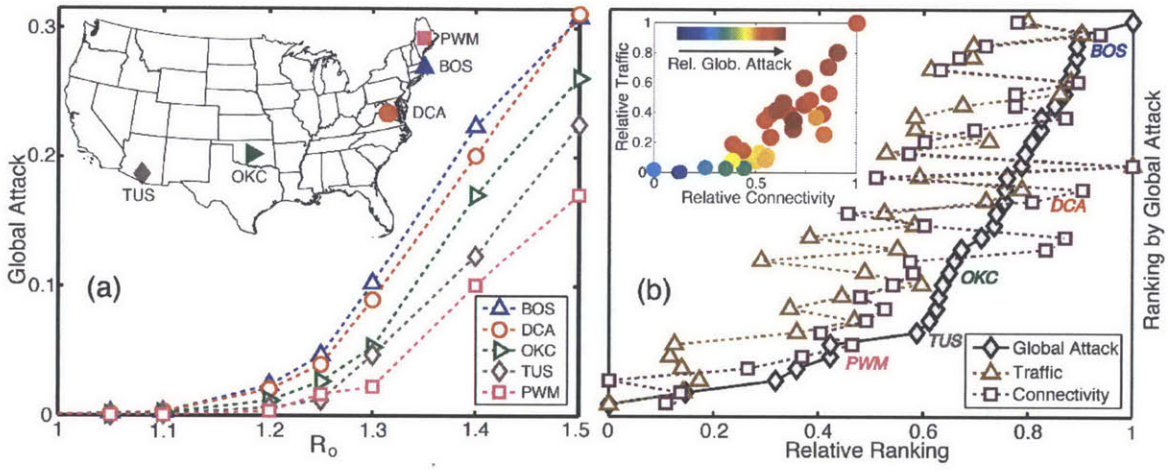

Figure 2-3: **Late-time spreading ability of different airports, measured by the global attack of an SIR epidemic that originates at each airport.** (a) Global attack as a function of reproductive number, for five different airports (see inset). We initialize the disease by infecting 10 randomly chosen individuals inside the subpopulation of consideration. We use $\mu^{-1} = 4$ days. Each point is the result of a Monte Carlo study averaging over 200 reaction and 20 mobility realizations and using 5×10^5 individuals. (b) Ranking of the 40 major airports in US in terms of their spreading ability measured by the normalized global attack. We compare the normalized global-attack ranking curve (black diamonds) to the ones that result from considering the airport's normalized degree (magenta squares) and the airport's normalized traffic (brown triangles). Also shown is the ranking of the airports shown in (a). Both degree and traffic provide effective rankings of influential late-time spreaders, which in this case can be understood from the good cross-correlation between the two (inset).

for highly contagious diseases (Bajardi et al., 2011), when the reaction and transport timescales are comparable.

The natural measure of physical spreading is the *total square displacement* (TSD) of the infected agents,

$$\mathrm{TSD} = \sum_{j=1}^{N_I} (\mathbf{x}_j - \langle \mathbf{x} \rangle)^2 \tag{2.7}$$

where N_I is the total number of infected individuals at time t, $\mathbf{x}_j$ is the position of the infected individual j, and $\langle \mathbf{x} \rangle$ denotes the position of the center of mass of infected individuals. The TSD increases with time as the infected agents, initially all in the same node, spread through the air transportation network by traffic and contact individuals at the connecting and destination nodes.

We compare the TSD for 40 major airports in the US, 10 days after the infection starts at each of those airports, and a reproductive number $R_0 = 1.5$. The random walk described by the infected agents is asynchronous (heterogeneous travel times and waiting times), traffic-driven (quenched disorder in the network fluxes), non-Markovian (recurrent individual mobility patterns) and non-conservative (appearance and disappearance of infected agents due to infection and recovery). This complexity requires that the transport and contagion processes be time-resolved, an essential feature of our model.

We rank all 40 airports according to their TSD at early times. The curve of ordinal ranking vs. normalized TSD is markedly concave, indicating that only a handful of airports are very good spreaders (Fig. 2-4). The list of early-time super-spreaders is led by J. F. Kennedy (JFK) , Los Angeles International (LAX), Honolulu (HNL), San Francisco (SFO) , Newark Liberty (EWR), Chicago O'Hare (ORD) and Washington Dulles (IAD).

We perform a sensitivity analysis with respect to the reproductive number, R_0, and the number of days after which the TSD is measured (Fig. 2-5). Clearly, a higher reproductive number leads to a more aggressive spread of the disease, and therefore larger values of the total square displacement at the same time. From its definition, it is also clear that the TSD increases with time, at least until saturation.

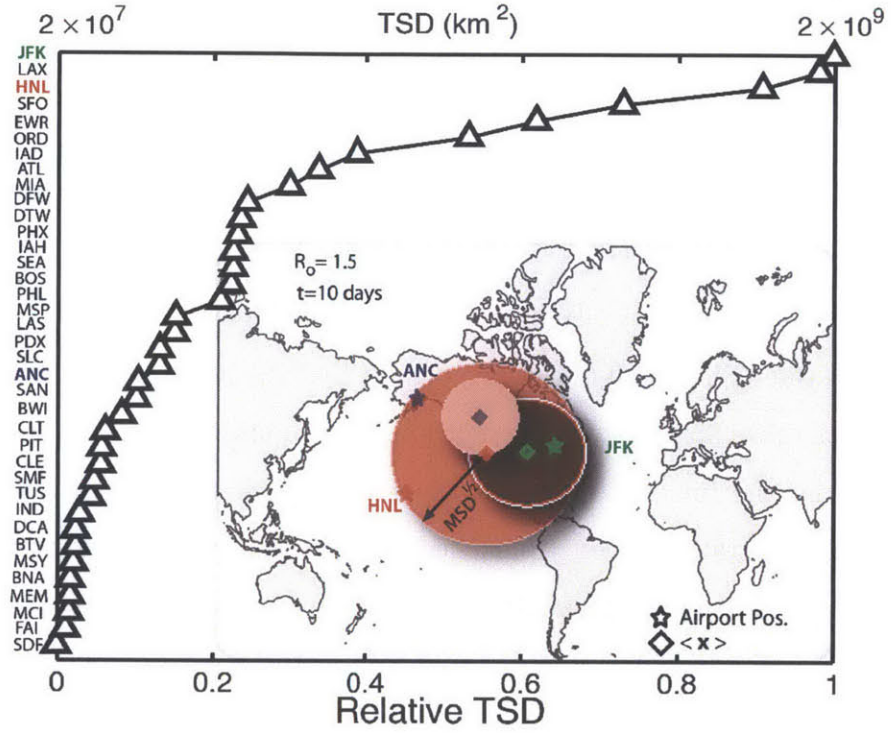

Figure 2-4: **Ranking of influential spreaders by the normalized early-time mean square displacement of infectious individuals.** We initialize the disease by infecting 10 individuals from each specific airport (see inset), and use $\mu^{-1} = 4$ days. Each point is the result of a Monte Carlo study averaging over 100 reaction and 20 mobility realizations and using 5×10^5 individuals. (Inset) Graphical representation of the mean position of infected individuals, 10 days after the outbreak from three different locations. The circle radius denotes the geographic extension of the infectious cloud (as measured by the square root of the Mean Square Displacement (Nicolaides et al., 2010) of infected individuals) while their color represents the number of infected at the same time (dark colors denote large number of infected).

Importantly, while the absolute value of TSD depends strongly on the R_0 and the time of calculation, the ranking of influential spreaders according to TSD appears to be rather insensitive to these parameters, at least for times in the order $t \sim 5 - 20$ days (Fig. 2-5*b*).

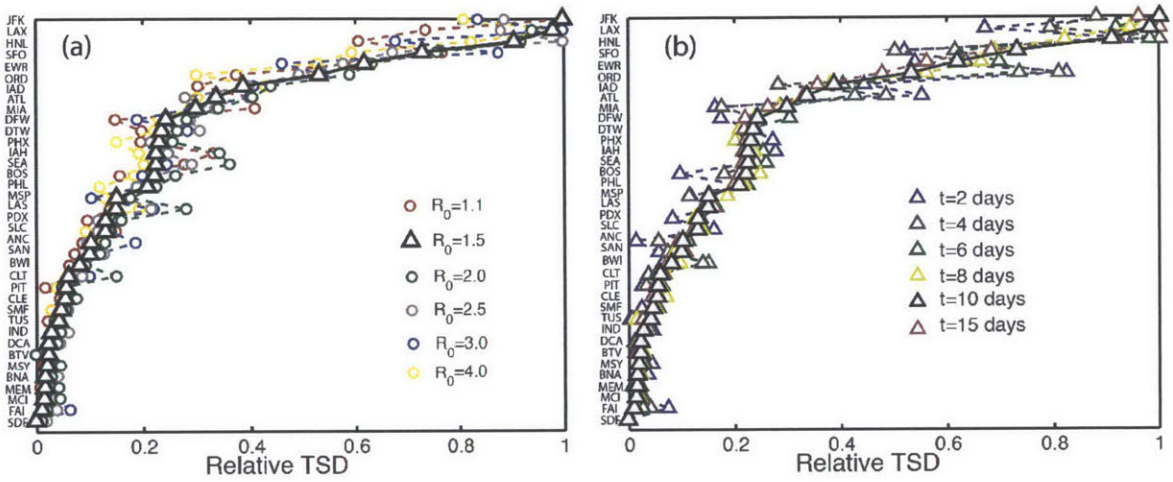

Figure 2-5: *Ranking of influential spreaders by the normalized early-time Total Square Displacement* (a) for different reproductive numbers, 10 days after the disease is initiated. (b) Ranking of influential spreaders by the normalized early-time Total Square Displacement at different times from the initiation of the disease. We use $R_0 = 1.5$ and $\mu^{-1} = 4$ days. Each point in the above plots is the result of a Monte Carlo study averaging over 100 reaction and 20 mobility realizations and using 5×10^5 individuals.

It is instructive to compare the TSD-ranking curve with the rankings provided by existing metrics of centrality and influential spreading, including the normalized degree (Barabási and Albert, 1999) (Fig. 2-6*a*), traffic (Fig. 2-6*b*), betweenness centrality (Guimerá et al., 2005) (Fig. 2-6*c*) and k-shell centrality (Kitsak et al., 2010) (Fig. 2-6*d*). Similar results to those from total traffic are obtained with the eigenvector centrality of the weighted mobility matrix (not shown). All of these metrics deviate significantly from the empirical simulations. For instance, HNL causes large physical spreading, even though it is the airport with the second lowest number of connections, and its traffic is only ~20% of that of Atlanta International (ATL). Equally surprising is that ATL has *both* the largest degree and the largest traffic, yet it comes in 8th place, with an early-time spreading power as low as ~30% that of the best spreader (Fig. 2-6*a*,*b*). Betweenness centrality is able to identify the poor spread-

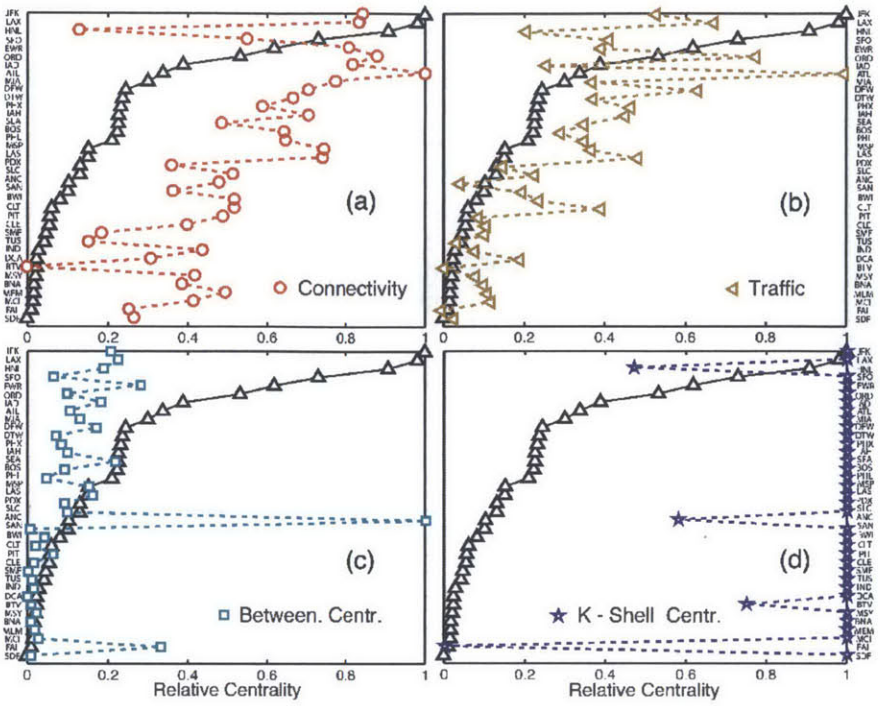

Figure 2-6: **Ranking of influential early-time spreaders by existing metrics. Shown are the results from the model simulations (black triangles), and comparison with the ranking provided by existing metrics of centrality** Shown are the results from the model simulations (black triangles), and comparison with the ranking provided by existing metrics of centrality and late-time influential spreading. (a) Normalized degree. (b) Normalized traffic. (c) Normalized betweenness centrality. (d) Normalized k-shell centrality.

ers, but does not provide accurate ranking or spreading power among the good ones (Fig. 2-6c). For example, Anchorage International (ANC) has the largest betweenness centrality, yet it ranks low as an early-time spreader. The k-shell centrality, which has recently been proposed as an effective metric for identifying influential spreaders at late-time (Kitsak et al., 2010), gives no information about early-time spreading (Fig. 2-6d).

2.4.3 Geographic spreading centrality

It is clear that existing metrics of influential spreading do not properly capture the early-time spreading behavior. We hypothesize that the main reason for this disparity is that they do not account for geographic information and the network's traffic

spatial organization. To test this hypothesis we develop two null networks. As opposed to the reference models presented earlier, which were introduced to incorporate an increasing degree of realism and identify key factors affecting the late-time global attack, the null *networks* employ the *same* empirical model, but modify specific aspects of the network to test whether they have an important bearing on early-time spreading. Null network 1 has the same degree and traffic distributions as the original air transportation network, but changes the geographical information by randomizing the identity of the nodes. In null network 2, we eliminate the traffic quenched disorder by homogenizing outgoing probabilities across the nodes' links, but preserving the position of the nodes. We apply the same mobility and epidemic models and we rank the same airports according to TSD. We find that these rankings are always, for each realization of the null networks, profoundly dissimilar to that of the original network (Fig. 2-7*a*). This confirms the importance of the geographic location of airports, which affects spreading directionality, and the importance of traffic heterogeneity, which affects the routing dynamics, suggesting that both spatial relations and traffic structure are critical elements in early-time spreading.

We also performed a comparison between the detailed empirical model and a model that is identical in all aspects except in that it employs a simpler mobility model. In the simplified model, all agents behave statistically in the same way, with no travel history and with a single return rate (equal to the inverse of the mean waiting time at destinations). The choice of destination from a given origin is random, weighted by traffic from the origin-destination database. A constant time step $\Delta t = 1$ day is used, therefore removing the detailed mobility dynamics. We find that, while the evolution of the TSD does depend on the details of the mobility model, the ranking of spreading power exhibits little dependence (Fig. 2-7*b*), suggesting that individual mobility patterns can be neglected in the construction of a simple metric of influential spreading.

In the light of these observations, we propose a new metric to characterize the ability of an airport to spread an infection spatially at early times, the *geographic spreading centrality* (GSC). We express the vector of airport spreading centralities

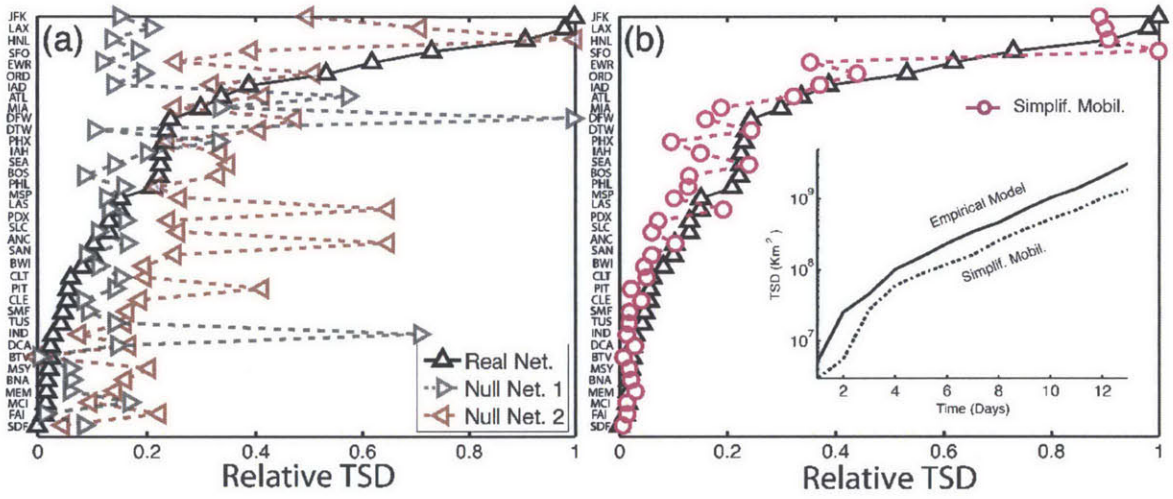

Figure 2-7: **Role of spatial organization, traffic quenched disorder, and mobility patterns, on early-time spreading.** (*a*) Shown is the TSD-ranking for individual realizations of two null networks testing the influence of (1) geographic locations of the nodes, and (2) heterogeneity in the traffic of the links. The dissimilarity between those rankings and that from the original network model strongly suggests that any effective measure of influential early-time spreaders must incorporate geography and traffic quenched disorder. (*b*) TSD-ranking for a simplified model of human mobility. Removing the detailed patterns of mobility affects the evolution of the predicted TSD (see inset for HNL airport) but does not affect the early-time spreading ranking significantly.

$\mathbf{C}_G = \{c_{G,i}\}$ as

$$\mathbf{C}_G = \left[\sum_{m=0}^{\infty} \frac{1}{2^m}\mathbf{\Omega}^m\right]\mathbf{S} = \mathbf{S} + \frac{1}{2}\mathbf{\Omega}\mathbf{S} + \frac{1}{2^2}\mathbf{\Omega}^2\mathbf{S} + \cdots, \tag{2.8}$$

where $\mathbf{\Omega} = [\omega_{ij}]$ is the normalized traffic flux matrix, with $\omega_{ij} = w_{ij}^f/T_i$, and where $\mathbf{S} = \{s_j\}$ is the vector of airport *spreading strengths* (Barrat et al., 2005), defined as

$$s_j = \frac{T_j}{k_j}\sum_{l=1}^{k_j} d_{jl}. \tag{2.9}$$

The spreading strength is a *local* measure that accounts for the node's traffic, degree, and spatial scale of influence. The overall spreading ability of a node, however, must reflect the spreading strength of its neighbors, its neighbors' neighbors, and so on. This notion has led to the classical understanding of the centrality of a node as a generalized eigenvalue problem (Bonacich, 1987), from which our definition of GSC in Eq. 2.8 follows naturally.

We compare the airport rankings predicted by GSC with those obtained from the model simulations, and find excellent quantitative agreement (Fig. 2-8), suggesting that GSC is a reliable *a priori* metric of influential early-time spreaders.

To quantify the correlation between the ranking provided by the TSD and the centrality measures, we use the Kendall tau (τ) rank coefficient (Kendall, 1938). This correlation coefficient indicates how rankings from two quantities are qualitatively correlated and takes a value of -1 if the two rankings are negatively correlated, 0 if the two rankings are independent, and $+1$ if they are positively correlated. The correlation coefficient of the rankings by TSD and connectivity (Fig. 2-6a) is equal to 0.53, by TSD and Traffic (Fig. 2-6b) is equal to 0.57, by TSD and betweenness centrality (Fig. 2-6c) is 0.48 and by TSD and k-shell centrality (Fig. 2-6d) is -0.02. The ranking by the proposed centrality (GSC) and by TSD (Fig. 2-8) are correlated with a Kendall tau of 0.87.

It is worth discussing the spreading power of specific airports in the light of the GSC ranking. Classical measures of centrality, such as total traffic or connectiv-

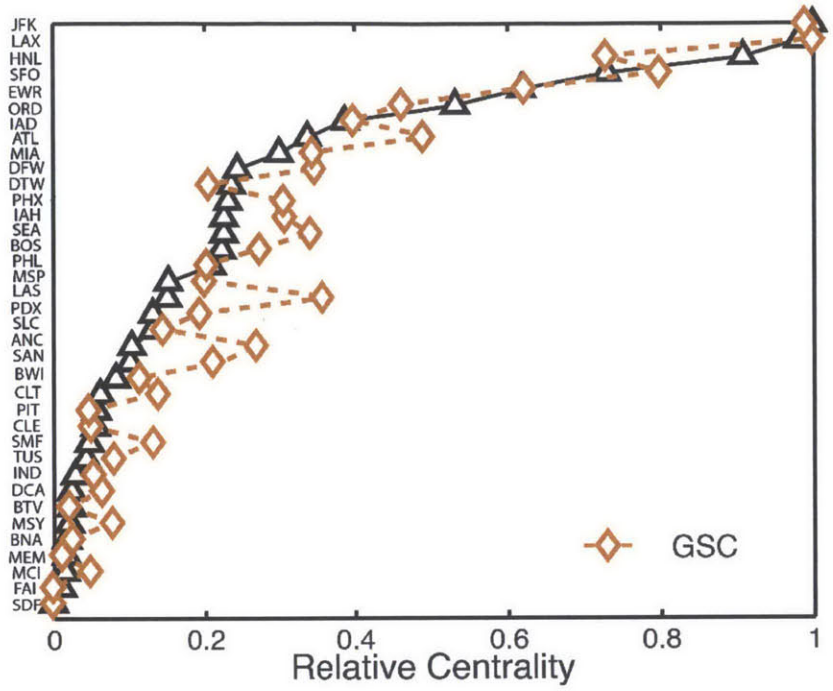

Figure 2-8: **Ranking of influential spreaders at early times from the geographic spreading centrality (GSC).** The GSC metric predictions are in quantitative agreement with the results from the Monte Carlo study on the empirical model.

ity, would suggest that Atlanta International airport (ATL) would have the largest spreading ability. This is clearly not the case, as it ranks 8th in terms of spreading power. The reason is that much of that traffic is of regional nature, within North America, and that many of the connected airports are not, themselves, strong spreaders. The GSC metric allows for a rationalization of the surprising fact that an airport like Honolulu (HNL) ranks third in early-time spreading, very close to JFK and LAX. Despite having a relatively low connectivity (Fig. 2-6*a*) and total traffic (Fig. 2-6*b*), HNL combines three important features that catalyze contagion spreading: (1) it is dominated by long-range travel; (2) it is well connected to other massive hubs, which are themselves powerful spreaders; and (3) it is geographically located such that East-West travel is balanced, thereby maximizing TSD growth. Importantly, these aspects are all captured in the definition of the geographic spreading centrality (Fig. 2-8).

2.5 Discussion

Characterizing the early-time behavior of epidemic spreading is critical to inform decisions during public-health emergencies, and to design regulations aimed at mitigating global pandemics. Here, we show that subpopulations that act as powerful spreaders of infectious diseases at early times—identified by the TSD during the first 10 days of the contagion— differ significantly from the central spreaders in terms of the late-time global attack.

Simulating the infectious dynamics during the initial stages of spreading requires a modeling framework in which transport and contagion processes are time-resolved. We develop a stochastic-agent mobility model through the air transportation network that relies on 3 years of US-centric air travel data and four months of individual travel itineraries. We use this database to build empirical distributions of waiting times at connecting airports and final destinations, and train the model to reproduce the recurrent mobility patterns of individuals. Our analysis demonstrates that the detailed spatiotemporal signatures of individual mobility patterns collectively impact epidemic spreading, especially in the range of low reproductive numbers.

Existing metrics of influential spreaders in networks were not designed to characterize the early-time spreading behavior. Here we propose a new metric, the geographic spreading centrality, which accounts for the local strength in terms of the node's traffic, degree and spatial scale of influence, as well as its global role within the network by incorporating the strength of its neighbors. This metric is able to successfully rank influential spreaders at early times, as evidenced by the agreement between the metric's prediction and detailed Monte Carlo simulations. The geographic spreading centrality opens the door to the quantitative understanding of spreading dynamics on other networks embedded in space, in which topology alone is insufficient to fully characterize the system (Barthélemy, 2011).

In the next Chapter, we incorporate behavioral changes in an effort to close the feedback loop of epidemic spreading and human mobility. We introduce changes in mobility behavior due to awareness of the epidemic, that we model as an additional

contagion process in the system. We predict the behavior of the disease spreading through a mean field approximation on a heterogeneous network topology and we explore the behavior of our model in the data driven scenario of disease spreading through the US commuting network. Our results open new doors to better understand public policy actions during emergencies and could serve as a good indicator to future health authorities policy.

Chapter 3

The price of anarchy in mobility-driven contagion dynamics

In this chapter, we study agent-driven contagion dynamics through transportation networks, coupled to the adoption of either selfish or policy-driven rerouting strategies. In analogy with the concept of price of anarchy in transportation networks subject to congestion, we show that maximizing individual utility leads to a loss of welfare for the social group, measured here by the total population infected after an epidemic outbreak. The results are published in the *Journal of the Royal Society Interface* (Nicolaides et al., 2013).

3.1 Motivation

Users of transportation networks adapt their routing strategies in response to public health emergencies. Changes in the patterns of individual mobility (Brockmann et al., 2006; González et al., 2008; Schneider et al., 2013) are elicited by awareness of the presence of the disease in nearby areas, or imposed by the decisions of policymakers. In an abstract sense, rational travelers adjust their paths to maximize an individual utility function, which depends on travel time and perceived exposure

to contagion risk. In contrast, policymakers attempt to enforce coordinated routing strategies that maximize the social welfare. Policy-driven routing, which is part of the so-called non-pharmaceutical interventions, may significantly reduce the frequency of infectious contacts, containing the transmission of diseases such as influenza (Hatchett et al., 2007). Among other factors, the societal utility function includes measures of the total number of affected individuals, the spatial footprint of the infection, the costs of profilaxis, vaccination and treatment, and the public perception of risk and comfort. The spread of the disease, and the adoption of self-initiated or coordinated rerouting strategies, are strongly coupled. Intuitively, the patterns of balcanvespignani11 through commuting networks modify the dynamics of the contagion process (Belik et al., 2011). Conversely, the dynamics of the disease affect the public perception of the emergency, and determine both the difficulties of implementing the policy and the individual incentives to reduce the risk of being exposed to the disease.

Research on the efficiency of transportation networks subject to congestion demonstrates that routing strategies aimed at maximizing individual utility often lead to a loss of welfare for the social group as a whole (Roughgarden, 2005; Youn et al., 2008; Roughgarden, 2003). Humans tend to follow shortest-path routes that formally minimize their travel times, but these selfish strategies may not yield the social optimum, in the sense that the average travel time increases (Pigou's example (Roughgarden, 2005)). Selfish routing also leads to the counterintuitive effect that network improvements may degrade network performance (Braess' paradox (Roughgarden, 2005)). These and other paradoxical scenarios raise a social dilemma between the pursuit of maximum individual utility and the search for social welfare. Within the framework of game theory, the best options for individual users yield a Nash equilibrium, not necessarily a social optimum (Nash, 1950; Koutsoupias and Papadimitriou, 1999; Youn et al., 2008). The ratio of the total cost of the Nash equilibrium to the total cost of the social optimum is commonly referred to as the the price of anarchy (Koutsoupias and Papadimitriou, 1999; Papadimitriou and Valiant, 2010), indicating the inefficiency of decentralization (Roughgarden, 2005; Youn et al., 2008), and the loss of social welfare due to the selfish behavior of agents in the system.

Here we extend the concept of price of anarchy to mobility-driven contagion dynamics. We study the influence of rerouting, elicited by individual awareness or imposed by public policy, on the dynamics of contagion through transportation networks. We assume that, when individuals are aware of the outbreak and allowed to choose their route from origin to destination, they tend to avoid traversing areas where the disease has been detected. Policymakers, in contrast, strive to enforce coordinated mobility patterns where individuals are segregated according to their health state, thus minimizing the number of infectious contacts. We measure social welfare through the density of infected populations, and define the price of anarchy as the loss of welfare due to selfish rerouting, compared to the policy-driven coordination.

The spread of individual awareness, and the adoption of policy, emerge as central drivers of behavioral change. The role of awareness, understood as knowledge that individuals are willing to act upon (Funk et al., 2010), has been previously studied in the context of network science, both as a mechanism that reduces susceptibility, and therefore infection rates (Bauch and Earn, 2004; Epstein et al., 2008; Funk et al., 2009, 2010; Perra et al., 2011), and as the trigger of self-initiated behavioral changes (Meloni et al., 2011). Meloni et al. (Meloni et al., 2011) analyzed the impact of selfish rerouting on mobility-driven epidemic spreading. In their model, individuals, with a certain probability, avoid traversing infected areas rather than following the shortest path to their destination. Their numerical simulations suggest that individual behavioral changes aimed at slowing down and containing the epidemic may give rise to the opposite effect. In particular, they show simulations where the invasion threshold does not seem to change in spite of rerouting, and the size of the outbreak depends non-monotonically on the traffic through the system; for low traffic, rerouting has a positive effect on the global outbreak, while for high traffic it increases the number of affected nodes.

An open problem is whether the propagation of awareness or policy adoption may enhance or mitigate the impact of an epidemic, in the context of mobility-driven contagion, and considering different rerouting strategies. At the origin of our study is the hypothesis that state-dependent routing behavior, elicited by propagating awareness,

exerts a powerful feedback on the contagion process, potentially impacting the invasion threshold and controlling the density of infected populations at long times. The policy question we want to address is whether the authorities should act to strictly enforce coordinated mobility strategies in the event of an outbreak, or individuals may be allowed to reroute freely without significant losses in social welfare. We test our hypothesis, and discuss its policy implications, through mean-field theories and Monte-Carlo simulations on synthetic and data-driven network models.

3.2 Infection Models

We assume that disease spreading and the propagation of behavioral changes share a common substrate—a commuting or mobility network— which we model as a heterogeneous, uncorrelated network (Albert and Barabási, 2002; Barabási, 2009). The nodes of the network represent populated areas, and the links indicate mobility between populations. The spread of the infection is driven by the mobility of individuals along the links connecting nodes. Individuals travel from an origin node towards a destination node, choosing their path according to a certain routing strategy (Fig. 3-1).

The mathematical epidemiology usually lies on the most simplistic compartmental models of SIR (Susceptible-Infected-Recovered) and SIS (Susceptible-Infected-Susceptible). The SIR model framework is appropriate for infectious diseases that confer lifelong immunity, such as measles or whooping cough (Anderson et al., 1992; Rohani et al., 2000), influenza–like illness or the severe acute respiratory syndrome (Fraser et al., 2009; Balcan et al., 2009; Colizza et al., 2007). The SIS model is predominantly used for sexually transmitted diseases (STDs), such as chlamydia or gonorrhoea, where repeat infections are common (Garnett and Anderson, 1996) as well as for rotaviruses and many bacterial diseases (Parashar et al., 2003).

Here, we are interested in the mechanisms that allow for the spread of the disease, irrespective of the long-term dynamics; that is, whether it will reach an equilibrium endemic state (SIS) or die out after an acute infection peak (SIR). We want to study

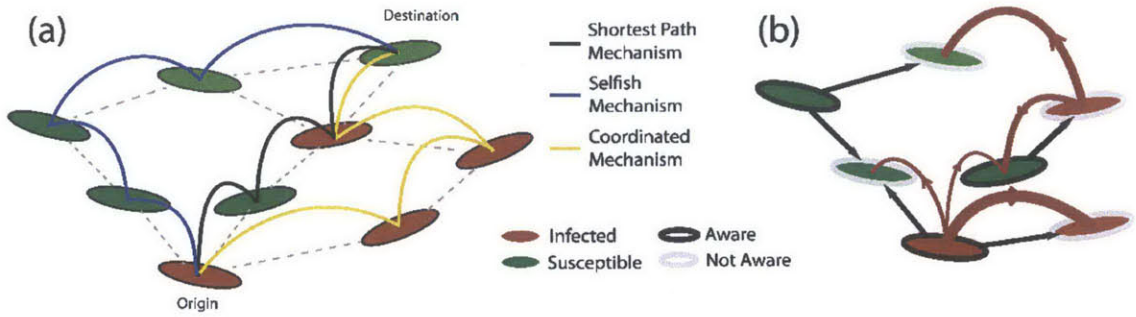

Figure 3-1: **Pictorial illustration of the network model.** (a) The three routing strategies studied in the model. An individual who is not aware of the disease travels to the destination through the shortest path (black). An aware individual that follows the coordinated routing and is located at an infected node is banned from visiting healthy nodes, and follows an "infected path" (yellow). An aware individual that adopts a selfish routing travels to its final destination following a "healthy path" (blue). (b) The two contagion processes in the network –disease spreading and adoption of rerouting behavior. The disease propagates via individual exchanges between the nodes (curved red arrows), while awareness is adopted by non-aware nodes through topological diffusion (black arrows).

the conditions by which the disease spreads through the network. In that sense, we are only interested in the early-time onset dynamics. From the perspective of the early time behavior, the time and spatial scales of the SIS and SIR models are similar.

We are interested in the mechanisms that allow for the spread of the disease in the first place, rather than on the long-term equilibrium. In particular, we emphasize how the spreading process is influenced—enhanced or abated—by mobility and behavioral factors. From a long-term, or equilibrium, perspective, some diseases reach an equilibrium endemic state within the population, while other infectious processes die out after an acute infection peak. The former are better modeled by the SIS model, whereas the latter are better described by the SIR. If one adopts the goal of understanding the early process of disease spreading, either to reach an endemic state or to decay, we believe that it is less important to assume a specific late-time fate of the disease. Hence, we choose to work the minimalistic SIS model which allows for a more detailed analytical study of the epidemic thresholds, but we also show that our conclusions are relevant for diseases better described by the SIR model.

3.2.1 Classical metapopulation models

The traditional approach to model disease spreading coupled to human mobility relies on metapopulation—or bosonic—models (Colizza et al., 2007; Colizza and Vespignani, 2007; Meloni et al., 2011; Vespignani, 2012; Nicolaides et al., 2012). In metapopulation networks, each node has an associated subpopulation of individuals. Infections, modeled as reaction processes, take place as a result of the interaction between individuals inside the subpopulations. For simplicity, it is customary to neglect the influence of internal heterogeneities, assuming full mixing, that is, that all individuals can become in contact with all other individuals. The infection spreads through the network driven by the mobility of individuals, which travel to other subpopulations in the network. In metapopulation networks, the modeler can incorporate a high level of detail in the specification of both the patterns of individual mobility, and the type of infection/reaction process. It is also possible to include behavioral changes elicited by feedbacks (Meloni et al., 2011). The main disadvantage of these sophisticated models is that, when they are coupled to other processes that share the same substrate, such as the spread of awareness about the disease, it is very difficult to develop analytical results to quantify and rationalize the results observed through numerical simulation.

We use a conceptual model of traffic-driven epidemics, originally proposed by Meloni *et al.* (Meloni et al., 2009). As we show later, the model captures the relevant features of mobility-driven disease spreading, coupled to awareness and behavioral changes, in the sense that the results are statistically equivalent to those obtained with a detailed metapopulation model. The advantage is that the simplicity of our model allows us to derive analytical predictions, with reinforce and justify the conclusions derived from simulation.

3.2.2 A conceptual, simplified model

In our model, the nodes can be in the different compartments of the infection model, and the infection spreads from node to node in the system through the exchange

of individuals. The model is fermionic, in the sense that the state of the node is an aggregate variable representing the state of its population. The mobility of individuals aims at representing the traffic heterogeneity in the system, but individuals do not have a particular state. Instead, they adopt the compartmental state of the nodes they traverse. While this description is not aimed at reproducing the detailed dynamics of human travel and recovery from infection, it is designed to capture the statistical signature of the coupled mobility-infection system. Furthermore, it allows us to make progress in the analytical description of the awareness-infection dynamics through the heterogeneous mean field theory.

The compartmental dynamics of the contagion process is given by the susceptible-infected-susceptible (SIS) model, where nodes in the network, at any given time, may be either infected or susceptible to the infection. A susceptible node becomes infected with rate β—the infection rate—when it receives an individual from an infected node. Thus, the larger the number of individuals a healthy node receives from infected nodes, the higher the probability for that node to become infected. An infected node recovers from the infection, becoming susceptible again, with rate μ—the recovery rate. We assume that individuals adopt the health state of the nodes they visit, regardless of their state at the origin or previous legs of their trip. The total number of individuals in the system is λN, where N is the total number of nodes in the network and λ parametrizes the intensity of traffic through the system (Meloni et al., 2009) (see Methods). The SIS model leads to a stationary endemic state in the limit of long times and system size (Castellano and Pastor-Satorras, 2012); we use the density of infected nodes at steady state as a measure of the intensity of the outbreak.

The role of network topology on epidemic spreading has attracted much attention (Pastor-Satorras and Vespignani, 2001; Colizza and Vespignani, 2007; Meloni et al., 2009; Vespignani, 2012). Various studies have demonstrated the impact of connectivity, through the statistics of the nodal degree—number of links of a node, k—and various measures of betweenness and centrality (Freeman, 1977; Barrat et al., 2004; Kitsak et al., 2010; Castellano and Pastor-Satorras, 2012; Nicolaides et al., 2012). We jointly quantify the impact of network topology and routing strate-

gies on the structure of traffic using the concept of *algorithmic betweenness* of a node i, b^i_{alg}, which is the fraction of individual trajectories that traverse that node, $b^i_{\text{alg}} = B^i_{\text{alg}} / \sum_j B^j_{\text{alg}}$, where B^i_{alg} is the number of individuals node i receives (Meloni et al., 2009; Balcan and Vespignani, 2011; Meloni et al., 2011). This quantity is to be understood as a time average.

3.3 Behavioral changes: awareness, rerouting, and policy

In addition to health state, we assign a state of awareness to each node. Awareness spreads through the network as a simple diffusive process (Pastor-Satorras and Vespignani, 2001) and, similar to the infection process, individuals adopt the awareness state of the nodes they traverse. Commuters leaving a non-aware node follow the shortest path between their origin and destination nodes, thus minimizing the number of steps along their path. Commuters leaving aware nodes change their routing strategy; we consider a self-initiated routing behavior, and a policy-driven, or coordinated, strategy. In the selfish rerouting strategy, aware individuals favor routes that avoid infected nodes, irrespective of their own state. In the policy-driven strategy, individuals coordinate their mobility patterns to minimize the global impact of the outbreak. A natural strategy to reach the social optimum is the segregation of travelers according to their health state. Thus, healthy individuals follow routes along which they have minimal exposure to the disease, while infected individuals are banned from visiting healthy nodes.

It is important to emphasize the limits of policy-driven strategies. Because there are costs associated to implementing the policy, it may not be feasible to enforce it, in spite of its potential benefits. Those costs are both material, due to the resources that need to be deployed for the policy to be successful, and in the form of loss of freedom for the individuals. We assume a canonical, perhaps unrealistic policy, where the coordinated action is a nearly-optimal strategy. The role of this idealized

coordinated action is to use is as a reference to define the price of anarchy, even if this strategy may be too costly in practice.

We implement the different routing strategies through a cost-function approach. At a given time, individuals move from the node they are located at to a neighbor node, in such a way that a certain cost function is minimized. In the *selfish* case, we adopt the cost function proposed in (Meloni et al., 2011),

$$C_j = -X_j + h\delta_j, \tag{3.1}$$

where X_j is either -1 if node j is infected, or $+1$ if it is susceptible. The term $h\delta_j$ is introduced to enforce the choice of shortest-path routes when two destinations are possible according to health state. Hence, $h \ll 1$ is a small positive number, and δ_j is equal to -1 if j is one step closer to the destination, 0 if the node j is at the same distance to the destination as node i is, and $+1$ otherwise. In the *coordinated* strategy, we propose the following cost function:

$$C_j = \begin{cases} X_j + h\delta_j & \text{if } i \text{ is Infected} \\ -X_j + h\delta_j & \text{if } i \text{ is Susceptible} \end{cases}$$

In the case of an SIR type of infection, in both strategies, individuals located at recovered nodes move through the shortest path to their destination.

In our model, rerouting is a stochastic process: we define the *degree of awareness* of an aware node i, $\omega_i(t)$, as the probability that an individual inside that node abandons the shortest path and adopts either the selfish or coordinated strategies.

3.4 Mean-field Theory

In the heterogeneous mean-field (HMF) approach (Pastor-Satorras and Vespignani, 2001; Meloni et al., 2009; Balcan and Vespignani, 2011; Vespignani, 2012), nodes with the same number of links—or degree k— are deemed statistically identical. Hence, we may replace nodal variables by degree-aggregates, and seek balance laws for the

evolution of the density of nodes of a given degree that are, e.g. infected or aware. This modeling framework has been successfully applied to describe a wide variety of dynamical processes in complex networks, from epidemic spreading (Pastor-Satorras and Vespignani, 2001) and activator-inhibitor systems (Nakao and Mikhailov, 2010), to coupled oscillators (Arenas et al., 2008) and voter models (Baronchelli et al., 2011). Assuming that there are no topological or dynamic correlations in our system, the HMF approach offers a clear framework to derive analytical expressions for the epidemic threshold in network models of contagion (Pastor-Satorras and Vespignani, 2001; Meloni et al., 2009; Vespignani, 2012; Castellano and Pastor-Satorras, 2012).

Consider the evolution of the relative density of *infected* nodes, $\rho_k(t)$, as well as the relative density of *aware* nodes, $\rho_k^{aw}(t)$, with degree k. The mean-field evolution equations for the two spreading processes are:

$$\frac{\partial \rho_k(t)}{\partial t} = -\mu \rho_k(t) + \beta \lambda N b_{\text{alg}}^k \left(1 - \rho_k(t)\right) \Phi(t), \tag{3.2}$$

$$\frac{\partial \rho_k^{aw}(t)}{\partial t} = -\mu^{aw} \rho_k^{aw}(t) + \beta^{aw} k \left(1 - \rho_k^{aw}(t)\right) \Psi(t). \tag{3.3}$$

The first terms on the right hand side represent *recovery* from infection and loss of awareness, respectively, while the second terms model *activation*. The activation term of the infection process reflects the probability that a node of degree k belongs to the healthy class, $(1 - \rho_k)$, and is infected with rate β when it receives an individual from an infected node, hence the factor $\beta \lambda N b_{\text{alg}}^k$. The probability that an individual travels through a link that points to an infected node, Φ, has contributions from the aware individuals, as well as from the non-aware ones. The activation term for the awareness process quantifies the probability that non-aware nodes, $(1 - \rho_k^{aw})$, become aware via a neighbor node. The probability of this event is proportional to the adoption rate β^{aw}, the number of neighbors, k, and the probability that a given link points to an infected node, Ψ. The structure of the probabilities Φ and Ψ is key for the dynamics of the coupled system (3.2)–(3.3), and determines the critical parameter values beyond which an outbreak propagates through the network: the invasion thresholds. The above mean field representation assumes that the time scales of the epidemic process,

the mobility of individuals, and the spread of awareness, are the same. More precisely, we assume that the characteristic time scale for all these processes is one day.

3.4.1 Invasion thresholds

We assume that the awareness process (Eq. 3.3) is independent from the infection process, so the classical results for diffusive processes in networks apply (Pastor-Satorras and Vespignani, 2001). For an uncorrelated network, Ψ takes the form

$$\Psi = (\sum_{k'} k' P(k') \rho_{k'}^{aw}) / \langle k \rangle, \tag{3.4}$$

and the invasion threshold is simply (Pastor-Satorras and Vespignani, 2001):

$$\beta_c^{aw} = \langle k \rangle / \langle k^2 \rangle, \tag{3.5}$$

where $\langle \cdot \rangle$ denotes averaging over all the nodes in the network. When the adoption rate, β^{aw}, is larger than this critical value, the spreading of awareness causes an endemic state in the system, with a nonzero fraction of aware nodes at steady state. Topological heterogeneities reduce the critical value of the awareness activation rate; in particular, it has been shown that the threshold vanishes for infinite size, scale-free networks with degree exponent $2 \leq \gamma < 3$ (Pastor-Satorras and Vespignani, 2002; Castellano and Pastor-Satorras, 2012).

One of the central contributions of this study is the derivation of the invasion threshold for the infection process, which is subject to strong feedback from the spread of awareness. In the case of selfish rerouting, the probability Φ can be written as

$$\Phi \approx \frac{1}{\langle b_{\text{alg}} \rangle} \sum_{k'} P(k') b_{\text{alg}}^{k'} \rho_{k'} \left[(1 - \rho_{k'}^{aw}) + (1 - \omega) \rho_{k'}^{aw} + \frac{\omega \rho_{k'}^{aw}}{1 - \rho_{k'}} \right]. \tag{3.6}$$

The first term inside the brackets models the influence of individuals that travel from the non-aware nodes, while the second models the non-rerouting individuals that travel from aware nodes, and the third one the rerouting individuals. Imposing

stationarity in Eq. (3.2), we arrive at the epidemic threshold condition,

$$\beta > \beta^c_{\text{null}} = \frac{\langle b_{\text{alg}} \rangle}{\langle b^2_{\text{alg}} \rangle} \frac{\mu}{\lambda N}, \tag{3.7}$$

Remarkably, the critical infection rate does not depend on the dynamics of the adoption process; more precisely, the threshold is independent from the degree of awareness, ω, and from the density of aware nodes in the system, ρ^{aw}.

Coordinated rerouting changes the structure of Φ. Our numerical simulations suggest that, in the policy-driven case, the contribution from the aware individuals is much smaller than that from non-aware ones. Intuitively, infected individuals that adopt the policy travel along paths of infected nodes, vastly reducing the frequency of infectious contacts. Hence, considering only the contribution of the non-aware individuals, the probability Φ reads

$$\Phi \approx \frac{1}{\langle b_{\text{alg}} \rangle} \sum_{k'} P(k') b^{k'}_{\text{alg}} \rho_{k'} \left[(1 - \rho^{aw}_{k'}) + (1 - \omega) \rho^{aw}_{k'} \right]. \tag{3.8}$$

This form of Φ, leads to the invasion threshold

$$\beta > \beta^c = \frac{\langle b_{\text{alg}} \rangle}{\langle b^2_{\text{alg}} \rangle - \omega \langle b^2_{\text{alg}} \rangle_\omega} \frac{\mu}{\lambda N}, \tag{3.9}$$

where $\langle b^2_{\text{alg}} \rangle_\omega = \sum_{k'} P(k') (b^{k'}_{\text{alg}})^2 \rho^{aw}_{k'}$ denotes the second moment of the algorithmic betweenness over the *aware* nodes in the system.

The above thresholds reveal a fundamental difference between the selfish and policy-driven routing strategies: with coordinated routing, awareness plays a central role in the onset of the outbreak, while self-initiated changes do not alter the threshold. The rich phase diagram of conceptual outcomes for the system yields further intuition of the relevance of policy action (Fig. 3-2). In the region where both policy and the disease itself are able to spread through the network, a more effective enforcement of the coordinated routing policy increases the invasion threshold, rendering a system that is more resistant against epidemic outbreaks (Fig. 3-2a). In contrast, even a broad-based adoption of self-initiated rerouting is unable to increase the invasion

threshold, implying that decreasing individual contagion risk may not decrease the societal risk (Fig. 3-2b). In the following sections we provide quantitative measures of these differences, through simulations on synthetic and realistic mobility networks.

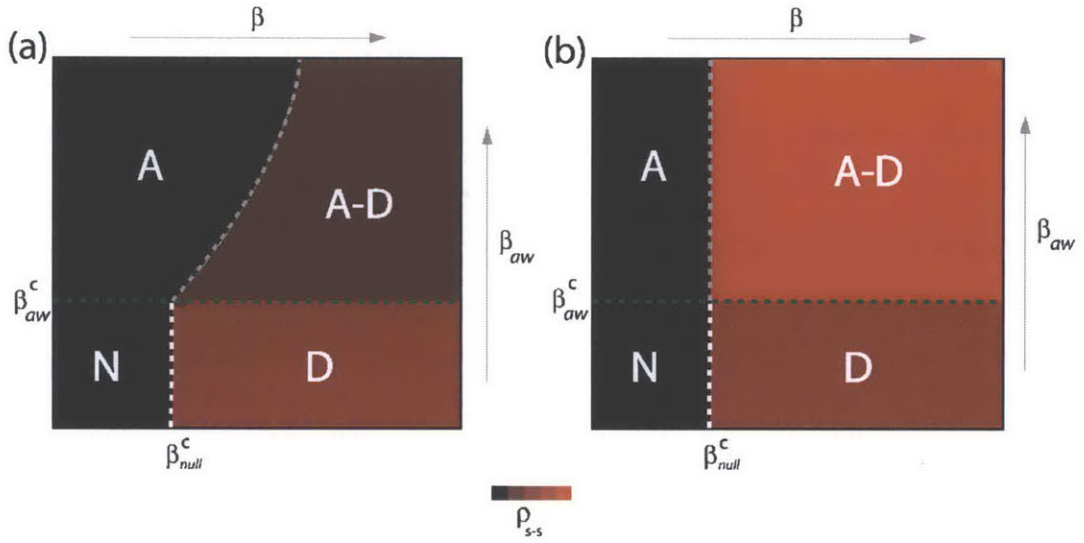

Figure 3-2: **Phase diagram of the coupled contagion processes at steady state.** The phase diagram for the prevalence of the two spreading processes in the case of coordinated (a) and selfish (b) awareness. The diagram is divided in four regions: (N) Neither disease spreading, nor awareness adoption, cause an outbreak in the system. (D) The prevalence of the disease causes an endemic state, while awareness dies out exponentially fast. (A) There is an endemic state of awareness in the system at equilibrium, while the disease dies out exponentially fast. (A-D) Both awareness and disease spread through the system and reach endemic states. The mean field assumption predicts that the invasion threshold changes in the presence of coordinated awareness in the system, but it remains unchanged in the case of selfish awareness. With black we denote the absorbing phase and with red the active phase for the disease spreading. The boundary curve between the region A and A-D in (a) represents the epidemic threshold condition. For a given network topology and a set of parameters, it is given by Eq. (2.6).

3.5 Numerical simulations

3.5.1 Monte Carlo simulations of conceptual model on synthetic networks

To investigate the intensity of the epidemic as a function of network topology and model parameters, and to discuss the price of anarchy when selfish routing is allowed, we perform Monte-Carlo simulations on synthetic uncorrelated, scale-free networks of N=5000 nodes.

We consider scale-free graphs generated by the uncorrelated configuration model (Catanzaro et al., 2005) with power law degree distribution $P(k) \sim k^{-\gamma}$ and $2 \leq k \leq \sqrt{N}$. A number of λN individuals are initially placed in the system randomly and uniformly. Individuals move through shortest paths to randomly chosen destinations with velocity of one node-to-node jump per time step. Once the mobility process reaches equilibrium, we infect randomly 1% of the nodes. We assume that these initially selected nodes are also aware of the disease. Individuals inside aware nodes are forced to travel through the system according to the coordinated or selfish routine strategies by minimizing the corresponding cost functions, taking into account the value of the degree of awareness ω. We implement both the SIS and SIR compartmental models. In the SIS model, a node can be either susceptible to the disease or infected. An infected node becomes healthy with a recovery rate μ. For simplicity and without loss of generality, we set $\mu = 1$. A susceptible node becomes infected with probability

$$P_{\text{inf}} = 1 - (1 - \beta\delta t)^{\nu_I}, \tag{3.10}$$

where β is the infection rate and ν_I is the number of individuals the node receives from an infected node in the time interval $(t, t+\delta t)$. In the SIR model, a node can be in three discrete states: susceptible to the disease, infected, or recovered/immune. An infected node recovers and becomes immunized with a recovery rate μ. At the same time, a susceptible node becomes infected with probability P_{inf}, as described above. Synchronously, we model the diffusion of awareness as an additional contagion process

in the system. An aware subpopulation forgets about the information with rate μ^{aw}, which we set equal to 1. On the other hand, a non-aware node adopts the information with probability

$$P_{\text{adp}}^{aw} = 1 - (1 - \beta^{aw}\delta t)^{k_{aw}}, \tag{3.11}$$

where β^{aw} is the rate of spreading of awareness, and k_{aw} is the number of aware neighbors. When the system reaches equilibrium, we compute the density of infected nodes, ρ.

We measure the density of infected nodes at equilibrium, averaged over 100 realizations, both for the case of coordinated action and selfish rerouting. We consider networks with two different levels of node-degree heterogeneity, $\gamma = 2$ and $\gamma = 3.3$.

We illustrate the importance of the spread of policy adoption by first considering that all nodes in the network are aware, and the degree of awareness ω is constant. In this particular case, the invasion threshold (3.9) becomes

$$\beta > \frac{\langle b_{\text{alg}} \rangle}{\langle b_{\text{alg}}^2 \rangle} \frac{\mu}{(1-\omega)\beta\lambda N}. \tag{3.12}$$

This prediction agrees nicely with our numerical simulations (Fig. 3-3), which show a mild dependence on network topology.

Figure 3-4 summarizes the main theoretical contributions of this study. We compare the analytical results derived using the HMF approach with Monte-Carlo simulations of the full coupled model, with either policy-driven (Fig. 3-4a) or selfish (Fig. 3-4b) rerouting. Our HMF theory accurately predicts the different transitions observed in the numerical simulations, confirming the conceptual phase diagram depicted in Fig. 3-2. These numerical simulations allow us to quantify the *price of anarchy* as the difference between the density of infected nodes at equilibrium for the coordinated and selfish strategies (Fig. 3-4c). The salient features of the behavior of this system arise from the strong nonlinearities induced by the coupling: while the selfish rerouting seems to reduce the intensity of the infection for mild diseases, it has a negative effect for more aggressive diseases, causing a larger fraction of infected subpopulations. The loss of welfare due to selfish rerouting, compared to the

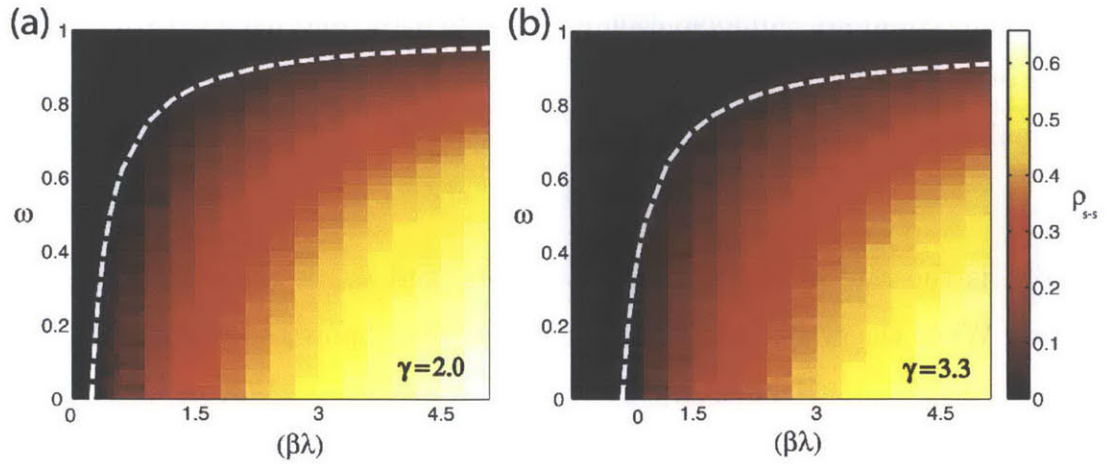

Figure 3-3: **Monte-Carlo simulations with global policy adoption**. We show the density of infected nodes at the steady state, as a function of the degree of awareness, ω, and the product of the infection rate by the traffic parameter $\beta\lambda$, for two scale-free networks with different level of heterogeneity. We use networks of size N=5000 nodes, and average the results over 100 realizations. The white dashed lines show the theoretical invasion threshold, calculated with Eq. 3.12.

policy-driven action, denotes the price of anarchy during disease spreading in mobility networks (Fig. 3-4c). The price of anarchy increases for more heterogeneous network substrates (smaller γ), and as the enforcement of policy increases (larger β^{aw}). It is also higher for more aggressive diseases (larger β).

To test our hypothesis about the use of the SIS contagion model, and the generality of the conclusions, we present Monte-Carlo simulations using the SIR compartmental model (Figure 3-5). The SIR model is more appropriate to describe Influenza-like diseases (eg. H1N1 (Fraser et al., 2009; Balcan et al., 2009)), and the severe acute respiratory syndrome (SARS) (Colizza et al., 2007). We first explore the early-time onset dynamics for both SIR and SIS models (Fig. 3-5a), and we conclude that at early times, time and spatial scales of both models are similar. Furthermore, in (Fig. 3-5b) we present the global attack of an SIR outbreak (density of subpopulations that experienced the infection) in a policy driven scenario as a function of the disease reaction rate and the adoption of awareness rate. It is clear that similar conclusions for the effect of a policy driven action can also be drawn in the case of an SIR infection model (see (Fig. 3-4a)).

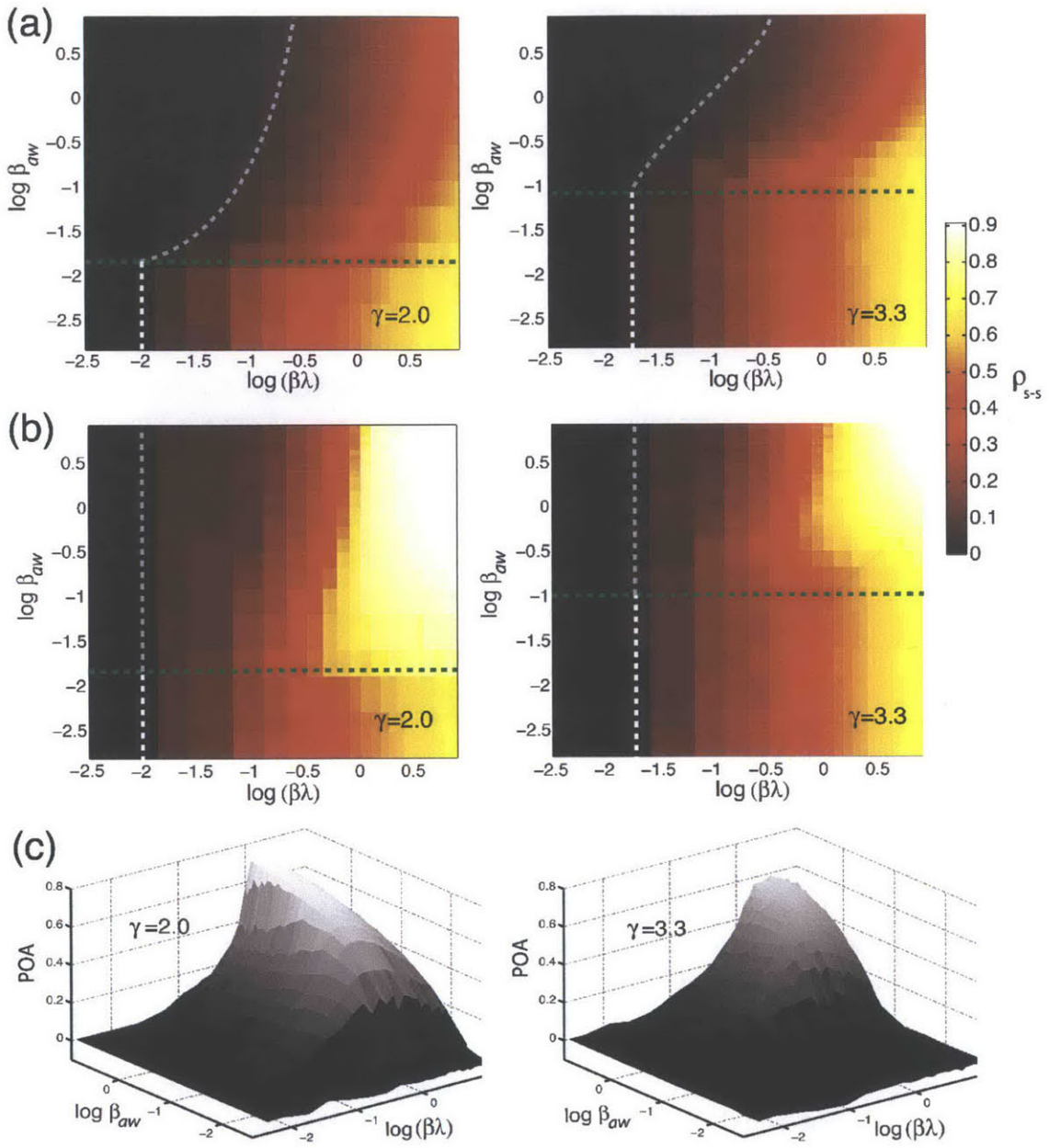

Figure 3-4: **Monte-Carlo simulations with spreading policy/awareness**. (a) Density of infected nodes the steady state, ρ, as a function of the product $\beta\lambda$ and the adoption of awareness rate β^{aw} that initiates *policy made* rerouting behavior, for two synthetic scale free networks with different degree exponents. (b) Density of infected nodes at steady state as a function of the infection rate and the adoption of awareness rate that initiates *selfish* rerouting behavior, for the same synthetic networks. Shown with dashed lines are the predictions of the mean field assumption for the phase diagram separation thresholds. (c) The price of anarchy as a function of the two reaction rates for the two synthetic networks. We use recovery rates, μ=μ^{aw}=1 and we set the degree of awareness $\omega = 0.8$ for both the coordinated and selfish rerouting strategies. The results are averaged over 100 realizations.

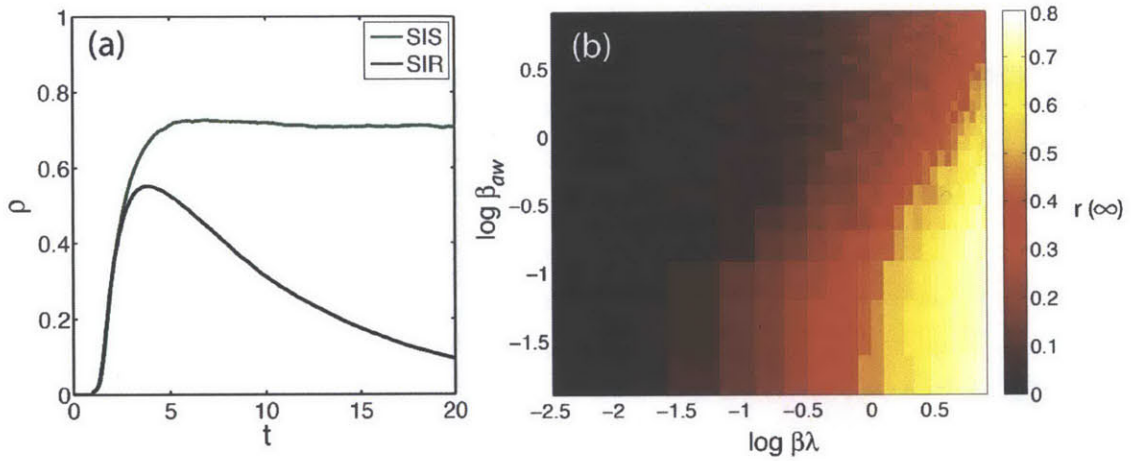

Figure 3-5: **Policy driven behavioral changes on an SIR epidemic model** (a) Time evolution of the density of infected subpopulations, under a policy driven behavior, for the SIS and SIR infection models. The infection rate is $\beta = 2.25$, the awareness adoption rate has value $\beta_{aw} = 0.25$, and we set the traffic parameter to $\lambda = 5$. The degree of awareness is equal to $\omega = 0.8$ and the recovery rates are equal to $\mu = 1/4$ and $\mu_{aw} = 1$. (b) The global attack ($r(t = \infty)$) as a function of the product $\beta\lambda$ and the adoption of awareness rate β^{aw} that initiates *policy driven* rerouting behavior. In both settings, the substrate network is uncorrelated, scale-free with degree exponent equal to $\gamma = 3.2$ and of size N=5000 nodes (subpopulations). The degree of awareness is equal to $\omega = 0.8$ and we use unit recovery rates. The results are averaged over 50 realizations.

3.5.2 Comparison between our conceptual model and a classical metapopulation model.

Using numerical simulations and simple theoretical arguments, we compare the predictions of our conceptual model with those of a more detailed metapopulation model, where we introduce behavioral changes. Since the parameters used by these two models are different, so it is not obvious in principle how these calculations should be compared. We put forward a rescaling argument that allows to quantify whether the two models yield the same behavior under equivalent infection rates and traffic density.

We consider a metapopulation network of size N and degree distribution $P(k)$, where each node i represents a subpopulation with V_i individuals. We set the population size proportional to the topological betweenness. We assume the mobility process in which each individual in a node i, with a probability p, travels to a destination node j that is randomly selected with probability proportional to its size V_j. Travelers reach their destinations following the shortest path. A convenient representation of the system is provided through quantities defined in terms of the degree k. Lets assume that a small set of initially infected subpopulations of degree k, $\{D_k^0\}$, is experiencing an outbreak with $R_0 = \beta_m/\mu_m > 1$, where β_m and μ_m are the infection and recovery rates respectively. In the early stage of the process, the number of subpopulations experiencing an outbreak is small and the disease spreading at the level of metapopulation system can be described as a branching process, using a tree-like approximation relating the infected subpopulations D_k^n at generation n to the infected subpopulations D_k^{n-1} at generation $n-1$ (Colizza and Vespignani, 2007; Balcan and Vespignani, 2011). The average number of infected individuals at equilibrium in a subpopulation k of population V_k is αV_k, where α is a disease dependent parameter equal to $\alpha = (R_0 - 1)/R_0$ (Barthélemy et al., 2010). Each infected individual stays in the infectious state for an average period μ_m^{-1}. Thus, the total number of infected

individuals circulating through the network at the $n-1$ generation is

$$w^{n-1} = (p\alpha/\mu_m) \sum_{k'} D_{k'}^{n-1} V_{k'}. \tag{3.13}$$

Those individuals can trigger the start of the epidemic in a susceptible subpopulation i with probability $[1 - R_0^{-\gamma_i^{n-1}}]$, where γ_i^{n-1} is the number of infectious individuals in generation $n-1$ that have visited the subpopulation. The number of infected individuals that will pass through a subpopulation of degree k will be proportional to the algorithmic betweenness $\gamma_k^{n-1} = b_{\text{alg}}^k w^{n-1}$. For the n^{th}-generation we have:

$$D_k^n = N_k(1 - D_k^{n-1}/N_k)[1 - (R_0^{\gamma_k^{n-1}})^{-1}], \tag{3.14}$$

where the second term on the right hand side is the probability that the subpopulation is not already seeded by infected individuals and the last term is the probability that the new seeded population will experience an outbreak. At the early times of the process and for $R_0 \sim 1$, equation (3.14) can be approximated by

$$D_k^n = (R_0 - 1)(p\alpha/\mu_m) N_k b_{\text{alg}}^k \sum_{k'} D_{k'}^{n-1} V_{k'}. \tag{3.15}$$

Considering at equilibrium $V_k = \bar{V} b_{\text{alg}}^k / \langle b_{\text{alg}} \rangle$, where $\bar{V}$ is the average population size and by defining $\Theta^n = \sum_k D_k^n b_{\text{alg}}^k$, we have that:

$$\Theta^n = (R_0 - 1)(p\alpha/\mu_m) \bar{V} N (\langle b_{\text{alg}}^2 \rangle / \langle b_{\text{alg}} \rangle) \Theta^{n-1}. \tag{3.16}$$

The above self-consistent equation defines the global invasion threshold:

$$\frac{(R_0 - 1)^2}{R_0} > \frac{\langle b_{\text{alg}} \rangle}{\langle b_{\text{alg}}^2 \rangle} \frac{\mu_m}{p \bar{V} N}. \tag{3.17}$$

This threshold condition unveils the influence of the model parameters, as well as the mobility patterns, on the spreading dynamics. The invasion threshold decreases with total traffic, and the condition $R_0 > 1$ for the global invasion is recovered for high

mobility rates. Furthermore, the threshold condition depends on the routing protocol, through the first and second moments of the algorithmic betweenness. Equations (3.17) and (3.7) suggest that the conceptual and metapopulation models should be compared by considering the traffic-weighted infection rates $\beta_c^*=\beta\lambda$ and $\beta_m^*=(R_0-1)^2/R_0p\bar{V}$, respectively.

We take this scaling relationship one step further, and use it to investigate whether the theoretical results derived for our conceptual model are also representative in the case of coordinated, policy driven behavior in a metapopulation model. Using the expression for the invasion threshold in the conceptual model, Eq. 3.9, and introducing the effective infection rate of the metapopulation model, we arrive at the invasion threshold:

$$\frac{(R_0-1)^2}{R_0} > \frac{\langle b_{\rm alg}\rangle}{\langle b_{\rm alg}^2\rangle - \omega\langle b_{\rm alg}^2\rangle_\omega}\frac{\mu_m}{p\bar{V}N} \tag{3.18}$$

To test this theory, we plot the intensity of the infection computed using the conceptual and metapopulation models for various parameter values, and rescale the axes according to the above effective infection rates (Fig. 3-6). Qualitatively, the two models seem to be equivalent in their predictions of the role of coordinated action on disease spreading. Their quantitative match is also remarkable, in particular the fact that, under the suggested rescaling, the invasion threshold derived for our conceptual model seems to capture the transition in the metapopulation model as well. We conclude that the two models are basically equivalent from a statistical viewpoint.

3.5.3 Data-driven simulations

We apply our methodology to a more realistic scenario where the substrate for the spreading processes is the commuting network of the United States. The nodes in the network represent the counties in the contiguous US, and the links between nodes represent daily commuting fluxes. We perform Monte-Carlo simulations, with commuters distributed among the different counties according to traffic. Epidemic spreading in the US commuting network is mainly driven by high traffic disorder, which leads to very heterogeneous spreading patterns. In Fig. 3-7 (inset), we show a sample re-

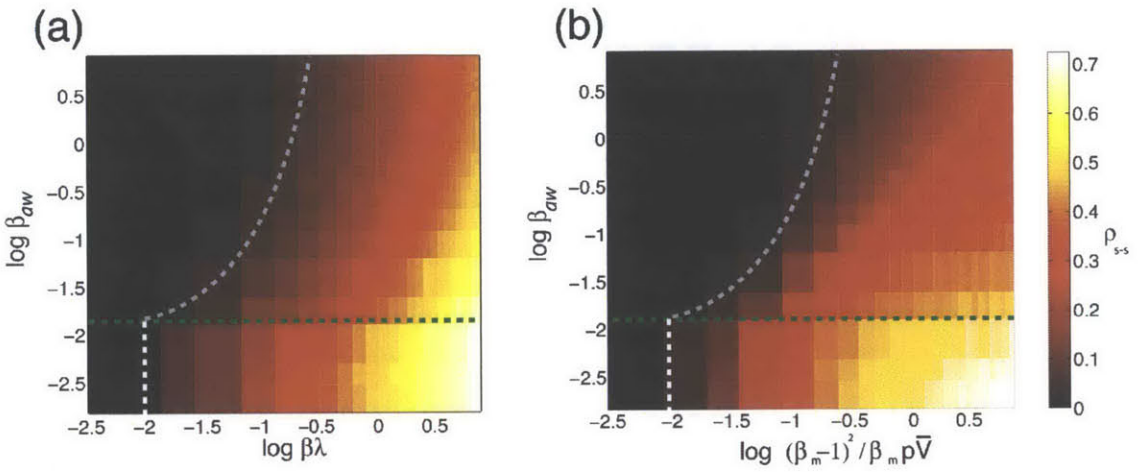

Figure 3-6: **Comparison between the *conceptual* and *metapopulation* models: Monte-Carlo simulations with spreading policy/awareness.** (a) *Conceptual model.* Density of infected nodes at the steady state, ρ, as a function of the product $\beta\lambda$ and the adoption of awareness rate β^{aw} that initiates *policy driven* rerouting behavior. Shown with dashed lines are the predictions of the mean field assumption for the phase diagram separation thresholds. (b) *Metapopulation model.* Density of infected subpopulations (subpopulations with $I_i/V_i > 1\%$) at steady state as a function of the quantity $p\bar{V}(\beta_m - 1)^2/\beta_m$ and the adoption of awareness rate β^{aw} that initiates *policy driven* rerouting behavior. The green dashed line is the threshold of the awareness spreading β_c^{aw}, the white line represents the threshold in Eq. 3.17, and the grey line is the a priori prediction of the invasion threshold in the presence of policy driven awareness in the metapopulation model (Eq. 3.18). The substrate network is uncorrelated, scale-free with degree exponent equal to $\gamma = 2$ and of size N=5000 nodes (subpopulations). The degree of awareness in the both cases is equal to $\omega = 0.8$ and all recovery rates are set equal to unit. For the metapopulation model we use a total number of $V = 3 \times 10^6$ individuals. The results are averaged over 50 realizations.

alization of our model, where we visualize the spreading patterns of infection and awareness, respectively, for both the selfish and coordinated rerouting strategies, two weeks after an outbreak at NYC county.

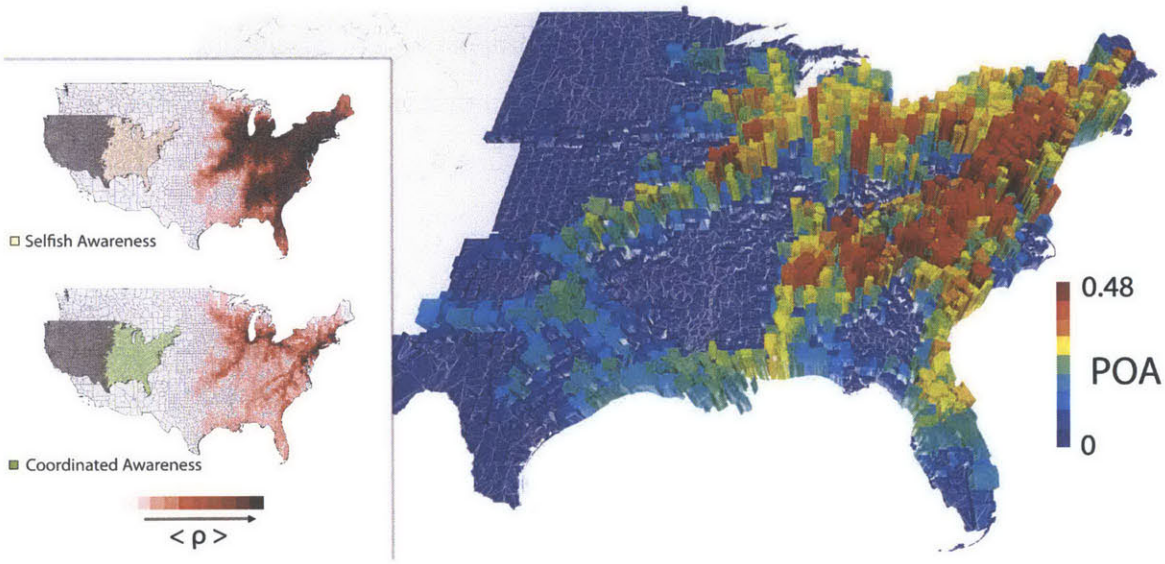

Figure 3-7: **Coupled information and epidemics in the US commuting network.** The price of anarchy, two weeks after an epidemic starts from each county in the East Coast of the United States. The results are averaged over 50 realizations. (Inset) A snapshot of the epidemic two weeks after it is initiated at the New York City county, in the presence of *selfish* and *coordinated* awareness. Dark red colors denote high densities of infection, $\langle \rho \rangle$. We average over 100 mobility and disease realizations. We use total traffic parameter $\lambda = 5$, reproductive numbers, R_0=2, R_0^{aw}=1.25, and recovery rates, μ=μ^{aw}=1 day^{-1}. We further assume that individuals spend 1/3 days (eight working hours) at the final destination before returning back home, and 1/24 days (one hour) at each of the intermediate counties. We set the degree of awareness to $\omega = 0.8$ for both the coordinated and selfish rerouting strategies.

The propagation of a disease depends strongly on the position of the initial seeding, due to traffic and topology heterogeneities (Nicolaides et al., 2012; Kitsak et al., 2010). The different system responses to an outbreak suggests that policy decisions need to account for the properties of the network as a whole, but also about the specific local transmission mechanisms and mobility patterns. We compute the price of anarchy for an infection starting at each one of the 2654 counties in the Eastern part of the contiguous US (Fig. 3-7). Due to traffic disorder we observe strong spatial heterogeneity in the price of anarchy values, depending on the origin of the outbreak.

Individuals living near the major interstate highways of the East Coast (I-80 from New York City to San Francisco, CA; I-85 from Petersburg, VA to Montgomery, AL; and I-95 from the Canadian Border to Miami, FL) have the incentive of a fast commute to neighboring counties. As a consequence, counties surrounding those commuting corridors have, in general, a larger proportion of commuters compared to other regions in the US. In contrast, the POA for epidemic spreading in areas of low commute flux is small, and therefore imposing policy-initiated action does not render substantial benefits for the containment of the epidemic.

Implementation. Data on commuting trips between counties in the United States is available online [1]. The files were compiled from Census 2000 responses to the long-form questions on where individuals worked. The files provide data at the county level for residents of the 50 states and the District of Columbia (DC). The data contain information on 34000 commuters in N=3141 counties. We build the non-symmetric traffic matrix F where its entry F_{ij} denotes the number of individuals traveling from county i to county j and, by considering only commuting flow up to 25 miles *outside the borders* of each county, we construct the immediate neighbor flux matrix F^{im}. We initialize the system by randomly placing $5N$ individuals in the system. The destination of an individual located at county i is chosen randomly among the set neighbor counties $\{j\}$ (i.e. $F_{ij} > 0$), with probability proportional to the flux F_{ij}. Individuals travel through the system following a "shortest path" to their destinations by maximizing the total traffic of the route segments, $\sum F^{im}_{kl}$. After this training period reaches equilibrium, we assign "home counties" to individuals, and we add the additional mobility rule of recurrent patterns. We make sure that the distribution of the population assigned after this initial training period correlates well with the actual populations from the census dataset. We then infect the county of consideration. We assume that this initially selected county is also aware of the disease. With probability ω, individuals inside aware nodes reroute from the "shortest path" to their destinations and follow either a policy-driven or a selfish path (Fig. 3-8). We implement the epidemic and awareness spreading models as described for the

[1] http://www.census.gov/population/www/cen2000/commuting/index.html

synthetic network simulations. We average our results over 50 model realizations.

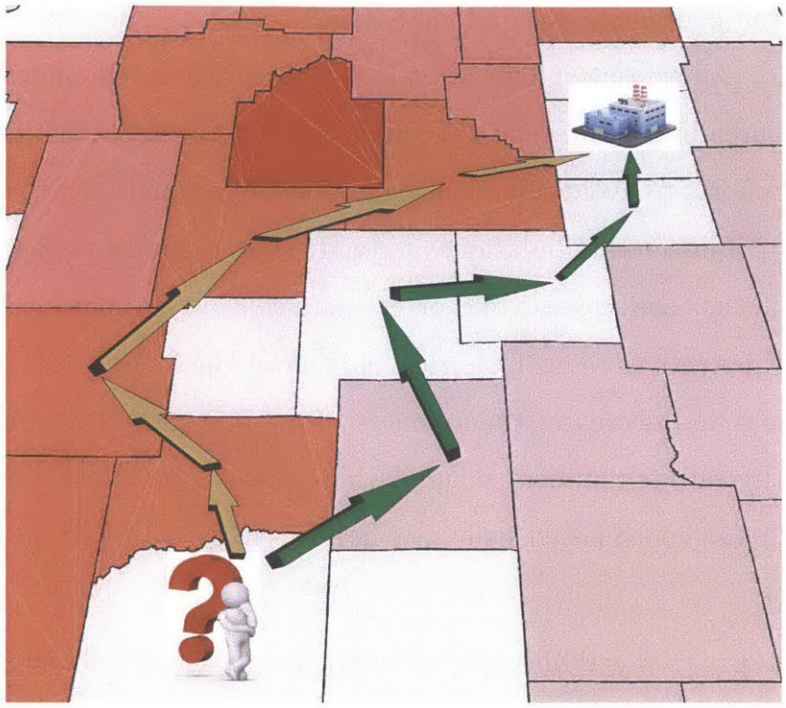

Figure 3-8: **The social dilemma for choosing the path to the destination in the US commuting network during an event of epidemic spreading.** An aware individual reroute from the "shortest path" to their destinations and follow either a policy-driven (green) or a selfish path (brown).

3.6 Discussion

Feedbacks between human behavior, policy action and the dynamics of contagion through mobility networks shape the footprint of infectious outbreaks, alter the disease transmission mechanisms, and determine the suitability of policies aimed at abating the epidemic by reducing the frequency of infectious contacts. In this study, we test the hypothesis that rerouting behavior, elicited by propagating awareness, can fundamentally change the patterns of disease spreading through mobility networks, both in terms of the invasion threshold and the total intensity of the outbreak. Our theoretical and simulation results support this hypothesis, and reveal a rich phase diagram of potential outcomes depending on the rerouting strategy and the dynamics of contagion and awareness.

Consistent with previous simulation studies on self-initiated behavioral responses in mobility-driven contagion, we find that selfish rerouting does not change the invasion threshold. As we report here, this is true even when the epidemic process is coupled to the spreading of awareness. From a policy perspective, this result suggests that individual efforts to avoid infectious contacts may not help the social welfare; in fact, the density of infected populations at steady state is higher in the case of selfish rerouting than in the base scenario of shortest-path routing. In contrast, policy-driven coordination, which strives to mitigate the epidemic at the societal level, increases the invasion threshold and decreases the intensity of the infection. The price of anarchy, which we quantify through numerical simulations, reveals the essential differences between the selfish and coordinated strategies in terms of their impact on the spreading of the epidemics.

These results pose a social dilemma, where policymakers and social agents need to find a balance between the pursuit of individual utility and the preservation of social welfare. In this study we show that the price of anarchy is related to the nature of the disease, the topology of the network substrate, and the resources deployed to enforce adoption of coordinated action. Such a quantitative analysis should provide valuable predictions to inform policy decisions about whether coordinated rerouting should be strongly enforced, or self-initiated behavior is allowed. Our map of the price of anarchy in the Eastern part of the United States illustrates this dilemma and suggests strategies for the deployment of measures intended to contain an outbreak. The price of anarchy to epidemic spreading exhibits strong heterogeneity, controlled by the proximity to major commuting corridors like the interstate highways. This categorical identification of population centers, ranked by their price of anarchy, may provide disease-control authorities with a priori information of the benefits of implementing mobility restrictions in the event of an outbreak.

In the next Chapter, we introduce and study extensively a fresh dynamical systems on a heterogeneous network topology that is able to self-organize input signals into localized and quantized patterns. We relate the characteristics of the pattern formation to both the topological properties of the network and to the nonlinear

structure of the underlying process, and discuss the implications of our findings in modeling cell assemblies of memories in brain activity.

Chapter 4

Self-organization and quantized states in neural activity

In this chapter, we introduce a simple model of excitation and inhibition on heterogeneous neural networks that can potentially explain the formation of cell assemblies in brain activity through self-organization. We demonstrate the properties of such neural network processes with a minimal-ingredients model of neuron activation and interaction within a complex network. The requirements are minimal and general: simple local dynamics based on typical activation potentials, and interactions between neurons that induce short-range anti correlation and long-range correlation in activity. The simplicity of our assumptions make our results very powerful. The formation of localized, robust neuron assemblies can be explained by simple neuron interaction without differential strengthening of the connections between neurons, the so called synaptic plasticity. We relate the characteristics of the model to both the topological properties of the network and the nonlinear structure of the underlying local process and we discuss applications in brain activity functional modeling.

4.1 Motivation

The mapping human brain function has emerged as one of the most fascinating challenges of the 21st century (Insel et al., 2013). Network theory is becoming a vitally

important ingredient in this endeavor (Bullmore and Sporns, 2009, 2012; Chavez et al., 2010; Sporns, 2011). The human brain' s capacity to process and store information arises from the collective behavior of 10^{11} neurons that form a highly heterogeneous interconnected network. Transmission of information between neurons occurs via the diffusion of electrical pulses along the wiring connecting different neurons. Therefore, understanding the topology and the dynamics of the brain from a network perspective, is a crucial step towards understanding the function of the human' s most complex organ.

The increasing amount of brain data has push forward our understanding of brain structure from two mainly points of views: structural and functional networks. Structural brain network, also known as "connectome" can be derived through *diffusion tensor imaging* (DTI) (Iturria-Medina et al., 2008) or *diffusion spectrum imaging* (DSI) (Hagmann et al., 2007) that both are based on diffusion of water through myelinated nerve fibers in the brain. On the other hand there is a plenty of methods through which neuroscientists derive functional network structure that based on the signal correlation between neurons or areas in brain. An important distinction must be made between these two kind of networks in terms of edge definition (Sporns et al., 2004; Bullmore and Sporns, 2009; Sporns, 2013). Edges in structural networks refer to aspects of the physical infrastructure of brain connectivity, that is, synaptic connections jointly comprising the wiring diagram, whereas edges in functional networks reflect aspects of statistical dependencies among neuronal time series corresponding to simple correlation's or covariance or more sophisticated measures of nonlinear coupling or casual dependence. These differences in edge definition entails clear differences in the way structural and functional networks should be analyzed, interpreted and used. For example information communication processes, governing the propagation of information along the connections do not occur in functional networks. Instead functional networks are manifestation of these processes as they unfold in the structural connectome. Therefore studies aiming to identify important features of communication processes in brain are more appropriately carry out on structural than on functional networks (Sporns, 2013).

The sight or the sound of a familiar concept triggers a cascade of brain processes that creates a representation leading to the recognition of the concept, the recollection of details related to it and the generation of new memories. For decades neuroscientists have debated how memories are stored. This debate continues today with two extreme theories. The first one suggests that single neurons hold the recollection, say, of the grandmother or of a famous movie star. The second and dominant until recently view suggests that the perception of any specific individual or object is accomplished by the collective activity of many millions if not billions of nerve cells, what Nobel laureate Charles Sherrington use to call "a millionfold democracy". However, recent experiments suggests that the reality lies somewhere between these two theories (Quiroga et al., 2005). That is, a particular concept triggers the firing of no more than a million and some times some thousands of neurons out of about more than a billion in the medial temporal lobe (MTL) of the brain, known as *cell assemblies* (Quiroga et al., 2005; Quiroga, 2012; Quiroga et al., 2008) (Fig. 4-1A).

4.2 Cell Assemblies

Cell assemblies, or small subsets of neurons that fire synchronously, are the functional unit of the cerebral cortex in the Hebbian theory of mental representation (Palm, 1981; Kelso et al., 1986; Lansner et al., 2002; Dudai, 2004; Buzsáki, 2010). The connection between localized activation patterns and specific memories or perceptions is well documented (Quiroga et al., 2005; Quiroga, 2012), and associative learning forms the basis of our current understanding of the brain's structure and function (Reijmers et al., 2007; Neves et al., 2008; Lansner, 2009). The role of localized neuron assemblies as functional building blocks for mental process such as memory, learning and perception, raises the key question of how these groups of neurons become associated. Conceptually, cell assemblies are attractors, in the sense that the activation of some of these neurons triggers a cascade that results in the persistent firing of the whole group (Lansner et al., 2002; Tsodyks, 2005).

Finding the underlying mechanisms behind the creation of these attracting lo-

cal structures is a long-standing problem in neuroscience (Kelso et al., 1986; Fuster, 2000; Han et al., 2007; Silva et al., 2009). Hebb proposed synaptic plasticity as the basic mechanism for neural assembly formation: the repeated, contingent, activation of a group of neurons would strengthen their neural connections, making them more sensitive to each other's activation and increasing the likelihood of synchronous firing (Kelso et al., 1986). Hence, in the Hebbian framework learning is mainly the physiological process of strengthening synaptic connections among groups of neurons (Fuster, 2000; Martin and Morris, 2002). A weaknesses of synaptic plasticity as a general mechanism for neural assembly is that the time scales for neural association and learning must be related to the time needed for the strengthening of synaptic connections. In fact, the times are often much faster, indicating that in some cases association may be a consequence of self-organization rather than synaptic plasticity. Another conceptual difficulty of the Hebbian postulate of synaptic plasticity is that it does not provide a basis for encoding and computation in the brain. The fundamental mechanisms for information processing—the basic elements of the brain's internal language—remain unknown (Gallistel and Matzel, 2013).

Here we show that simple mechanisms of neural interaction allow for robust cell assembly creation through self-organization, without the need for reinforcement and synaptic plasticity. We demonstrate the properties of such neural network processes with a minimal-ingredients model of neuron activation and interaction within a complex network. The requirements are minimal and general: simple local dynamics based on typical activation potentials, and interactions between neurons that induce short-range anti correlation and long-range correlation in activity. The simplicity of our assumptions make our results very powerful. The formation of localized, robust neuron assemblies can be explained by simple neuron interaction without differential strengthening of the connections between neurons. Our model is compatible with Hebbian learning as well. Selective strengthening of certain synaptic connections through synaptic plasticity would also lead to the association of certain groups of neurons.

Our results suggest a new mechanisms for the formation of cell assemblies in neural

networks, and point to two mechanisms for learning and representation in the brain: Hebbian synaptic plasticity, and self-organization. We conceptualize these correlations as the interaction of neurons with first and second neighbors in the network. The proposed mechanism is compatible with Hebbian learning. In fact, the robustness of the localized activation patterns provides a powerful feedback reinforcing the action of synaptic plasticity. Our model also indicates a mechanism for learning based on self-organization. The conceptual advantage of self-organization in the brain is that the formation of cell assemblies is much faster that when synaptic plasticity is needed as an explanation.

An important question that will be addressed in the future is whether localized, self-organized network structures can provide better algorithms for machine learning and artificial intelligence. Because of their robustness and localization, self-organized structures may provide an encoding mechanism for information processing and computation in the brain. They may form the basic elements of the brain's internal language.

4.3 Mathematical Model

The complex neuron dynamics, and the interactions between neurons, are often conceptualized as dynamical process in network theory (Bullmore and Sporns, 2009; Chialvo, 2010; Bullmore and Sporns, 2012). The objective is to identify minimal-ingredients computational models of the functional and structural organization of the brain. The ultimate goal is to simulate the processes that enable the human brain to perform its tasks related to perception, memory and computation. Such minimal models of neural activity comprise a series of neuron compartments, with local dynamics and interactions between neurons. Computational models can be highly detailed, mimicking the various neuron functional types and the brain's structural and physiological organization (Izhikevich and Edelman*, 2008). Here we propose a minimal-ingredients, phenomenological model of neural excitation, inhibition and interaction a network with heterogeneous connectivity. Our goal is not to provide

a physiologically detailed representation of the human brain, but rather to demonstrate that a simple combination of local excitation and relaxation of individual units, and generic excitatory/inhibitory interactions between connected units, leads to self-organization, and explains the spontaneous formation of cell assemblies without the need for synaptic plasticity or reinforcement.

We conceptualize our neural networks as large sets of interconnected relaxation oscillators. A neuron's membrane potential is driven by local excitation-relaxation dynamics, in the spirit of the Hodgkin or FitzHugh-Nagumo models (Hodgkin and Huxley, 1952; FitzHugh, 1961), and by the interaction with other neurons in the network via exchanges through the links, or *synapses*, connecting them. Neuron activation is thus a consequence of ionic fluxes that cause changes in chemical potential across the neuron membrane—*neuron spiking* (Hodgkin and Huxley, 1952). In dimensionless quantities, the proposed excitation-inhibition model for the evolution of membrane potential, u_i, and resting variable w_i, in each neuron $i = 1, ..., N$ is given by:

$$\frac{du_i}{dt} = f(u_i, \mu) - w_i + I_i \tag{4.1}$$

$$\frac{dw_i}{dt} = C(u_i - w_i) \tag{4.2}$$

where potentials, u_i, are referred to the resting potential, which becomes the zero-level in our model. Locally, the membrane potential behaves as a relaxation oscillator, where

$$f(u_i, \mu) = -(1 + \mu)u_i + 1.5u_i^2 - u_i^3 \tag{4.3}$$

is a dynamic forcing term (Fig. 4-1 B), C is a decay coefficient, and μ acts as a bifurcation parameter. Synaptic currents are represented by the interactions with other neurons in the network. The topology of the network, with N nodes, is defined by the adjacency matrix, A, whose elements, A_{ij}, take values of 1, if nodes i and j $(i, j = 1, ..., N)$ are connected, and 0 otherwise. The degree—or connectivity—of a node i is therefore given by $k_i = \sum_{j=1}^{N} A_{ij}$.

The structure of the interactions between neurons is one of the key ingredients of our model. Rather than a simple, diffusive type process, that induces short-range correlation, we propose a two-level interaction structure that induces anti correlation in the short range (nearest-neighbors, or first-order connectivity), and long-range correlation (second-nearest neighbors, or second-order connectivity) (Fig. 4-1 C). These interactions can be simply achieved by assuming that neurons interact in a dual network. Neurons interact with first neighbors in one of the networks. In the second network, their networks are the second-order neighbors in the first networks. Mathematically, we express the integration of synaptic contributions as

$$I_i = -2\sum_{j=1}^{N} L_{ij}^{(2)} u_j - \sum_{j=1}^{N} L_{ij}^{(4)} u_j. \tag{4.4}$$

The simplest form for the interaction matrices representing these correlation/anti-correlation effects while ensuring conservation of charge is based on network representation of laplacian and bilaplacian operators. The network-laplacian matrix is (Nakao and Mikhailov, 2010)

$$L_{ij}^{(2)} = A_{ij} - k_i \delta_{ij}, \tag{4.5}$$

while we adopt the network-bilaplacian matrix

$$L_{ij}^{(4)} = (A^2)_{ij} - (k_i + k_j)A_{ij} + k_i^2 \delta_{ij} \simeq \sum_{l=1}^{N} L_{il}^{(2)} L_{lj}^{(2)}, \tag{4.6}$$

where the $(A^2)_{ij}$ matrix has information about second order nodal connectivity and takes nonzero values if node i is two jumps away from node j. The eigenvalues Λ_α and eigenvectors $\mathbf{\Phi}^\alpha = (\phi_1^{(\alpha)}, .., \phi_N^{(\alpha)})$ of the Laplacian matrix $L_{ij}^{(2)}$ are determined by

$$\sum_{j=1}^{N} L_{ij}^{(2)} \phi_j^{(\alpha)} = \Lambda_\alpha \phi_i^{(\alpha)}, \tag{4.7}$$

with $\alpha = 1, .., N$. The bilaplacian, $L_{ij}^{(4)}$, has the same eigenvectors as $L_{ij}^{(2)}$ (i.e. $\mathbf{\Phi}^\alpha$)

and its eigenvalues are the square of those of $L_{ij}^{(2)}$, Λ_α^2.

4.4 Connection to the Swift-Hohenberg equation

Equations (4.1) -(4.4), modeling the evolution of neuron membrane potential, can be interpreted as a network representation of the Swift-Hohenberg equation (Swift and Hohenberg, 1977),

$$\partial_t u = f(u, \mu) - \partial_{xx} u - \partial_{xxxx} u. \tag{4.8}$$

This equation is a prototype model for pattern formation in numerous natural systems. Its fascinating properties have been extensively studied recently (Cross and Hohenberg, 1993; Tlidi et al., 1994; Burke and Knobloch, 2006; Lloyd et al., 2010; Gomez and Nogueira, 2012). One the most intriguing features of solutions to the Swift-Hohenberg equation is the possibility of localized steady-state patterns (Fig. 4-1D). In two dimensions, solutions display characteristic hexagonal patterns (Burke and Knobloch, 2006; Lloyd et al., 2010). Our network model is more complex than the continuum Swift-Hohenberg model, due to the heterogeneity in network topology and connectivity, with is absent in the continuum case. Furthermore, our model has a decay term, implying that any activation patterns return to the resting potential. The question still remains of whether some of the properties of the Swift-Hohenberg properties are still inherited by our model.

4.5 Stability Analysis

To understand the properties of our neural network model, we begin with an investigation of the stability of the resting potential state, that is, the case in which all neurons are at the zero-level potential. The physiological and functional relevance of this analysis stems from the fact that, if the resting state is an unstable one, small random perturbations could trigger a large-scale, cascade of spontaneous firing of neurons in the brain. Under healthy operating conditions, the resting potential must be stable. Since neuron activation relaxes towards the resting potential in the absence

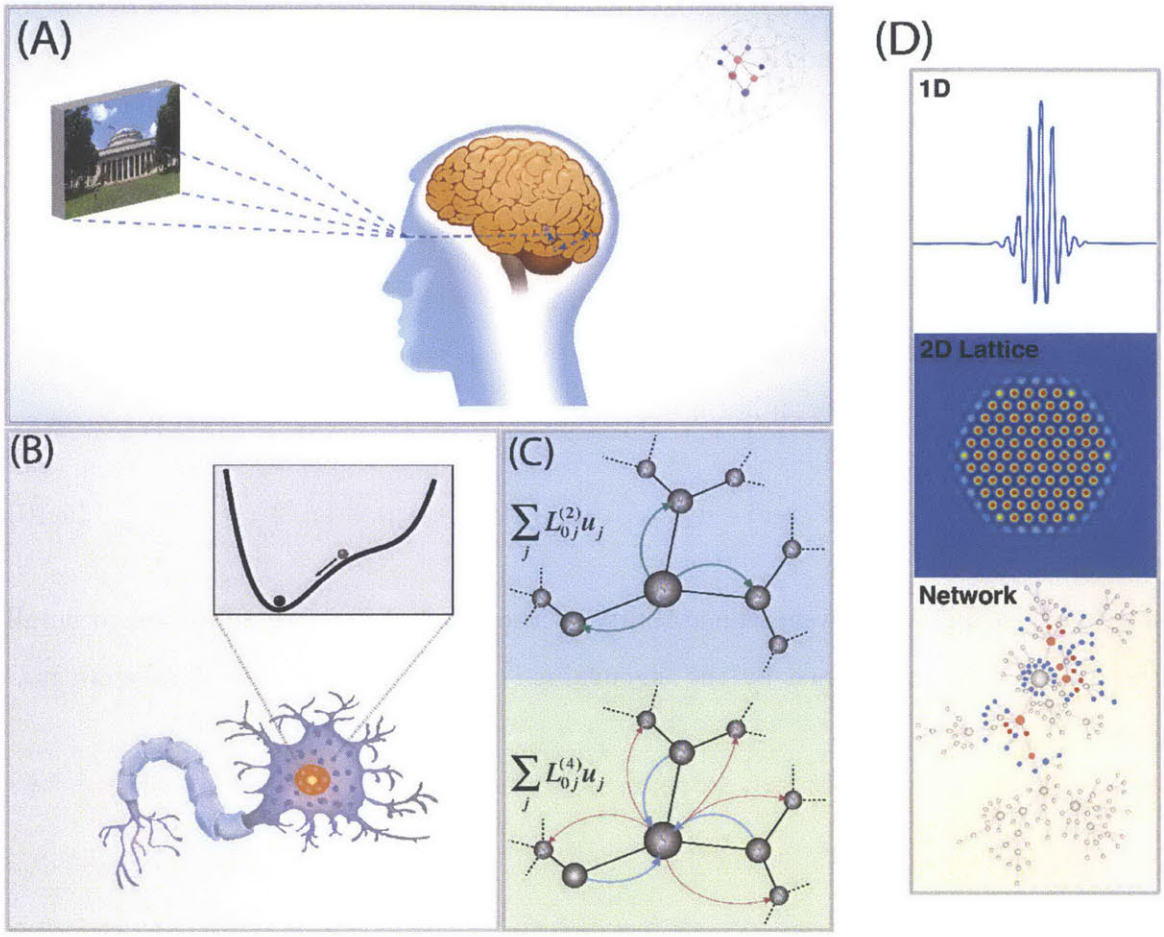

Figure 4-1: **Motivation and pictorial illustration of our dynamical model in neural networks.** **(A)** The sight of a familiar concept triggers a cascade of brain processes that creates a representation leading to the recognition of the concept through the firing of a finite number of neurons in the brain (cell assemblies). **(B)** Pictorial view of the local dynamics that neurons undergo. Locally, the membrane potential behaves as a relaxation oscillator, where $f(u, \mu)$ is a dynamic forcing term. In the inset we show the energy landscape, that behaves as a single well (at $u = 0$) with an inflection point, necessary condition for localized pattern formation. **(C)** Neurons interact with each other in the systems through diffusively transported species via the links (synapses) connecting them. The network–laplacian operator $L^{(2)}$ represents short range diffusion of the species in the system (top). The network–bilaplacian operator $L^{(4)}$ induces short range anti correlation with the nearest-neighbors, and long-range correlation with the second-nearest neighbors (bottom). **(D)** Typical localized patterns in 1D, 2D and in network topology. Quantized states are localized in the sense that they do not span the whole domain. In 1D the localized patterns is in the form of packets of waves restricted in a small area of the domain, where in 2D the localized patterns have hexagonal configuration. In network topology localized patterns usually span a neighborhood, or a community.

of input signals, our analysis refers to the short-time patterns of activation driven by the equation for the potential (4.1):

$$\frac{du_i}{dt} = f(u_i, \mu) - 2\sum_{j=1}^{N} L_{ij}^{(2)} u_j - \sum_{j=1}^{N} L_{ij}^{(4)} u_j, \qquad i = 1, ..., N. \tag{4.9}$$

Flat, stationary solutions to this equation satisfy $f(u_i) = 0$, where the membrane potential is equal for all nodes, $u_i = \bar{u}, \forall i = 1, ..., N$. For $f(u_i, \mu) = -(1+\mu)u_i + 1.5u_i^2 - u_i^3$, there are three uniform solution branches given by

$$u_0 = 0 \quad \text{and} \quad u_\pm = [1.5 \pm \sqrt{1.5^2 - 4(\mu+1)}]/2. \tag{4.10}$$

In a linear stability analysis, the stability of these flat stationary solutions to small perturbations is determined by the eigenvalues of the laplacian and bilaplacian matrices.

4.5.1 Linear stability analysis of flat solutions in time

In a linear stability analysis, the stability of these flat stationary solutions to small perturbations is determined by the eigenvalues of the laplacian and bilaplacian matrices. Introducing small perturbations, δu_i, to the uniform state $\bar{u}$, $u_i = \bar{u} + \delta u_i$, the linearized version of Eq. 4.9 takes the form

$$d\delta u_i/dt = f_u(\bar{u})\delta u_i - 2\sum_{j=1}^{N} L_{ij}^{(2)} \delta u_j - \sum_{j=1}^{N} L_{ij}^{(4)} \delta u_j. \tag{4.11}$$

By expanding the perturbation δu_i over the set of the laplacian eigenvectors,

$$\delta u_i = \sum_{\alpha=1}^{N} q_\alpha e^{\lambda_\alpha t} \phi_i^{(\alpha)}, \tag{4.12}$$

the linearized equation is transformed into a set of N independent linear equations for the different normal modes:

$$\lambda_\alpha = f_u(\bar{u}) - 2\Lambda_\alpha - \Lambda_\alpha^2, \quad \alpha = 1, \ldots, N, \tag{4.13}$$

where Λ_α are the eigenvalues of the laplacian matrix. The α-mode is unstable when Re λ_α is positive. Instability occurs when one of the modes (the critical mode) begins to grow. At the instability threshold, Re $\lambda_\alpha = 0$ for some α_c and Re $\lambda < 0$ for all other modes. In Fig. 4-2 we summarize the linear stability analysis of the flat states of our model on a scale-free network that we construct using the Barabási-Albert model (BA) (Barabási and Albert, 1999).

4.6 Localization: quantized response and robustness

4.6.1 Stability analysis of resting state in network topology

Localized activation patterns are possible due to the particular structure of the model, with short- and long-range interactions. Mathematically, the localized states are homoclinic orbits around the base resting state, $\bar{u} = u_0 = 0$. The existence of these homoclinic orbits can be studied using the technology developed for the linear stability analysis. Since homoclinic orbits leave the flat state as we approach a small neighborhood (cluster) of the network, the fixed point must have both stable and unstable eigenvalues. We linearize Eq. 4.9 around $u_0 = 0$, and expand the perturbations over the set of the laplacian eigenvectors $\delta u_i = \sum_{\alpha=1}^{N} q_\alpha \phi_i^{(\alpha)}$, $q_\alpha \ll 1$, arriving at the relation $f_u(0) - 2\Lambda_\alpha - \Lambda_\alpha^2 = 0$. Since the laplacian eigenvalues Λ_α, are real and non-positive, we can write them in the form $\Lambda_\alpha = -k_\alpha^2$. If $\mu > 0$, the topological eigenvalues at $u_0 = 0$ form a complex quartet, $k_\alpha = \pm i \pm \frac{\sqrt{\mu}}{2} + \mathcal{O}(\mu)$. For $\mu = 0$ they collide pairwise on the imaginary axis, and for $\mu < 0$ they split and remain on the imaginary axis, $k_\alpha = i(\pm 1 \pm \frac{\sqrt{-\mu}}{2}) + \mathcal{O}(\mu)$. For $\mu = -1$, two of the topological

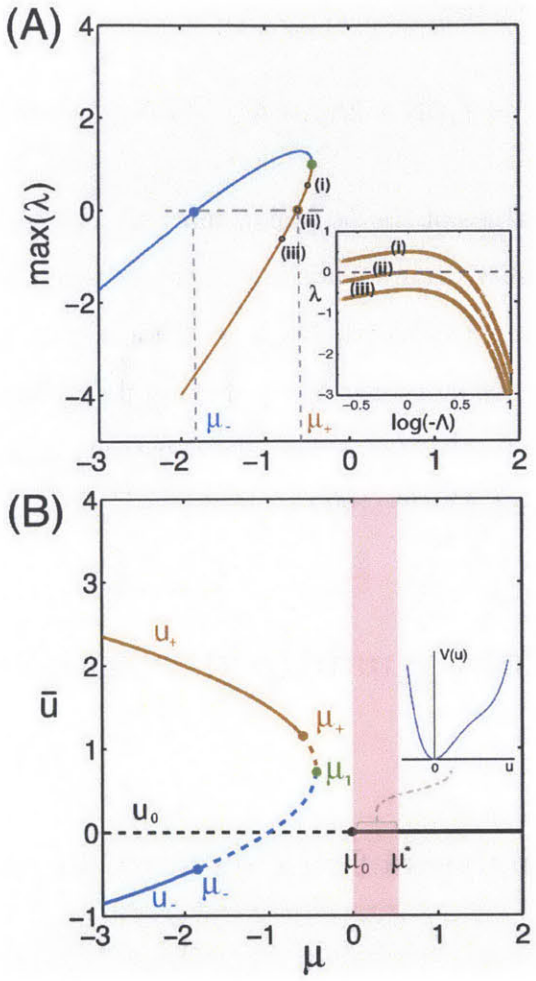

Figure 4-2: **Linear stability analysis of the flat stationary solutions of our model.** **(A)** maximum value of the growth rate λ as a function of the bifurcation parameter μ for the two flat stationary states u_+ (yellow) and u_- (blue) on a B-A network model with mean degree $\langle k \rangle = 3$ and size $N = 2000$. When the maximum value of λ is negative, the state is stable with respect to small non uniform perturbation. (Inset) The growth rate λ as a function of the Laplacian eigenvalue Λ (Eq. 4.13) for three different values of the bifurcation parameter μ as they indicated in the main diagram for the flat stationary solution u_-. **(B)** The flat stationary solutions u_0 and $u_\pm$ as a function of μ on the same network. Solid (dotted) lines represent stability (instability) with respect to small non-uniform perturbations. The labelled bifurcation points are $\mu_0 = 0$, $\mu_1 = -0.44$ and $\mu_+ = -0.62$ and $\mu_- = -1.82$. The pink shaded region is where we observe localized self-organization patterns with respect to the trivial solution u_0. For values of μ outside that region we get either global activation patterns (for $\mu < \mu_0$) or any perturbation relaxes back to the flat stationary solution (for $\mu > \mu^*$).

eigenvalues collide at the origin, while for $\mu < -1$ they move onto the real axis. These results are summarized in the Figure 4-4A. The topological eigenvalues in the neighborhood of $\mu = 0$ is an indication for a variety of topologically localized states for μ non-negative and close to zero (Burke and Knobloch, 2006)

4.6.2 Localized patterns by direct simulations

The existence of localized structures has important implications for the response of the network to a stimulus. When a neuron or group of neurons in the network receive an input stimulus, the activation will propagate within a certain neighborhood around the stimulated neurons, while the rest of the network will remain at the resting state. A neuron assembly will thus be spontaneously generated. The pattern of activation decays over time due to the relaxation term, but the short-term response is dominated by the localized firing pattern. To understand the onset of localized patterns for different model parameters and input stimuli, we construct the snaking bifurcation diagram of the resting state, as a function of the total potential energy of the stimulus and bifurcation parameter μ, in the vicinity of $\mu \simeq 0$. An initial stimulus is applied at the most central neighborhood in the network, defined as the most central node (using eigenvalue centrality) along with its first- and second-step neighbors.

In more details, we are interested in identifying a localized, steady state pattern of the dynamical system described by Eq. 4.9. We first choose a value for the bifurcation parameter μ^0 in the phase space where localized patterns are possible, i.e. $\mu^0 \epsilon (0, 0.6)$. We then initialize the activation potential profile on the network. In more details, we set $u_i = 0$ everywhere except at a small number of nodes in a randomly chosen neighborhood of the network, where we set $u_i = \hat{u} > 0$. Starting with this initial condition, we then advance the dynamical system in time, using implicit Runge-Kutta method. After a short time that depends on the time integration scheme and time step used, the dynamical model relaxes to a steady state $\mathbf{u}^0$ that is localized in nature, i.e. the activation area spans a small only fraction of the network (Fig. 4-3C). By computing the energy of that state $||\mathbf{u}^0||_{L2} = (1/N \sum_{j=1}^{N} u_j^0)^{1/2}$ we are able to draw a single point ($||\mathbf{u}^0||_{L2}$, μ^0) on the bifurcation diagram (Fig. 4-3B).

4.6.3 Numerical continuation

The full bifurcation diagram cannot be obtained by time-marching procedures. Instead, we use a pseudo-arclength continuation method to find a single branch of localized steady states in the bifurcation diagram.

In more details, we are interested in finding branches of localized states at the bifurcation diagram. These states are solutions of the time independent version of the system of equations (Eq. 4.9):

$$0 = f(u_i, \mu) - 2\sum_{j=1}^{N} L_{ij}^{(2)} u_j - \sum_{j=1}^{N} L_{ij}^{(4)} u_j, \quad i = 1..N. \tag{4.14}$$

We can rewrite the above equation of interest in vector form:

$$\mathbf{G}(\mathbf{u}, \mu) = -(1+\mu)\mathbf{u} + 1.5(\mathbf{u} \circ \mathbf{u}) - (\mathbf{u} \circ \mathbf{u} \circ \mathbf{u}) - 2\mathbf{L^{(2)}u} - \mathbf{L^{(4)}u} = \mathbf{0}, \tag{4.15}$$

where $\mathbf{u} = \{u_i\}$ is the vector of the main variable, μ is the bifurcation parameter, $\mathbf{L^{(2)}}$ and $\mathbf{L^{(4)}}$ are the laplacian and bilaplacian operators respectively. With -o-, we denote the element-wise product. Lets assume that initially a solution $(\mathbf{u}^0, \mu^0)$ of the above equation and an initial guess of the tangent vector $(\dot{\mathbf{u}}^0, \dot{\mu}^0)$ of the bifurcation branch at that point in the bifurcation space are given. Starting from that initial point and direction in the bifurcation diagram, we are interested in finding a branch of solutions using a numerical continuation algorithm. In order to allow for the solution branch to past a fold we use the Keller' s pseudo-arclength continuation of the initial solution (Keller, 1977). The main idea in pseudo-arclength continuation is to drop the natural parametrization by μ and use some other parameterization. Pseudo-arcrength continuation solves the following equations for $(\mathbf{u}^1, \mu^1)$:

$$\begin{aligned} \mathbf{G}(\mathbf{u}^1, \mu^1) &= \mathbf{0} \\ (\mathbf{u}^1 - \mathbf{u}^0)^T \dot{\mathbf{u}}^0 + (\mu^1 - \mu^0)\dot{\mu}^0 - \delta s &= 0, \end{aligned} \tag{4.16}$$

where the second equation is the plane, which is perpendicular to the tangent $(\dot{\mathbf{u}}^0, \dot{\mu}^0)$

at a distance δs from $(\mathbf{u}^0, \mu^0)$ (see Fig. 4-3).

Using Newton' s method this leads to the linear system:

$$\begin{pmatrix} \mathbf{G_u}(\mathbf{u}^{1(p)}, \mu^{1(p)}) & \mathbf{G}_\mu(\mathbf{u}^{1(p)}, \mu^{1(p)}) \\ (\dot{\mathbf{u}}^0)^T & \dot{\mu}^0 \end{pmatrix} \begin{pmatrix} (\Delta \mathbf{u}^1)^{(p)} \\ (\Delta \mu^1)^{(p)} \end{pmatrix} = - \begin{pmatrix} \mathbf{G}(\mathbf{u}^{1(p)}, \mu^{1(p)}) \\ (\mathbf{u}^1 - \mathbf{u}^0)^T \dot{\mathbf{u}}^0 + (\mu^{1(p)} - \mu^0)\dot{\mu}^0 - \delta s \end{pmatrix}, \tag{4.17}$$

$$\begin{aligned} (\mathbf{u}^1)^{(p+1)} &= (\mathbf{u}^1)^{(p)} + (\Delta \mathbf{u}^1)^{(p)} \\ (\mu^1)^{(p+1)} &= (\mu^1)^{(p)} + (\Delta \mu^1)^{(p)}, \end{aligned} \tag{4.18}$$

for p=0,1,2,..., where $\mathbf{G_u} = -(1+\mu)\mathbf{I} + 2 \times 1.5\mathbf{u} - 3(\mathbf{u} \circ \mathbf{u}) - 2\mathbf{L^{(2)}} - \mathbf{L^{(4)}}$. The new tangent direction vector is then computed by solving:

$$\begin{pmatrix} \mathbf{G}^1_\mathbf{u} & \mathbf{G}^1_\mu \\ (\dot{\mathbf{u}}^0)^T & \dot{\mu}^0 \end{pmatrix} \begin{pmatrix} \dot{\mathbf{u}}^1 \\ \dot{\mu}^1 \end{pmatrix} = \begin{pmatrix} \mathbf{0} \\ 1 \end{pmatrix}. \tag{4.19}$$

Note that in practice the system of Eq. 4.19, can be computed with one extra back-substitution and the orientation of the branch is preserved if the arclength δs is sufficiently small. The direction vector must be rescaled, so that indeed $||\dot{\mathbf{u}}^1||^2 + (\dot{\mu}^1)^2 = 1$.

Using this continuation method we are able to construct single *snaking* branches of localized states like the one shown in(Fig.4-4B). These branches show a characteristic "snaking" structure of localized states with varying activation energy $||\mathbf{u}^0||_{L2} = (1/N \sum_{j=1}^{N} u_j^0)^{1/2}$ (Fig. 4-4B). As the system jumps from one steady state branch to the next one, a new neighborhood in the network is being activated. Figure 4-4C visualizes the different steady localized states of the six different branches as they are spotted in the diagram of Figure 4-4B.

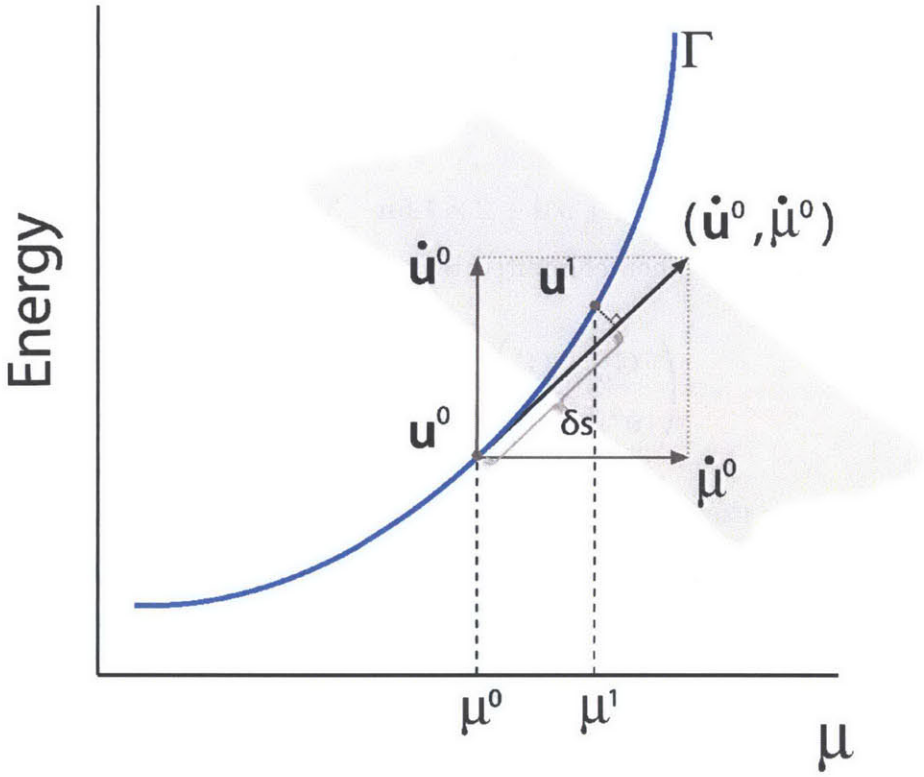

Figure 4-3: **Graphical interpretation of pseudo-arclength continuation on the bifurcation diagram.**

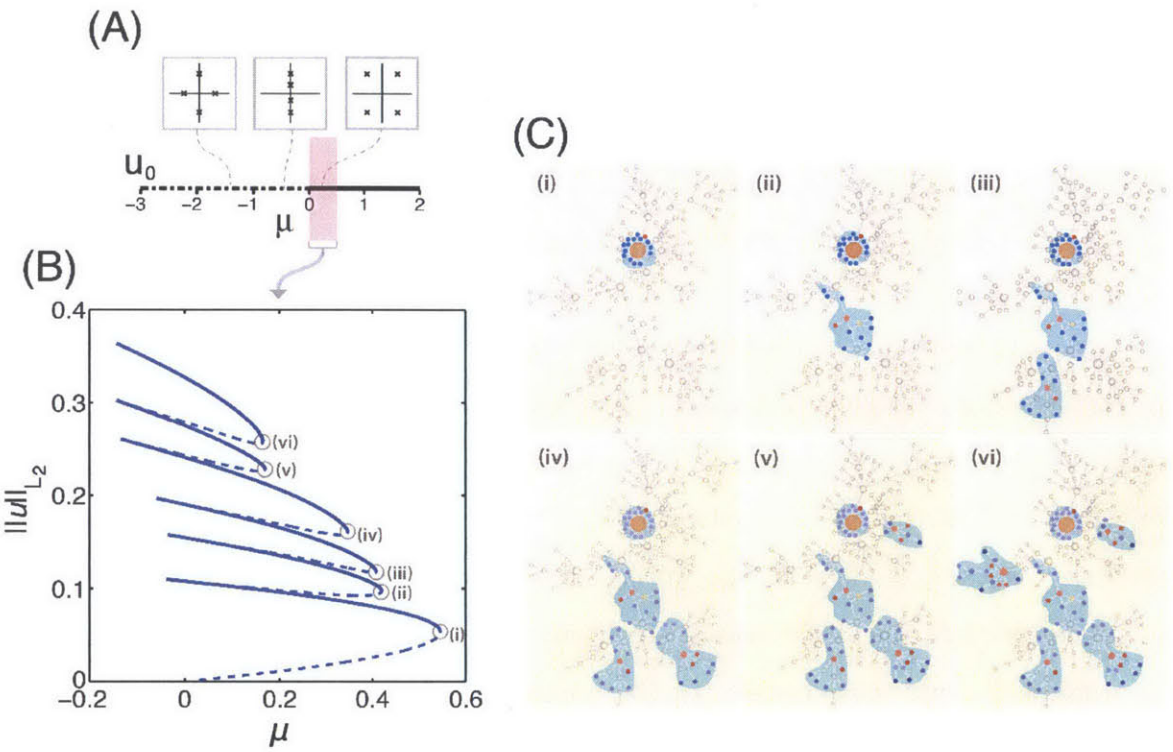

Figure 4-4: **Localized self-organized quantized patterns.** **(A)** Stability of the trivial flat stationary state of our model with respect to the values of the bifurcation parameter, μ. For positive values of μ the trivial stationary solution is stable with respect to uniform small random perturbations (solid line) while for negative values of μ this state becomes unstable (dotted line). Also shown in the insets are the topological eigenvalues of the trivial state as we tune the bifurcation parameter. The behavior eigenvalues in the neighborhood of $\mu = 0$ indicates the possibility for localized patterns in the neighborhood of small positive values of μ (pink shaded region). **(B)** A single branch of the bifurcation diagram in a Barabási-Albert network model of size $N = 200$ with mean degree equal to $\langle k \rangle = 3$ and minimum degree equal to 1. Solid (dotted) lines represent stable (unstable) localized solutions. **(C)** Visualization of the localized patterns corresponding to the states indicated on the bifurcation diagram (B). Gray-colored nodes are non-active ($u = 0$), red-colored nodes are active with $u > 0$ and blue-colored nodes are active with $u < 0$. The size of the node is proportional to its eigenvalue centrality.

4.6.4 Robustness of the localized states

The response of the system is quantized: the transition from one pattern of activation to another one is discontinuous as we vary the activation energy $||\mathbf{u}^0||$, or the parameter μ (Fig. 4-4B). These jumps in activation energy correspond to the addition of neighbor nodes to the cluster (Fig. 4-4C). The discontinuous—quantized—nature of the neural network response leads to robustness in the local patterns.

To gain insight into the robustness of the obtained localized patterns of activation, we performed a synthetic test where we initially stimulate a specific neighborhood in the network, where we set $u_i = \hat{u} \geq 0$ (i.e. a step-like function signal in network topology) and let the system evolve to equilibrium without decay. We then increase gradually the amplitude $\hat{u}$ of the initial signal, and record the final energy values of the equilibrium, localized states. In Figure 4-5 we show the resulting energy of the quantized state with respect to the amplitude of the initial input signal. For small amplitudes the perturbation relaxes back to the resting state, and no activation pattern is elicited. There is a threshold in the energy of the input stimulus beyond which robust quantized states are form. The states are robust in the sense that further increments in the input signal amplitude do not change the final equilibrium pattern.

The self-organized local structures are also robust in the sense of invariance with respect to random noise in the initial stimulus. To check if this property of the brain activity is also property of the proposed model we use Monte Carlo simulations. By increasing the ratio of the noise amplitude to the signal amplitude, we record the energy of the quantized state over many realizations (see Figure 4-6.). Note that the starting input signal (without noise) is the same step-like function on the same neighborhood as before with large amplitude that cause the formation of the robust quantized state of Figure 4-5. Random noise of small amplitude has no effect on the final states. When the signal-to-noise ratio decreases, we do observe a departure from the energy of the base equilibrium state (see Figure 4-6.).

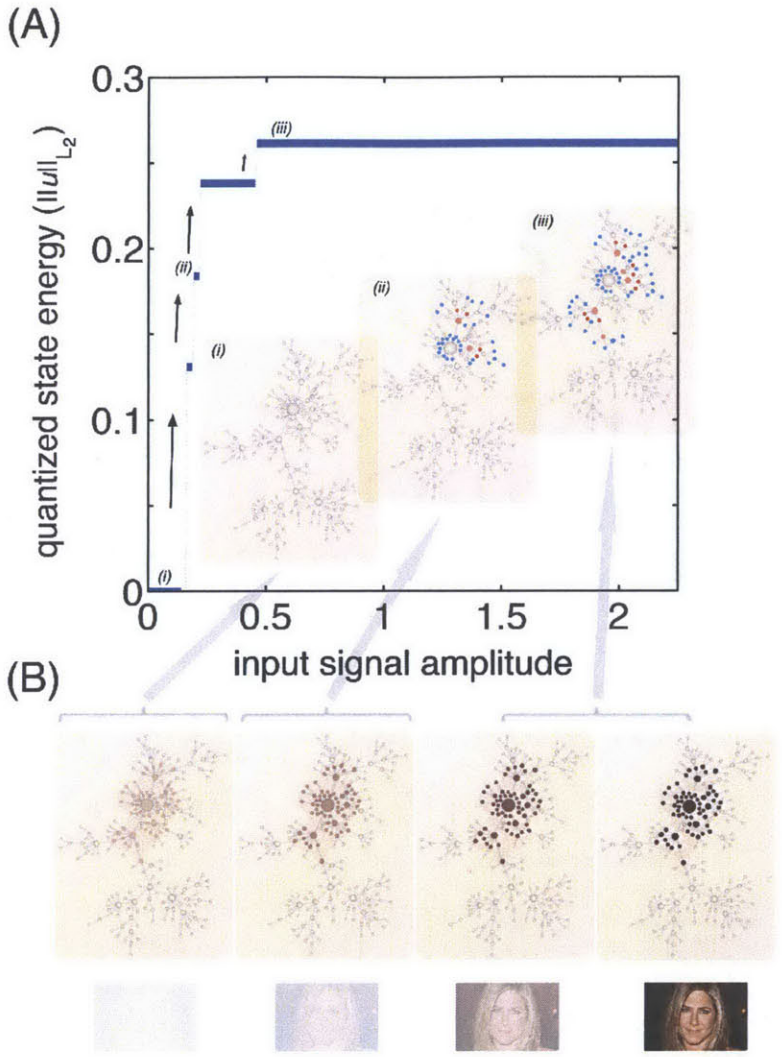

Figure 4-5: **Robustness of quantized patterns with respect to the input signal amplitude.** **(A)** Energy of the resulting quantized state with respect to the input signal amplitude $\hat{u}$ at the nearest and next-nearest neighbors of the best connected node in the system. When the amplitude is very small, the initial perturbation relaxes back to the trivial solution and no quantized state is formed (i). As the amplitude of the input signal is increased, fragile quantized states are formed (ii). When the amplitude of the input signal is larger than a threshold value, a very robust quantized state is formed (iii). Further increases in the input signal amplitude lead to the same quantized state. **(B)** Visualization of the input signal in our network topology. The amplitude increases from left to right. With the metaphor example of Jennifer Aniston, we illustrate the robustness with respect to the amplitude of the input signal that cell assemblies experienced in neural network activity. When we look at a picture with a small signal amplitude our brain usually ends up to wrong conclusions. When the amplitude of the signal we receive is large enough, that leads to the right conclusion about the concept that represents. Further increases in the input signal amplitude lead to the same conclusions (i.e. same cell assembly is getting fired).

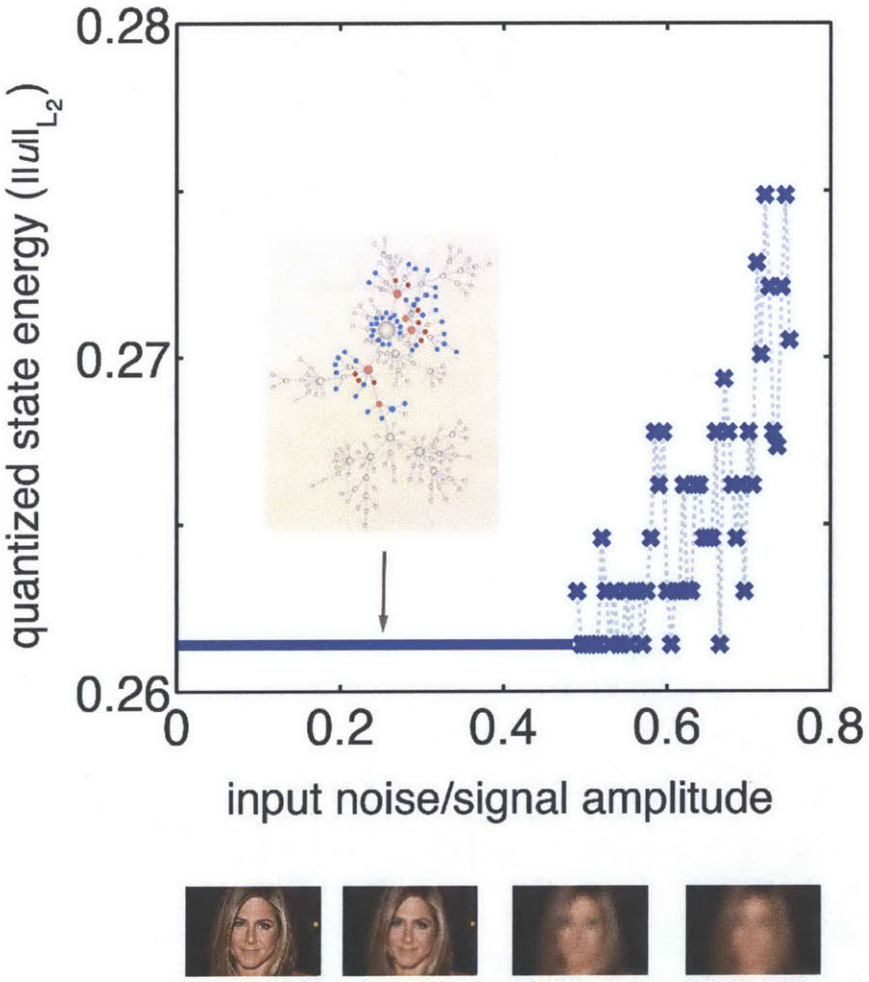

Figure 4-6: **Robustness of quantized patterns with respect to the noise over the signal amplitude of the input.** The energy of the resulting quantized state with respect to the ratio between the signal amplitude and the noise amplitude. Starting from the step-like input signal that gives the robust quantized state ("Jennifer Aniston"), we add random noise at the already perturbed neighborhood and we compute the energy of the resulting quantized state over 100 realizations. We use a Barabási-Albért, scale-free network of size N=200 and mean degree 4.

4.7 Global patterns and mean-field approximation

Our model predicts a range of parameter values where localized states disappear, replaced by global activation patterns. Physiologically, these may be related to pathological functional disorders that cause a transition from a normal brain activity to a synchronous, large scale activation of neurons—seizures (Penfield and Jasper, 1954; Fisher et al., 2005; Chavez et al., 2010). Mathematically, global patterns are possible when the non-active stationary solution is perturbed outside the parameter region of localized patterns ($\mu < 0$). These—global—Turing patterns can be understood and modeled using the mean-field approximation, a method that has been successfully used to describe a wide variety of dynamical processes in heterogeneous networks, from epidemic spreading (Pastor-Satorras and Vespignani, 2001; Nicolaides et al., 2013) and activator-inhibitor models (Nakao and Mikhailov, 2010; Kouvaris and Mikhailov, 2012), to coupled oscillators (Arenas et al., 2008) and voter models (Baronchelli et al., 2011).

4.7.1 Direct Simulations

We are interested in identifying a global activation, steady state pattern of the dynamical system described by Eq. 4.9. We first choose a value for the bifurcation parameter μ^0 in the phase space where global activity patterns are possible, i.e. $\mu^0 < 0$. We then initialize the activation potential profile on the network by setting $u_i = 0$ everywhere. By introducing a small, uniform distributed perturbation in the network, we advance the dynamical system in time, using implicit Runge-Kutta method. Since the trivial state for $\mu < 0$ is unstable with respect to small uniform noise, the initial exponential growth of the perturbation is followed by a nonlinear process leading to the formation of stationary Turing patterns (Nakao and Mikhailov, 2010; Turing, 1952) (Fig. 4-7).

4.7.2 Mean Field Approximation

Since in our model both the degree and two-jump degree play important in the formation of patterns we use a mean field approximation where we assume that all the

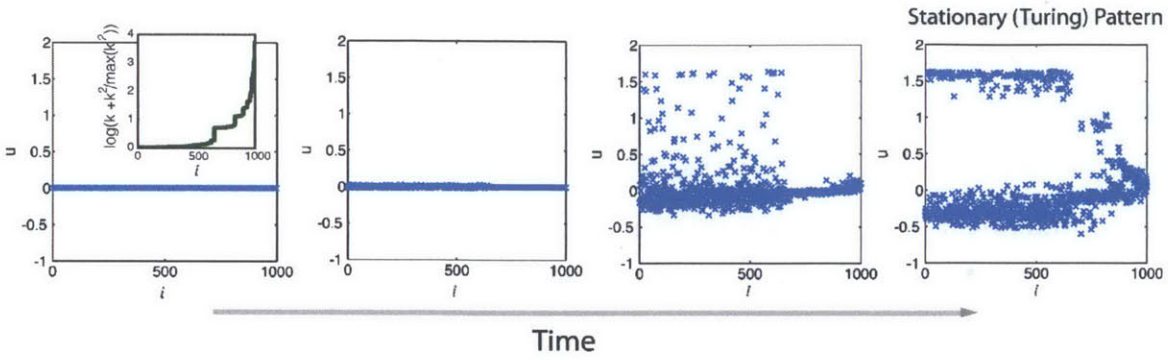

Figure 4-7: **The formation of global activation Turing patterns in a scale-free network.** The time evolution of the proposed model of Eq. 4.9 with bifurcation parameter equal to $\mu = -0.25$ on a scale-free network of size N = 1,000 nodes and mean degree $\langle k \rangle = 4$. The initial exponential growth of the perturbation is followed by a nonlinear process leading to the formation of stationary Turing patterns. We sort node index in increasing degree k. Nodes with the same degree are sorted with increasing two-jump connectivity $k^{(2)}$ (see inset at the very left figure).

nodes with the same degree and two-jump degree behave in the same way. We start by writing Eq. 4.9 in the form

$$\frac{du_i}{dt} = f(u_i) - 2(h_i - k_i u_i) - (l_i - g_i - k_i h_i + k_i^2 u_i), \tag{4.20}$$

where the local fields felt by each node, $h_i = \sum_{j=1}^{N} A_{ij} u_j$, $l_i = \sum_{j=1}^{N} (A^2)_{ij} u_j$ and $g_i = \sum_{j=1}^{N} A_{ij}(k_j u_j)$ are introduced. These local fields are then approximated as $h_i \simeq k_i H_u$, $l_i \simeq k_i^{(2)} H_{uu}$ and $g_i \simeq k^{(2)} H_u$, where $k_i = \sum_{j=1}^{N} A_{ij}$ is the degree and $k_i^{(2)} = \sum_{j=1}^{N} (A^2)_{ij}$ is the number of secondary connections of node i (two-jump degree). The global mean fields are defined by $H_u = (1/N) \sum_k N_k H_u^k$ where $H_u^k = (1/(kN_k)) \sum_{i \varepsilon k} \sum_j A_{ij} u_j$ and $H_{uu} = (1/N) \sum_{k^{(2)}} N_{k^{(2)}} H_{uu}^{k^{(2)}}$, where $H_{uu}^{k^{(2)}} = (1/(k^{(2)} N_{k^{(2)}})) \sum_{i \varepsilon k^{(2)}} \sum_j (A^2)_{ij} u_j$. Here, N_k is the number of nodes with degree k, $N_{k^{(2)}}$ is the number of nodes with $k^{(2)}$ number of two-jump neighbors and $N = \sum_k N_k = \sum_{k^{(2)}} N_{k^{(2)}}$ is the size of the network.

With this approximation, the individual model equation on each node interacts

only with the global mean fields H_u and H_{uu} and its dynamics is described by :

$$\begin{aligned} \frac{du(t)}{dt} = \quad & f(u) \quad -2\alpha(H_u - u) - \\ & - \quad [\beta H_{uu} - \alpha^2 H_u - \beta H_u + \alpha^2 u] \end{aligned} \tag{4.21}$$

We have dropped the index i, as all nodes obey the same equations and we introduced the parameters $\alpha(i) = k_i$ and $\beta(i) = k_i^{(2)}$. If the global mean fields H_u and H_{uu} are given (can be calculated from the global Turing pattern), as well as the parameters α and β for each node, the time independent version of above mean field equation can be written as a third degree algebraic equation that we solve N times. For each node i, we get three solutions u_i^l, $l = 1..3$ that can be stable or unstable depending on the sign (negative or positive) of the operator $f'|_{u_i^l} + 2\alpha - \alpha^2$. In the Figure 4-8 (for the small "toy" network used in Fig. 4-3)) as well as in Figure 4-9 for larger networks we plot the Turing patterns from direct simulations as well as the solution of the mean field equation. The mean field approximation fits *very well* the Turing global activation profile.

4.8 Discussion

Our results suggest a new mechanisms for the formation of cell assemblies in neural networks. Rather than relying on synaptic plasticity, we show that localized, robust cell assemblies are possible due to self-organization. The spontaneous creation of localized activation patterns is potentially ubiquitous due to the simple and general functional structure of the proposed conceptual model: local dynamics based on activation potentials, and interactions between neurons that induce short-range anti correlation and long-range correlation in activity. The proposed mechanism is compatible with Hebbian learning. In fact, the robustness of the localized activation patterns provides a powerful feedback reinforcing the action of synaptic plasticity. Our model also provides a mechanism for learning based on self-organization. The conceptual advantage of self-organization in the brain is that the formation of cell

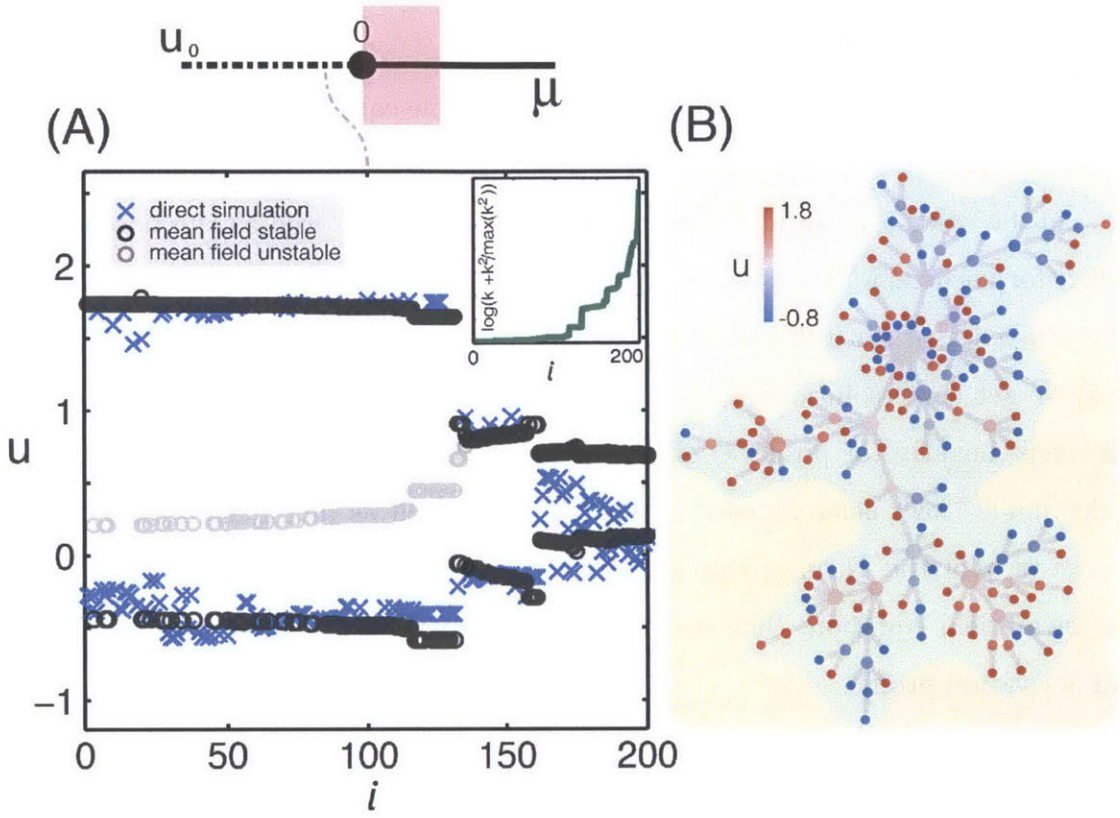

Figure 4-8: **Global self-organization patterns for our toy network model.** Global patterns are possible when the non-active stationary solution is perturbed outside the parameter region of localized patterns ($\mu < 0$). The initial exponential growth of the perturbation is followed by a nonlinear process leading to the formation of stationary Turing patterns. **(A)** The activation profile as a function of the node index i of a global stationary Turing pattern from direct simulation (blue crosses) is compared with the mean-field bifurcation diagram. Black curves indicate stable branches while grey curves correspond to unstable branches of a single activator–inhibitor system coupled to the computed global mean fields. We sort the node index in increasing connectivity k. Nodes with the same degree are sorted with increasing two-jump connectivity $k^{(2)}$ (see Inset). We use the same network model as in Fig. 2 and we set the bifurcation parameter equal to $\mu = -1/4$. **(B)** Visualization of the global activity pattern on the network topology.

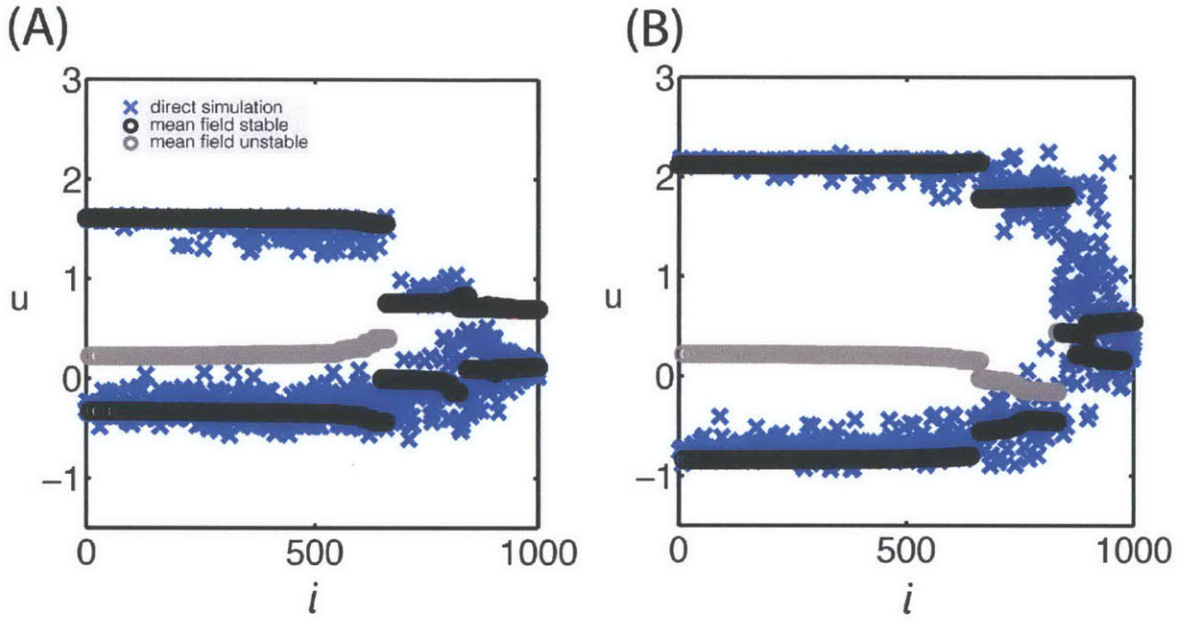

Figure 4-9: **Global self-organization patterns for large networks. (A)** The activation profile as a function of the node index i of global stationary Turing patterns from direct simulation for bifurcation parameters (A) $\mu = -0.25$ and (B) $\mu = -1.5$ are compared with the mean-field solution. The bifurcation parameter Black curves indicate stable branches while grey curves correspond to unstable branches of a single activatorinhibitor system coupled to the computed global mean fields, (A) $H_u = 0.119$ and $H_{uu} = 0.110$, (B) $H_u = 0.400$ and $H_{uu} = 0.389$. We sort the node index in increasing connectivity k. Nodes with the same degree are sorted with increasing two-jump connectivity $k^{(2)}$. The model is solved on a scale-free network with size $N = 1000$ and mean degree $\langle k \rangle = 4$.

assemblies is much faster that when synaptic plasticity is needed as an explanation.

Because of their robustness and localization, self-organized structures may indeed provide an encoding mechanism for information processing and computation in the brain. They may form the basic elements of the brain's internal language. An important question that will be addressed in the future is whether localized, self-organized network structures can provide better algorithms for machine learning and artificial intelligence.

Chapter 5

Conclusions & Future Directions

We have study several dynamical processes in complex network topologies with emphasis on the real world applications that may represent, from global scale epidemic spreading to better understand cell assemblies formation in neural activity networks. The results presented in the previous chapters clearly show that in order to understand the dynamics of complex networks it is essential to take into account the spatial and topological features of the substrate systems as well as the the "traffic" disorder they experience.

In Chapter 2, we developed a new metric for influential spreaders in reaction–diffusion processes through space embedded heterogeneous networks using as case study the spreading of an infectious disease through the air transportation network. We showed with detailed in modeling and resolved in time numerical simulations that the spreading influence of individual nodes in a space embedded network with traffic disorder is not only a function of the total outgoing traffic or the connectivity. However, a combination of properties is vital to understand and quantify the spreading influence of individual nodes. We showed that total traffic, connectivity, space embedding and traffic bias are all important ingredients in order to describe influential spreaders in contagion processes through realistic heterogeneous mobility networks.

Future work includes the study of mitigation strategies during a disease spreading scenario through the world air transportation network. Up to this point, we have only studied the influence of an airport when a disease initiated there, without consider-

ing any intervention scenarios. Many questions arise about the effects of mitigation strategies, particularly when these interventions are associated with a cost. For example, during a scenario of public health emergency, health authorities must take some measures in order to decrease the effect of a virus prevalence. However, these mitigation actions have to reflect an optimum scenario that maximizes the social welfare on the one hand but also minimizes the cost of intervention. In other words, we need to study the advantages in terms of the disease containment, of closing each one of the significant airports during an emergency as well as the associated cost. A data driven integrated study like this will better inform authorities about efficient and and low cost mitigation strategies during a health emergency event.

In Chapter 3, we incorporated behavioral changes in human mobility during a public health emergency, driven by a level of awareness of individuals in the system. We modeled the propagation of awareness as an additional contagion process in the system and we studied the effect of two kind of behavioral changes: a policy driven and a selfish behavior both in a metapopulation model and a simplified conceptual model of human mobility. Using results from high quality numerical simulations on heterogeneous artificial and data driven mobility networks, we define the price of anarchy in mobility driven epidemic spreading. Our results suggest that the structure of the underlying mobility network and its traffic properties have to be significantly considered before authorities impose policy initiated action for the containment of an epidemic.

Future work includes the study of the hypothesis we drew here in a realistic scenario of disease spreading, like the H1N1 case in 2009. This can incorporates human mobility data with a multiscale nature (air travel, commuting etc) as well as the dataset about cases of infected individuals around the world as a function of time. This will allow us to develop a model that can simulate what happened during that emergency scenario with up to a level of precision. Furthermore, having a model like this will help us implement and quantify the effect of hypothetical behavioral changes imposed by the authorities as well as understanding the price of anarchy in such a realistic scenario. This will give us confidence in notifying public policy authorities

about the effectiveness of intervention in mobility behavior of individuals during a infectious disease spreading event.

In Chapter 4, we introduced and studied a minimal-ingredients model of activation and interaction within a heterogeneous neural network that is able to self-organize input signals in globally connected network into localized and quantized patterns. The requirements are minimal and general: simple local dynamics based on typical activation potentials, and interactions between neurons that induce short-range anti correlation and long-range correlation in activity. This model can serve as a potential alternative to the theory of plasticity for the formation of cell assemblies (memory units) in a mammalian brains. Our results suggested that these self-organized, local patterns of activity can provide an encoding mechanism for information processing and neural computation.

Future work include the development of a neural classifier based on our proposed model of self-organization in network topology. This model will require encoding training of the cell assemblies (quantized patterns) based on the theory of synaptic plasticity. This can serve as can be the base for a classifier that can provide better algorithms for machine learning and artificial intelligence.

Bibliography

Albert, R. and A.-L. Barabási (2002). Statistical mechanics of complex networks. *Rev. Mod. Phys. 74*, 47–97.

Albert, R., H. Jeong, and A.-L. Barabási (1999). Internet: Diameter of the world-wide web. *Nature 401*, 130–131.

Anderson, R. M., C. Fraser, A. C. Ghani, C. A. Donnelly, S. Riley, N. M. Ferguson, G. M. Leung, T. Lam, and A. J. Hedley (2004). Epidemiology, transmission dynamics and control of sars: the 2002–2003 epidemic. *Philosophical Transactions of the Royal Society of London. Series B: Biological Sciences 359*(1447), 1091–1105.

Anderson, R. M., R. M. May, and B. Anderson (1992). Infectious diseases of humans: Dynamics and control. *Oxford: Oxford University Press*.

Apicella, C. L., F. W. Marlowe, J. H. Fowler, and N. A. Christakis (2012). Social networks and cooperation in hunter-gatherers. *Nature 481*, 497.

Aral, S. and D. Walker (2012). Identifying influential and susceptible members of social networks. *Science 337*(6092), 337–341.

Arenas, A., A. D. Guilera, J. Kurths, Y. Moreno, and C. Zhou (2008). Synchronization in complex networks. *Phys. Rep. 469*, 93–153.

Bajardi, P., C. Poletto, J. J. Ramasco, M. Tizzoni, V. Colizza, and A. Vespignani (2011). Human mobility networks, travel restrictions, and the global spread of 2009 H1N1 pandemic. *PloS ONE 6*, e16591.

Balcan, D., V. Colizza, B. Gonçalves, H. Hu, J. J. Ramasco, and A. Vespignani (2009). Multiscale mobility networks and the spatial spreading of infectious diseases. *Proc. Natl. Acad. Sci. USA 106*, 21484–21489.

Balcan, D., H. Hu, B. Goncalves, P. Bajardi, C. Poletto, J. J. Ramasco, D. Paolotti, N. Perra, M. Tizzoni, W. V. Broeck, et al. (2009). Seasonal transmission potential and activity peaks of the new influenza a (h1n1): a monte carlo likelihood analysis based on human mobility. *BMC medicine 7*(1), 45.

Balcan, D., H. Hu, B. Goncalves, P. Bajardi, C. Poletto, J. J. Ramasco, D. Paolotti, N. Perra, M. Tizzoni, W. V. Broeck, V. Colizza, and A. Vespignani (2009). A new measure of rank correlation. *BMC Medicine 7*, 45.

Balcan, D. and A. Vespignani (2011). Phase transitions in contagion processes mediated by recurrent mobility patterns. *Nat. Phys.* 7(7), 581–586.

Barabási, A.-L. (2009). Scale-free networks: A decade and beyond. *Science 325*, 412–413.

Barabási, A.-L. and R. Albert (1999). Emergence of scaling in random networks. *Science 286*, 509 -512.

Barabási, A.-L., N. Gulbahce, and J. Loscalzo (2007). Network medicine from obesity to the diseasome. *N. Engl. J. Med. 357*, 404–407.

Barnhart, C., D. Fearing, and V. Vaze (2010). Modeling passenger travel and delays in the national air transportation system. *Submitted to Operations Research.*

Baronchelli, A., C. Castellano, and R. Pastor-Satorras (2011). Voter models on weighted networks. *Phys. Rev. E 83*, 066117.

Barrat, A., M. Barthélemy, R. Pastor-Satorras, and A. Vespignani (2004). The architecture of complex weighted networks. *Proc. Natl. Acad. Sci. USA 101*, 3747–3752.

Barrat, A., M. Berthélemy, and A. Vespignani (2005). The effects of spatial constraints on the evolution of weighted complex networks. *J. Stat. Mech. 2005*(05), P05003.

Barthélemy, M. (2011). Spatial networks. *Phys. Rep. 499*, 1–101.

Barthélemy, M., C. Godréche, and J.-M. Luck (2010). Fluctuation effects in metapopulation models: Percolation and pandemic threshold. *J. Theory Biol. 267*, 554–564.

Bauch, C. T. and D. J. D. Earn (2004). Vaccination and the theory of games. *Proc. Natl. Acad. Sci. USA 101*, 13391–13394.

Bear, M. F. (1996). A synaptic basis for memory storage in the cerebral cortex. *Proc. Natl. Acad. Sci. USA 93*, 13453–13459.

Belik, V., T. Geisel, and D. Brockmann (2011). Natural human mobility patterns and spatial spread of infectious diseases. *Phys. Rev. X 1*, 011001.

Belykh, I., E. de Lange, and M. Hasler (2005). Synchronization of bursting neurons: what matters in the network topology. *Phys. Rev. Lett. 94*(18), 188101.

Bonacich, P. (1987). Power and centrality: a family of measures. *Am. J. Sociol. 92*, 1170–1182.

Brockmann, D., L. Hufnagel, and T. Geisel (2006). The scaling laws of human travel. *Nature 439*, 462–465.

Bullmore, E. and O. Sporns (2009). Complex brain networks: Graph theoretical analysis of structural and functional systems. *Nature Rev. Neurosci. 10*, 186.

Bullmore, E. and O. Sporns (2012). The economy of brain network organization. *Nature Rev. Neurosci. 13*, 336.

Burke, J. and E. Knobloch (2006). Localized states in the generalized swift-hohenberg equation. *Phys. Rev. E 73*, 056211.

Buzsáki, G. (2010). Neural syntax: Cell assemblies, synapsembles, and readers. *Annu. Rev. Physiol. 68*, 362.

Castellano, C. and R. Pastor-Satorras (2010). Thresholds for epidemic spreading in networks. *Phys. Rev. Lett. 105*, 218701.

Castellano, C. and R. Pastor-Satorras (2012). Competing activation mechanisms in epidemics on networks. *Sci. Rep. 2*(371).

Catanzaro, M., M. Boguñá, and R. Pastor-Satorras (2005). Generation of uncorrelated random scale-free networks. *Phys. Rev. E 71*, 027103.

Centola, D. (2010). The spread of behavior in an online social network experiment. *Science 329*(5996), 1194–1197.

Chavez, M., M. Valencia, V. Navarro, V. Lattora, and J. Martinerie (2010). Functional modularity of background activities in normal and epileptic brain networks. *Phys. Rev. Lett. 104*, 118701.

Chialvo, D. R. (2010). Emergent complex neural dynamics. *Nat. Phys. 6*(10), 744–750.

Colizza, V., A. Barrat, M. Barthélemy, A.-J. Valleron, and A. Vespignani (2007). Modeling the worldwide spread of pandemic influenza: baseline case and containment interventions. *PLoS Medicine 4*, 95–110.

Colizza, V., A. Barrat, M. Barthélémy, and A. Vespignani (2006). The role of airline transportation network in the prediction and predictability of global epidemics. *Proc. Natl. Acad. Sci. USA 103*, 1893–1921.

Colizza, V., R. Pastor-Satorras, and A. Vespignani (2007). Reaction-diffusion processes and metapopulation models in heterogeneous networks. *Nat. Phys. 3*, 276–282.

Colizza, V. and A. Vespignani (2007). Invasion threshold in heterogeneous metapopulation networks. *Phys. Rev. Lett. 99*, 148701.

Consortium, C. S. M. E. et al. (2004). Molecular evolution of the sars coronavirus during the course of the sars epidemic in china. *Science 303*(5664), 1666–1669.

Cross, M. C. and P. C. Hohenberg (1993). Pattern formation outside of equilibrium. *Rev. Mod. Phys. 65*, 851–1112.

DallAsta, L., A. Baronchelli, A. Barrat, and V. Loreto (2006). Nonequilibrium dynamics of language games on complex networks. *Phys. Rev. E 74*(3), 036105.

Dudai, Y. (2004). The neurobiology of consolidations, or, how stable is the engram? *Annu. Rev. Psychol. 55*, 51–86.

Dune, J., R. J. Williams, and N. D. Martinez (2002). Food-web structure and network theory: The role of connectance and size. *Proc. Natl. Acad. Sci. USA 99*, 12917–12922.

Eagle, N., A. Pentland, and D. Lazer (2009). Inferring friendship network structure by using mobile phone data. *Proc. Natl. Acad. Sci. USA 106*(36), 15274.

Epstein, J. M., D. M. Goedecke, F. Yu, R. J. Morris, D. K. Wagener, and G. V. Bobashev (2007). Controlling pandemic flu: the value of international air travel restrictions. *PloS one 2*(5), e401.

Epstein, J. M., J. Parker, D. Cummings, and R. A. Hammond (2008). Coupled contagion dynamics of fear and disease: Mathematical and computational explorations. *PLoS ONE 3*, e3955.

Erdos, P. and A. Rényi (1960). On the evolution of random graphs. *Publ. Math. Inst. Hung. Acad. Sci 5*, 17–61.

Faloutsos, M., P. Faloutsos, and C. Faloutsos (1999). On power-law relationships of the internet topology. *Comput. Commun. Rev. 29*, 251–262.

Ferguson, N. M., D. A. Cummings, C. Fraser, J. C. Cajka, P. C. Cooley, and D. S. Burke (2006). Strategies for mitigating an influenza pandemic. *Nature 442*(7101), 448–452.

Fisher, R. S., W. van Emde Boas, W. Blume, C. Elger, P. Genton, P. Lee, and J. E. Jr. (2005). Epileptic seizures and epilepsy: Definitions proposed by the international league against epilepsy (ilae) and the international bureau for epilepsy (ibe). *Epilepsy 46*, 470–472.

FitzHugh, R. (1961). Impulses and physiological states in theoretical models of nerve membrane. *Biophys. J. 1*, 445.

Fraser, C. et al. (2009). Pandemic potential of a strain of influenza A(H1N1): Early findings. *Science 324*, 1557–1561.

Fraser, C., C. A. Donnelly, S. Cauchemez, W. P. Hanage, M. D. Van Kerkhove, T. D. Hollingsworth, J. Griffin, R. F. Baggaley, H. E. Jenkins, E. J. Lyons, et al. (2009). Pandemic potential of a strain of influenza a (h1n1): early findings. *science 324*(5934), 1557–1561.

Freeman, L. C. (1977). A set of measures of centrality based on betweenness. *Sociometry 40*, 35–41.

Freeman, L. C. (1979). Centrality in social networks: Conceptual clarification. *Social Networks 1*, 215–239.

Funk, S., E. Gilad, C. Watkins, and V. A. A. Jansen (2009). The spread of awareness and its impact on epidemic outbreaks. *Proc. Natl. Acad. Sci. USA 106*, 6872–6877.

Funk, S., M. Salathé, and V. A. A. Jansen (2010). Modelling the influence of human behaviour on the spread of infectious diseases: a review. *J. R. Soc. Interface 7*, 1247–1256.

Fuster, J. M. (2000). Cortical dynamics of memory. *Int. J. Psychophysiol. 35*, 155–164.

Gallistel, G. R. and L. D. Matzel (2013). The neuroscience of learning: Beyond the hebbian synapse. *Annu. Rev. Psychol. 64*, 169–200.

Garnett, G. P. and R. M. Anderson (1996). Sexually transmitted diseases and sexual behavior: insights from mathematical models. *J. Infect. Dis. 174*, S150 -S161.

Girvan, M. and M. E. J. Newman (2002). Community structure in social and biological networks. *Proc. Natl. Acad. Sci. USA 99*(12), 7821–7826.

Goffman, W. and V. Newill (1964). Generalization of epidemic theory. *Nature 204*(4955), 225–228.

Gomez, H. and X. Nogueira (2012). A new space–time discretization for the swift–hohenberg equation that strictly respects the lyapunov functional. *Communications in Nonlinear Science and Numerical Simulation 17*, 4930.

Gómez, S., A. Arenas, J. Borge-Holthoefer, S. Meloni, and Y. Moreno (2010). Discrete-time Markov chain approach to contact-based disease spreading in complex networks. *EPL 89*, 38009.

González, M. C., H. J. Herrmann, J. Kertész, and T. Vicsek (2007). Community structure and ethnic preferences in school friendship networks. *Physica A: Statistical mechanics and its applications 379*(1), 307–316.

González, M. C., C. A. Hidalgo, and A.-L. Barabási (2008). Understanding individual human mobility patterns. *Nature 453*, 779–782.

Guimera, R. and L. A. L. Amaral (2005). Functional cartography of complex metabolic networks. *Nature 433*, 895–900.

Guimerá, R., S. Mossa, and L. A. N. Amaral (2005). The worldwide air transportation network: Anomalous centrality, community structure, and cities' global roles. *Proc. Natl. Acad. Sci. USA 102*, 7794–7799.

Hagmann, P., M. Kurant, X. Gigandet, P. Thiran, V. J. Wedeen, R. Meuli, and J.-P. Thiran (2007). Mapping human whole-brain structural networks with diffusion mri. *PLoS ONE 2*, e597.

Han, J.-H., S. A. Kushner, A. P. Yiu, C. J. Cole, A. Matynia, R. A. Brown, R. L. Neve, J. F. Guzowski, A. J. Silva, and S. A. Josselyn (2007). Neuronal competition and selection during memory formation. *Science 316*, 457.

Hancock, K., V. Veguilla, X. Lu, W. Zhong, E. N. Butler, H. Sun, F. Liu, L. Dong, J. R. DeVos, P. M. Gargiullo, T. L. Brammer, N. J. Cox, T. M. Tumpey, and J. M. Katz (2009). Cross-reactive antibody responses to the 2009 pandemic h1n1 influenza virus. *N. Engl. J. Med. 361*(20), 1945–1952.

Hatchett, R. J., C. E. Mecher, and M. Lipsitch (2007). Public health interventions and epidemic intensity during the 1918 influenza pandemic. *Proc. Natl. Acad. Sci. USA 104*, 7582–7587.

Hodgkin, A. A. and A. F. Huxley (1952). A quantitative description of membrane current and its application to conduction and excitation in nerve. *J. Physiol. 117*, 500.

Hollingsworth, T. D., N. M. Ferguson, and R. M. Anderson (2006). Will travel restrictions control the international spread of pandemic influenza? *Nature medicine 12*(5), 497–499.

Hufnagel, L., D. Brockmann, and T. Geisel (2004). Forecast and control of epidemics in a globalized world. *Proc. Natl. Acad. Sci. USA 101*, 15124–15129.

Insel, T. R., S. C. Landis, and F. S. Collins (2013). The nih brain initiative. *Science 340*, 687.

Iturria-Medina, Y., R. C. Sotero, E. J. Canales-Rodríguez, Y. Alemán-Gómez, and L. Melie-García (2008). Studying the human brain anatomical network via diffusion-weighted mri and graph theory. *Neuroimage 40*, 1064–1076.

Izhikevich, E. M. and G. M. Edelman* (2008). Large-scale model of mammalian thalamocortical systems. *Proc. Natl. Acad. Sci. USA 105*(9), 3593–3598.

Jeong, H., B. Tombor, R. Albert, Z. N. Oltvai, and A.-L. Barabási (2000). The large-scale organization of metabolic networks. *Nature 407*, 651–654.

Jin, E. M., M. Girvan, and M. E. Newman (2001). Structure of growing social networks. *Phys. Rev. E 64*(4), 046132.

Keller, H. B. (1977). Numerical solution of bifurcation and nonlinear eigenvalue problems. *Applications of Bifurcation Theory*, 359–384.

Kelso, S. R., A. H. Ganong, and T. H. Brown (1986). Hebbian synapses in hippocampus. *Proc. Natl. Acad. Sci. USA 83*, 5326–5330.

Kempe, D., J. Kleinberg, and E. Tardos (2005). Influential nodes in a diffusion model for social networks. In *Automata, languages and programming*, pp. 1127–1138. Springer.

Kendall, M. (1938). A new measure of rank correlation. *Biometrika 30*, 81–89.

Kim, M., M. Bertram, M. Pollmann, A. von Oertzen, A. S. Mikhailov, H. H. Rotermund, and G. Ertl (2001). Controlling chemical turbulence by global delayed feedback: pattern formation in catalytic co oxidation on pt (110). *Science 292*(5520), 1357–1360.

Kitsak, M., L. K. Gallos, S. Havlin, F. Liljeros, L. Muchnik, and H. E. Stanley (2010). Identification of influential spreaders in complex networks. *Nat. Phys. 6*, 888–893.

Kitsak, M., L. K. Gallos, S. Havlin, F. Liljeros, L. Muchnik, H. E. Stanley, and H. A. Makse (2010). Identification of influential spreaders in complex networks. *Nat. Phys. 6*, 888–893.

Kondo, S. and T. Miura (2010). Reaction-diffusion model as a framework for understanding biological pattern formation. *Science 329*, 1616.

Koutsoupias, E. and C. H. Papadimitriou (1999). Worst-case equilibria. *Proceedings of the 16th Annual Symposium on Theoretical Aspects of Computer Science*, 404–413.

Kouvaris, N. E. and H. K. A. S. Mikhailov (2012). Traveling and pinned fronts in bistable reaction-diffusion systems on networks. *Plos One 7*, e45029.

Lansner, A. (2009). Assortative memory models: from the cell-assembly theory to biophysically detailed cortex simulations. *Trends in Neurosciences 32*, 178–186.

Lansner, A., E. Fransén, and A. Sandberg (2002). Cell assembly dynamics in detailed and abstract attractor models of cortical associative memory. *Theory Biosci. 122*, 19–36.

Lazer, D., A. Pentland, L. Adamic, S. Aral, A.-L. Barabsi, D. Brewer, N. Christakis, N. Contractor, J. Fowler, M. Gutmann, T. Jebara, G. King, M. Macy, D. Roy, and M. Van Alstyne (2009). Computational social science. *Science 323*(5915), 721–723.

Lloyd, A. L. and R. M. May (2001). How viruses spread among computers and people. *Science 292*(5520), 1316–1317.

Lloyd, D. J. B., B. Sandstede, D. Avitabile, and A. R. Champneys (2010). Localized hexagon patterns of the planar swifthohenberg equation. *SIAM J. Appl. Dyn. Syst. 7*, 1049–1100.

Lohr, S. (2012). The age of big data. *New York Times*.

Martin, S. J. and P. G. M. Morris (2002). New life in an old idea: The synaptic plasticity and memory hypothesis revisited. *Hippocampus 12*, 609–636.

Meloni, S., A. Arenas, and Y. Moreno (2009). Traffic-driven epidemic spreading in finite-size scale-free networks. *Proc. Natl. Acad. Sci. USA 106*, 16897–16902.

Meloni, S., N. Perra, A. Arenas, S. Gómez, Y. Moreno, and A. Vespignani (2011). Modeling human mobility responses to the large-scale spreading of infectious diseases. *Sci. Rep. 1*, 62.

Milgram, S. (1967). The small world problem. *Psychology today 2*(1), 60–67.

Montoya, J., S. L. Pimm, and R. V. Solé (2006). Ecological networks and their fragility. *Nature 420*, 259.

Nakao, A. and A. S. Mikhailov (2010). Turing patterns in network-organized activatorinhibitor systems. *Nat. Phys. 6*, 544–550.

Nash, J. F. (1950). Equilibrium points in n-person games. *Proc. Natl. Acad. Sci. USA 36*, 48–49.

Neves, G., S. F. Cooke, and T. V. Bliss (2008). Synaptic plasticity, memory and the hippocampus: a neural network approach to causality. *Nat. Rev. Neurosci. 9*(1), 65–75.

Newman, M. E. J., S. H. Strogatz, and D. J. Watts (2001). Random graphs with arbitrary degree distributions and their applications. *Phys. Rev. E 64*, 026118.

Nicolaides, C., L. Cueto-Felgueroso, M. C. González, and R. Juanes (2012). A metric of influential spreading during contagion dynamics through the air transportation network. *PLoS ONE 7*, e40961.

Nicolaides, C., L. Cueto-Felgueroso, M. C. González, and R. Juanes (2014). Quantized states and self-organization in brain activity. *submitted for publication.*

Nicolaides, C., L. Cueto-Felgueroso, and R. Juanes (2010). Anomalous physical transport in complex networks. *Phys. Rev. E 82*, 055101(R).

Nicolaides, C., L. Cueto-Felgueroso, and R. Juanes (2013). The price of anarchy in mobility driven contagion dynamics. *J. R. Soc. Interface 10*(87), 20130495.

Onnela, J.-P., S. Arbesman, M. C. González, A. L. Barabási, and N. A. Christakis (2011). Geographic constraints on social network groups. *PloS ONE 6*, e16939.

Palm, G. (1981). Towards a theory of cell assemblies. *Proc. Natl. Acad. Sci. USA 39*, 181–194.

Papadimitriou, C. and G. Valiant (2010). A new look at selfish routing. In *Innovations in Computer Science (ICS2010).*

Parashar, U. D., E. G. Hummelman, J. S. Bresee, M. A. Miller, and R. I. Glass (2003). Global illness and deaths caused by rotavirus disease in children. *Emerg. Infect. Dis. 9*, 565–572.

Pastor-Satorras, R. and A. Vespignani (2001). Epidemic spreading in scale-free networks. *Phys. Rev. Lett. 86*, 3200–3203.

Pastor-Satorras, R. and A. Vespignani (2002). Epidemic dynamics in finite size scale-free networks. *Phys. Rev. E 65*, 035108(R).

Penfield, W. G. and H. H. Jasper (1954). *Epilepsy and the functional anatomy of the human brain.* Oxford England: Little, Brown.

Perra, N., D. Balcan, B. Gonçalves, and A. Vespignani (2011). Towards a characterization of behavior-disease models. *PLoS ONE 6*, e23084.

Quiroga, Q. R. (2012). Concept cells: The building blocks of declarative memory functions. *Nature Rev. Neurosci. 13*, 587–597.

Quiroga, Q. R., G. Kreinman, C. Koch, and I. Fried (2008). Sparse but not grandmother-cell coding in the medial temporal lobe. *Trends Cogn. Sci. 12*, 87–91.

Quiroga, Q. R., L. Reddy, G. Kreiman, C. Koch, and I. Fried (2005). Invariant visual representation by single neurons in the human brain. *Nature 453*, 1102–1107.

Reijmers, L. G., B. L. Perkins, N. atsuo, and M. Mayford (2007). Localization of a stable neural correlate of associative memory. *Science 317*, 1230.

Rohani, P., D. J. D. Earn, and B. T. Grenfell (2000). The impact of immunization on pertussis transmission in england and wales. *Lancet 355*, 285–286.

Roughgarden, T. (2003). The price of anarchy is independent of the network topology. *J. Comp. Sys. Sci. 67*, 341–364.

Roughgarden, T. (2005). *Selfish Routing and the Price of Anarchy.* MIT Press, Cambridge.

Rutherford, A., M. Cebrian, S. Dsouza, E. Moro, A. Pentland, and I. Rahwan (2013). Limits of social mobilization. *Proc. Natl. Acad. Sci. USA 110*, 6281–6286.

Schneider, C. M., V. Belik, T. Couronné, Z. Smoreda, and M. C. González (2013). Unravelling daily human mobility motifs. *J. R. Soc. Interface 10.*

Schweitzer, F., G. Fagiolo, D. Sornette, F. Vega-Redondo, A. Vespignani, and D. R. White (2009). Economic networks: The new challenges. *Science 325*, 422.

Silva, A. J., Y. Zhou, T. Rogerson, J. Shobe, and J. Balaji (2009). Molecular and cellular approaches to memory allocation and neural circuits. *Science 326*, 391.

Smoller, J. (1983). Shock waves and reaction-diffusion equations. In *Research supported by the US Air Force and National Science Foundation. New York and Heidelberg, Springer-Verlag (Grundlehren der Mathematischen Wissenschaften. Volume 258), 1983, 600 p.*, Volume 258.

Song, C., T. Koren, P. Wang, and A.-L. Barabási (2010). Modelling the scaling properties of human mobility. *Nat. Phys. 6*, 818 823.

Spirin, V. and L. A. Mirny (2003). Protein complexes and functional modules in molecular networks. *Proc. Natl. Acad. Sci. USA 100*(21), 12123–12128.

Sporns, O. (2011). *Networks of the brain.* MIT Press.

Sporns, O. (2013). Making sense of brain network data. *Nature Methods 10*, 491.

Sporns, O., D. R. Chialvo, M. Kaiser, and C. C. Hilgetag (2004). Organization, development and function of complex brain networks. *Trends Cogn. Sci. 8*, 418.

Steyvers, M. and J. B. Tenenbaum (2005). The large-scale structure of semantic networks: Statistical analyses and a model of semantic growth. *Cognit. Sci. 29*(1), 41–78.

Swift, J. and P. C. Hohenberg (1977). Hydrodynamic fluctuations at the convective instability. *Phys. Rev. A 15*, 319.

Tlidi, M., P. Mandel, and R. Lefever (1994). Localized structures and localized patterns in optical bistability. *Phys. Rev. Lett. 73*, 640.

Toole, J. L., M. Cha, and M. C. González (2012). Modeling the adoption of innovations in the presence of geographic and media influences. *PLoS ONE 7*, e27528.

Tsodyks, M. (2005). Attractor neural networks and spatial maps in hippocampus. *Annu. Rev. Physiol. 44*, 168–169.

Turing, A. M. (1952). The chemical basis of morphogenesis. *Phil. Trans. R. Soc. Lond. B 237*, 37–72.

Vanag, V. K. and I. R. Epstein (2001). Inwardly rotating spiral waves in a reaction-diffusion system. *Science 294*(5543), 835–837.

Vespignani, A. (2012). Modelling dynamical processes in complex socio-technical systems. *Nat. Phys. 8*, 32–39.

Wang, D., C. Song, and A.-L. Barabási (2013). Quantifying long-term scientific impact. *Science 342*(6154), 127–132.

Watts, D. J. and S. H. Strogatz (1999). Collective dynamics of 'small-world' networks. *Science 286*, 509–512.

Youn, H., M. T. Gastner, and H. Jeong (2008). Price of anarchy in transportation networks: Efficiency and optimality control. *Phys. Rev. Lett. 101*, 128701.