Boston University. Spring 2014.
PY538: Econophysics.
Topology, correlations and opinion weighting in a stochastic model of opinion formation.

Bernardo J. Zubillaga Herrera. BU-ID: U99149765

In this project, we study the inclusion of asymmetries in the influences of agents in a model of stochastic opinion formation dynamics proposed by Bartolozzi, Leinweber and Thomas. The model is based on scale-free complex networks, in which the nodes represent agents that formulate binary opinions on a particular issue under the influence of their local environment (first neighbors) and the global behavior of the network. The model is defined in analogy with systems of interacting spin 1/2 particles in thermal contact with a heat bath. One can interpret the financial agents as nodes within the network and the two possible states (spin projections) of each agent represent buying/selling decisions regarding a financial asset. The return of the asset is defined as the instantaneous magnetization of the system. The weighting of influences of the agents' opinions is taken to be correlated with the degrees of such agents within the network in such a way that nodes of higher degree have greater influence on their neighbors than nodes of lower degree. Also, we propose a way to have the model exhibit time correlations by means of volatility clustering. The dependence of the model on the network topology is studied.

I. INTRODUCTION

The study of social dynamics has been a recent focus of investigation within the field of complex systems. The application of the science of complex networks in the study of real social networks has consistently suggested the existence of fundamental organization principles in such systems. Studies of sexual contact networks, the World Wide Web, the Internet, actor collaboration networks, scientific publication networks and financial networks have revealed that, even though the natures of these networks seem unrelated and independent, there are various common characteristics between them[1–4].

Complex networks are ideal substrates not only for representing networks of social interaction, but also for studying social dynamics defined on the structure of the complex networks. In particular, physicists have studied opinion formation dynamics in social contexts and have proposed models based on microscopic interaction rules between individuals and their surroundings. In typical models, individuals are prone to the influence of their immediate environment (neighborhood of acquaintances) and to the global influence of society as whole.

One useful social system with which to compare the results of these models is a financial market, where a wealth of information of price histories of financial assets has been recorded for decades (thanks to the power of computers). Such statistics are widely available. Therefore, it is not surprising that microscopic social dynamics models have been tailored to study these systems.

II. FUNDAMENTAL NOTIONS ON COMPLEX NETWORKS

A *network* or *graph* is defined as a set of elements, called *nodes* or *vertices*, linked by *connections* or *edges*. A connection represents a binary relation between the pair of nodes it links. Networks have a simple graphical representation: nodes are represented by a set of points and the edges are represented by lines linking the points. Each pair of nodes linked by an edge is called a pair of *neighbor nodes*. The *degree of connectivity*, k, (or simply the *degree*) of a node is defined as the number of edges linked to that node. For the purposes of this project, the degree of a node coincides with its number of neighbors.

Up until the decade of 1950−1960, the study of networks was led by mathematicians in the formal theory of graphs. The *random network model* of Hungarian mathematicians Paul Erdős and Alfréd Rényi was an important advance in the study of real-world networks, which seemed to have no apparent principles of organization or laws of construction. However, as technology advanced, computers allowed for a more thorough study of the properties of the topologies of real networks. The subsequent studies of these networks have systematically revealed the existence of non-trivial topological characteristics. Phenomena such as small world properties, high clustering probabilities and power law degree distributions have been identified as key elements in the description of the topologies of real networks[1,2].

In a complex network, not all nodes have the same number of edges and, hence, not all nodes have the same degree. To describe and quantify the dispersion of the degrees of the nodes in a network, we define a *degree distribution* function, $P(k)$. This function provides the probability of finding a node in the network with degree equal to k.

The probability distribution function ($P(k)$) of the degree (k) of a node in a real complex network, shows typical quantitative characteristics in the previously mentioned systems (WWW, Internet, collaboration networks, etc.). Scale invariant degree distribution functions are ubiquitous (i.e. a power law), at least in the tails: $P(k) \sim k^{-\gamma}$. Networks that show this feature are referred to as *scale-free*[5].

Due to the ubiquitous appearance of scale-free degree distributions in social complex networks, it is reasonable to use these topologies for the underlying financial networks in simulations of microscopic market dynamics[6,7]. To this end, we use the Barabási-Albert model, which is a stochastic network construction algorithm for scale-free networks.

The Barabási-Albert model generates scale-free networks with degree distribution:

$$P(k) = 2m^2 k^{-\gamma} \tag{1}$$

with $\gamma = 3$. The algorithm is based on two fundamental mechanisms: *network growth* and *preferential attachment*[1–4]. Starting with a small network, a new node with m connections is added at each time step; this is what we mean by network growth. The new node will connect to m of the previously existing nodes in the network, according to a probabilistic criterion called *preferential attachment*, given by the following rule: The new node will link to the i- th node of the network with probability $\Pi(i) = k_i / \Sigma_j k_j$.

Fig. (1) shows a scale-free complex network, constructed using the Barabási-Albert algorithm, for $N = 200$ nodes and $m = m_0 = 5$, using Monte Carlo simulations. m_0 is number of nodes in the initial network (before implementing the Barabási-Albert algorithm). Fig. (2) shows the power law degree distributions of equation (1), obtained by Monte Carlo simulations of the Barábasi-Albert algorithm, for a set of 100 networks of a total of 2000 nodes each.

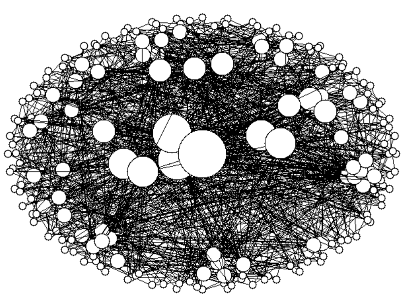

Figure 1: Scale free complex network, constructed using Barabási-Albert algorithm, for $N = 200$ nodes and $m = m_0 = 5$. Nodes of greater size have a higher degree. Note the presence of a few nodes of very high degrees (*hubs*).

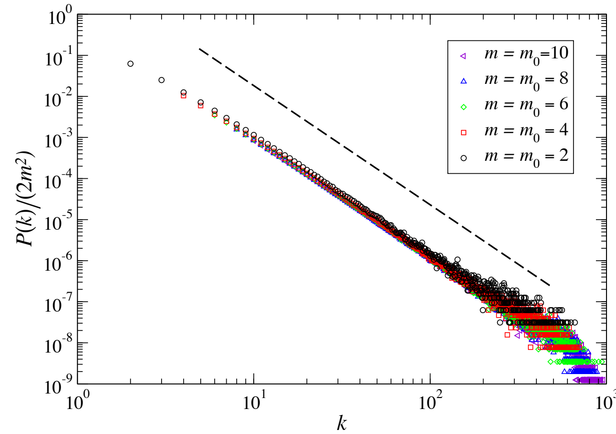

Figure 2: Degree distributions for scale-free networks constructed using the Barabási-Albert model for different values of m. Note that the dashed line corresponds to a power law given by k^{-3}.

The study of opinion formation dynamics in social environments and its potential applications to the study of financial market dynamics has recently gained importance in the field of complex systems. Multiple examples of microscopic models of opinion formation and financial market dynamics have been proposed in the past two decades. In particular, the application of models based on interactions between spin 1/2 particles (Ising-like systems) to the study of financial market and opinion formation dynamics has been of considerable interest to physicists[8–11].

In this project, we will include and investigate the effects of asymmetries in the influences of agents in one of such spin models, proposed by Bartolozzi, Leinweber and Thomas[9]. In order to do so, financial variables (i.e. return and volatility) are defined within the context of a stochastic opinion formation model. We shall be interested in the study of the probability distributions of these variables. The effects of the previously mentioned asymmetries will be quantified by a weighting factor, which is a function of the degrees of the nodes of the underlying networks. We will explore the behaviour of the tails of the return distributions, their variances and their kurtosis with respect to the weighting factor.

Also, the model will be further modified with the purpose of having it show volatility clustering, an empirical aspect that is not reproduced by the original model.

III. FINANCIAL VARIABLES TO CONSIDER

Financial markets are complex systems where a great number of investors (financial agents) interact with one another and react to to the arrival of external and internal information with the result of determining the price for a given asset, such as commodities, stock, derivatives (for example, futures and options), foreign currency and bonds.

The concepts of risk and reward are essential in the context of financial markets. These are quantified typically in terms of returns and volatility, which are defined in terms of the price history of the asset.

Let $S(t)$ denote the price of an asset at time t. The *return* of a financial asset is defined in terms of the price $S(t)$ as:[8,12,13]

$$r_\tau(t) = \frac{S(t+\tau) - S(t)}{S(t)} \tag{2}$$

where τ is the time scale between two successive prices in the time series (for example: seconds, minutes, hours, days...). Thus, the return is a measure of the relative change of the price of the asset; specifically, it measures the percentile increase ($r_\tau(t) > 0$) or decrease ($r_\tau(t) < 0$) of the price between 2 instances of time. (If the price does not change between two successive instances of time, then $r_\tau(t) = 0$).

We define the *normalized returns* in terms of equation (2) as:

$$\tilde{r}_\tau(t) = \frac{r_\tau(t) - \langle r_\tau \rangle}{\sigma_\tau} = \frac{r_\tau - \langle r_\tau \rangle}{\sqrt{\langle r_\tau^2 \rangle - \langle r_\tau \rangle^2}} \tag{3}$$

where σ_τ is the standard deviation of the time series of prices and $\langle r_\tau \rangle$ is the average return[8]. Obviously, the time series of normalized returns has a zero mean and a unit variance.

Another quantity of interest in finance is *volatility*, which is a measure of investment risk for a given asset.[8,12,13]. The volatility of an asset is usually defined in finance as the standard deviation of the distribution of returns over a certain time window. Since real return distributions are typically almost symmetric, negative returns are almost as likely as positive returns and, therefore, the standard deviation of the return distribution is a reasonable measure of the risk of negative returns. Another definition takes the volatility to be the average of the absolute values of the returns over a certain time window. However, for our purposes, the volatility of the financial asset is defined locally in time as the absolute value of the returns[8]:

$$\tilde{v}_\tau(t) = |\tilde{r}_\tau(t)| \tag{4}$$

The *standard model of finance* is comprised of two mathematical models that are typically used to describe the behavior and dynamics of stock shares and the pricing of options[8] . In particular, we are talking about *Geometric Brownian Motion* (which is used to model the dynamics of stock) and the Black-Scholes theory of option pricing, whose underlying asset is stock. The Geometric Brownian assumption for the modeling of stock behavior predicts two major characteristics of the behavior of returns: that the distribution of returns is Gaussian and that returns are not correlated in time.

The perception of validity of the Geometric Brownian Motion in describing the form of empirical return distributions was well established and undisputed until 1963, when Mandelbrot observed that the actual distribution functions exhibit heavy tails, which depart from the Gaussian assumption. Eugene Fama confirmed this observation when studying the return distribution for the DJIA. It is well known that the actual distributions of returns of real stock, indices, commodities and foreign exchange rates fail to conform to this Gaussian prediction[8]. This can be seen clearly in Fig. (3).

The presence of heavy tails in return distributions implies that they are leptokurtic, meaning that the kurtosis, defined as the fourth cumulant of the distribution, is greater than the kurtosis of a normal distribution (which has a kurtosis $K_\tau \simeq 3$). The actual form of the tails has been identified as a power law. This means that the probability of extreme fluctuations in financial time series is much greater than predicted by the Gaussian model, which is reflected in discontinuities or jumps in the real time series. Many functional forms have been proposed for the description of the actual return distributions, including stable Lévy distributions, truncated Lévy distributions and Student's -t distributions, although there is presently no consensus as to the actual true form[8,12].

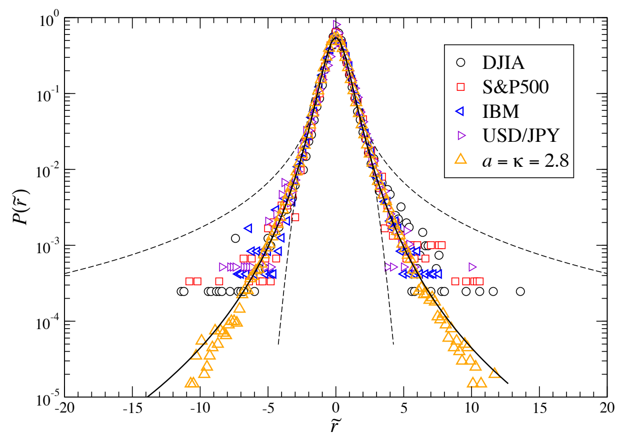

Figure 3: Return distributions of real financial time series and a return distribution from the microscopic model. We show the distribution of daily returns for the DJIA and S&P500 indices, IBM quote and the foreign exchange rate between the US dollar and the Yen. The yellow curve corresponds to Monte Carlo simulations of the microscopic model. for $a = \kappa = 2.8$ and a total of 2000 nodes. The inner dashed line is a Gaussian distribution, while the outer dashed line is a Lorentzian distribution. Both real data and simulations show heavy tails. The continuous dark line corresponds to a fit of the model to a Student's t- distribution. Source of daily price histories used: *Yahoo! Finance*.

Some social psychologists have asserted that the reasons that explain these heavy tails are rooted in psychology of masses. Although it is very likely that the time evolution of the price will be continuous (i.e.,having small returns), future expectations are subject to information flow in the market and *herding behavior* and may exhibit big fluctuations that lead to heavy tails[8,14].

A question of practical importance in finance is the one of memory/correlations in financial time series. Let us define the autocorrelation function of the time series of normalized returns as:

$$C_{r,\tau}(t-t') = \langle \tilde{r}_\tau(t) \tilde{r}_\tau(t) \rangle \tag{5}$$

Eugene Fama, one of the fathers of the *Efficient Market Hypothesis*, studied these autocorrelation functions of returns empirically for a wide variety of assets, noticing that relative price changes seem to be uncorrelated in time. For time scales of observation typically larger than 30 minutes, the autocorrelation functions of return shows this characteristic lack of correlations[8,12,13]. Below this time scale, when analyzing high frequency data, correlations are, in general, no longer vanishing. Fig. (4) shows autocorrelation functions of returns for several real time series at a daily time scale, exhibiting the lack of memory/inertia in the time series.

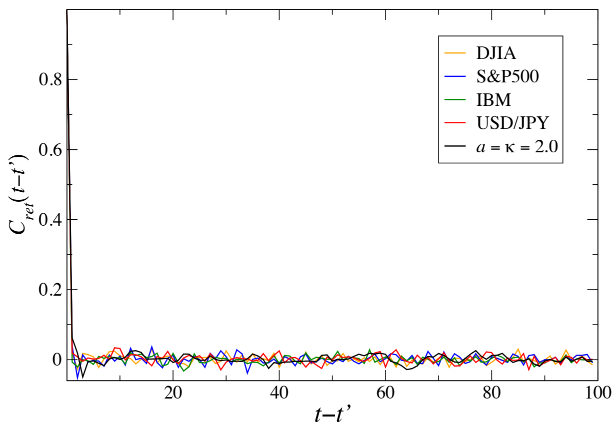

Figure 4: Autocorrelation functions of real financial time series of returns. We show the ACF of daily returns for the DJIA and S&P500 indices, IBM quote and the foreign exchange rate between the US dollar and the Yen. The black curve corresponds to the autocorrelation function of a time series obtained by Monte Carlo simulations of the microscopic model, with $a = \kappa = 2.0$ and a total of 2000 nodes. The graph clearly shows the lack of temporal correlations in the returns, both in real financial data (for daily returns) and the simulations Source of daily price histories used: *Yahoo! Finance*.

We shall define the autocorrelation function of the volatility as:

$$C_{v,\tau}(t-t') = \langle \tilde{v}_\tau(t) \tilde{v}_\tau(t) \rangle \tag{6}$$

This second order time correlation function does show memory in real financial time series. In fact, the long memory effect is typically of a scale invariant nature, meaning that there is no typical correlation time in the system, and therefore the autocorrelation function for the volatility behaves a power law (with some cut-off at the tail). The presence of long term correlations in volatility is a reflection of the volatility clusters that are ubiquitous in financial time series[8,12,13]. Fig. (5) shows this explicitly for several real time series.

IV. ISING MODELS AND MEAN FIELDS

The fundamental starting point for the opinion formation dynamics that we will later define is the statistical mechanics of spin 1/2 particle in contact with a heat bath (Ising model). Consider a single spin 1/2 in contact with a heat bath at temperature T in the presence of a magnetic field H. The particle has a state $\sigma = \pm 1$ that depends on the probabilities computed in the framework of the canonical ensemble.

Let E_+ and E_- be the energies of the particle when it has spin orientations $+1$ and -1, respectively. These energies arise from the interaction of the intrinsic magnetic moment ($\mu_\pm = \pm\mu_B$, where μ_B is the Bohr magneton), corresponding to $\sigma = \pm 1$, with the magnetic field H. Therefore, $E_\pm = -\mu_\pm H = \mp \mu_B H = \mp E$, where $E = \mu_B H$. The partition function for the particle is:

$$Z = e^{-\beta E_+} + e^{-\beta E_-} = e^{+\beta E} + e^{-\beta E} \tag{7}$$

where $\beta = 1/k_B T$ y k_B is Boltzmann's constant. Thus, the probability p_+ that the spin will have orientation $+1$ is given by the Boltzmann factor is given by[15,16]:

$$p_+ = \frac{e^{-\beta E_+}}{Z} = \frac{e^{\beta E}}{e^{\beta E} + e^{-\beta E}} = \frac{1}{1 + e^{-2\beta E}} \tag{8}$$

The probability, p_-, that the spin has a state -1 is $p_- = 1 - p_+$. If there are N spin 1/2 in the presence of a heat bath, the magnetization, M, of the system is given by :

$$M = \frac{1}{N} \sum_{i=1}^{N} \mu_B \sigma_i \tag{9}$$

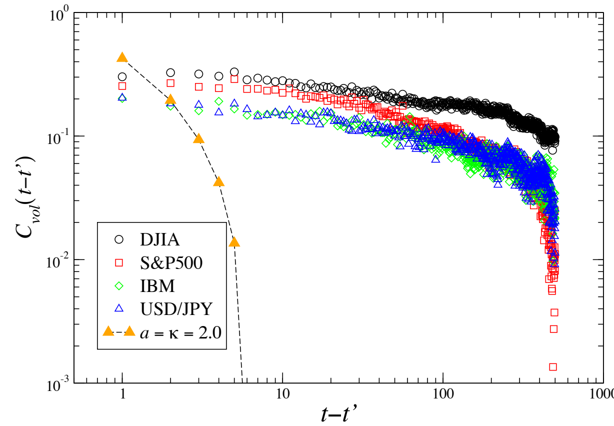

Figure 5: Autocorrelation functions of real financial time series of volatilites. We show the ACF of volatilities for the DJIA and S&P500 indices, IBM quote and the foreign exchange rate between the US dollar and the Yen. The dashed curve corresponds to the autocorrelation function of a volatility time series obtained by Monte Carlo simulations of the microscopic model, with $a = \kappa = 2.0$ and a total of 2000 nodes. The temporal lag $t - t'$ is shown in a daily scale for the real financial data. The graph shows the existence of time correlations of long range in the volatility of real financial time series. However, the volatility time series generated by microscopic model clearly fails to show long range memory, due to the model's inability of producing volatility clusters. Source of daily price histories used: *Yahoo! Finance*.

where σ_i is the state of the $i-$th spin. In what follows, we normalize the magnetic moment: $\mu_B \equiv 1$. Thus, $M = \sum_i \sigma_i/N$.

The Ising model proper does consider interactions between spins, which modify the total Hamiltonian of the system. Interaction energies or *exchange energies* are given by the Hamiltonian:

$$-\mathcal{H}(\sigma) = \sum_{\langle ij \rangle} J_{ij}\sigma_i\sigma_j + h\sum_i \sigma_i \tag{10}$$

for a spin configuration (σ), where h is the external field in units of the Bohr magneton ($h = H/\mu_B$) and J_{ij} is the magnitude of the coupling between spins i, j. The sum in the first term is taken over first neighbors (indicated by $\langle ij \rangle$). Here the spins occupy the nodes of a complex network and the first neighbors of the i- th spin are located in nodes which are linked to it. In terms of the Hamiltonian (10), the expression for the total partition function in terms of all possible spin configurations; $\{\sigma\}$; is given by:

$$\mathcal{Z} = \sum_{\{\sigma\}} \exp[-\beta\mathcal{H}(\sigma)] \tag{11}$$

Nevertheless, the opinion formation model we shall present, which is based on spin 1/2 systems, does not conform strictly to the Ising model for N spins. Instead, the model is defined as a sort of mean field approximation. For each spin, σ_i, there is a local field given by I_i. I_i is defined as a contribution of 2 terms: one global interaction term (due to the average spin in the network) and a local interaction term (due to the average of neighbouring spins). Thus the state of spin, σ_i, is computed in terms of the Hamiltonian of a one spin 1/2 particle in presence of an external magnetic field, which is given by I_i, requiring the use of the one particle partition function given by (7).

V. OPINION FORMATION MODEL OF BARTOLOZZI, LEINWEBER AND THOMAS

The microscopic opinion formation model that we will deal with has been proposed by Bartolozzi, Leinweber and Thomas[9]. It studies the influences of social interactions with first neighbors (local interactions) and with the society as a whole (global interaction) in the process of opinion formation of individuals. This discrete time model represents individuals as nodes on a scale free network of N nodes. The mechanisms of opinion formation are assumed to be stochastic heat bath dynamics with feedback. Each individual is forced to assume one of two possible states or opinions (spin orientations) at each time step.

Let $\sigma_i(t) = \pm 1$ represent the two possible states (opinions) of the i-th individual at time t. At each time step, the opinion of each individual is updated according to the following probabilistic prescription: $\sigma_i(t+1) = +1$, with probability p_i, and $\sigma_i(t+1) = -1$, with probability $1 - p_i$. These probabilities are calculated in analogy with the statistical mechanics of spin 1/2 particles in thermal equilibrium with a heat bath, whose temperature is formally defined as $T \equiv k_B^{-1}$, where k_B is Boltzmann's constant. Therefore, recalling (8), we have:[15,16]

$$p_i(t) = \frac{1}{1 + \exp[\,-2I_i(t)\,]} \tag{12}$$

Let $k_i = N_i$ be the number of first neighbors (degree) of the i -th node. The quantities $I_i(t)$ are calculated individually for every agent according to:

$$I_i(t) = a\xi(t)\frac{1}{N_i}\sum_{j=1}^{N_i} \sigma_j(t) + h_i\eta_i(t)r(t) \tag{13}$$

The quantities $\xi(t)$ and $\eta_i(t)$ are uniformly distributed random variables in the interval $[-1, 1]$, without any correlation in time or in the topological structure of the network.

The first term in (13) represents the degree of conviction with which the i -th node responds to the influences or decisions of his first neighbors. This is clear from the presence of the local average opinion over the N_i neighbors of node i. In the first term, $\xi(t)$ has the same value for every node in the network.

In the model, the parameter a is constant in time and is the same for all the network. It represents the strength or amplitude of the interactions or influences exerted by the first neighbors of the node in question and, therefore, it is referred to as the local interaction parameter.

The second term in (13) represents the degree of conviction with which the i -th node responds the global orientation or average opinion ($r(t)$) of the network. This global orientation of the network (i.e., the *"magnetization"*) is given by:

$$r(t) = \frac{1}{N}\sum_{i=1}^{N} \sigma_i(t) \tag{14}$$

In the global interaction term in Eq. 13, h_i is chosen individually for the i -th node as a uniformly distributed random variable in the interval $[0, \kappa]$. κ is constant in time and it represents the strength of the influences exerted by the global orientation of the network on the node; hence, it is referred to as the global interaction parameter.

Thus, a and κ are the fundamental parameters of the system and define the local and global degrees of influence, respectively.

As seen before, the microscopic model includes two sources of uncertainty in the dynamics of decision making. First, the opinion formation dynamics are explicitly stochastic, since the state of each individual changes according to a probabilistic rule. Second, the probabilities themselves are stochastic, since they depend on the random variables, $\xi(t)$, $\eta_i(t)$ and h_i. These multiple sources of stochasticity are justified by the authors of the model as necessary to mimic the intrinsic uncertainties and the complex human nature of opinion formation in social backgrounds.

In the context of finance, the nodes may be interpreted as *financial agents* and the two possible states for each agent $\sigma_i = +1[-1]$ represent intentions of buying or selling a unit of a certain financial asset, such as a share of stock.

The interactions and influences between agents, both locally and globally, are taken from the previously defined model. Thus, a financial agent is subject to the influence of his immediate environment (first neighbors) in deciding whether to buy or sell a share of stock. And he is also subject to the influence of the global orientation of the market as whole. The tendency to imitate his neighbors (or not) is affected by the amplitude and sign of the stochastic coupling constant $\xi(t)$, which controls the nature of the local interaction, making it ferromagnetic or antiferromagnetic.

The law of supply and demand is included in the model by defining the average opinion of the agents within the financial network ($r(t)$) as the return of the asset at time t. Recalling (2), we will take take the return between two successive times, in this discrete time microscopic model, to be:

$$r(t) = \frac{S(t+1) - S(t)}{S(t)} \tag{15}$$

It is in this sense that we say that the law of supply and demand is included in the model, since the average opinion in the network quantifies the demand for the asset at any given time, which moves the price of the asset.

Likewise, the normalized returns are now, according to (3):

$$\tilde{r}(t) = \frac{r(t) - \langle r \rangle}{\sigma} = \frac{r - \langle r \rangle}{\sqrt{\langle r^2 \rangle - \langle r \rangle^2}} \tag{16}$$

where σ is the standard deviation of the time series of prices and $\langle r_\tau \rangle$ is the average return. The volatility of the financial asset is defined locally in time as the absolute value of the returns:

$$\tilde{v}(t) = |\tilde{r}(t)| \tag{17}$$

The general behavior of a typical return distribution produced by this model is shown in Fig. 3, where it is clear that the real financial distributions compare quite well to the ones produced by the model, showing heavy tails. The distribution can be reasonably fitted by a Student's t- distribution which exhibits power laws in its tails.

VI. THE EFFECT OF TOPOLOGY

In this section, we explore the effect that the underlying topology of the network has on the return distributions of the simulated time series of the model. In particular, we generate networks of different topologies by varying the 'preferential attachment' probability in the Barabasi-Albert model by making it non-linear. The probability that a new node will attach to node i with degree k_i will now be

$$\Pi(i) = k_i^\alpha / \Sigma_j k_j^\alpha \tag{18}$$

where α is a tunable parameter that changes the attractiveness of nodes in the network to new nodes in the growth process. The original Barabasi-Albert model is recovered when $\alpha = 1$.

The effect of this *non-linear preferential attachment* on the degree distributions is shown in Fig. (6). Clearly the degree distribution is very sensitive to the exponent α[1]. The only case where a power law distribution is present is for $\alpha = 1$. For $\alpha < 1$ (sublinear regime), the degree distribution progressively loses its tail as α is reduced to 0 and hubs disappear from the network. In this case, $\alpha = 0$, the new nodes attach randomly to previously existing nodes and there is no preferential attachment and the degree distribution is a decreasing exponential function[1]: $P(k) = (e/m)\exp(-k/m)$.

For $\alpha > 1$ (superlinear regime), the degree distribution develops hubs of ever increasing degree as α increases. For $\alpha > 2$, the immense majority of the nodes have very small degree (equal to m) and are connected to very few *gel nodes*, which are hubs that are connect to the rest of the nodes in the network.

Now we examine the effect of the topology of the network on the response of the opinion formation model defined in the previous section by doing Monte Carlo simulations of the dynamics of the systems for different values of α. For each α, 100 different networks with $m = 5$ were generated and for each of them the dynamics of the model was simulated for 10000 time steps each with $a = \kappa = 2.5$. The return distributions obtained from the model are shown in Fig. (7).

From this figure, it is clear that system's dynamics are sensitive to the underlying topology of the network. The dependence of the shape of the return distribution on the topology is striking.

In the superlinear regime, the spread of the return distribution increases with increasing α, which seems to suggest that the presence of these very few but extremely highly connected individuals in the network is highly conducive to the presence of large scale synchronizations in the orientations of all the individuals in the network. Indeed, this can be seen in Fig. (8), where we plot the variance of the return distribution for different values of α.

In the sublinear regime, we see that there is little change in the variance of the distribution of returns. Indeed, the return distributions for $\alpha = 0.25$ and $\alpha = 1$ are qualitatively and quantitatively very similar to one another. So, at least in this range of values, the system's dynamics seems to be indifferent to the underlying network structure. However, there seems to be an important

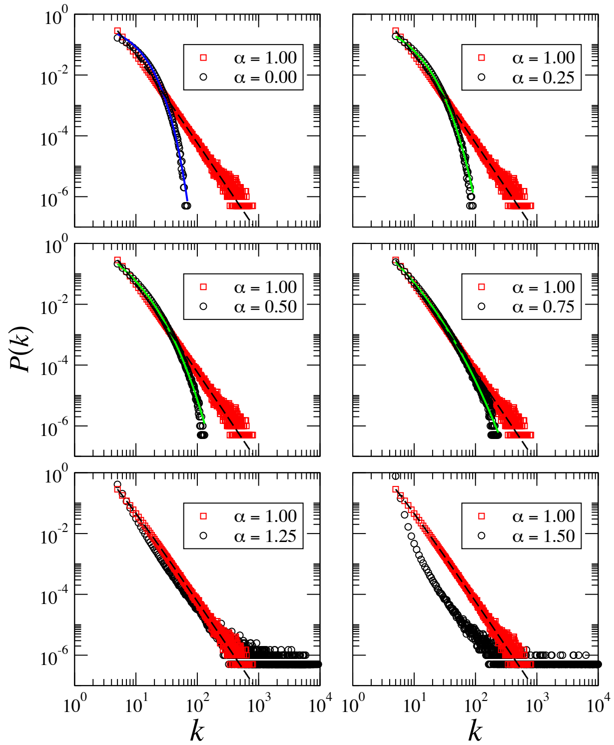

Figure 6: Degree distributions for different values of α. The red histograms correspond to the scale-free networks of the original Barabasi-Albert model $\alpha = 1$ and serve as a reference to compare with networks that exhibit non-linear preferential attachment (shown in black). For each α, 100 networks were generated with 20000 nodes each. The blue curve is an exponential distribution (theoretically predicted).

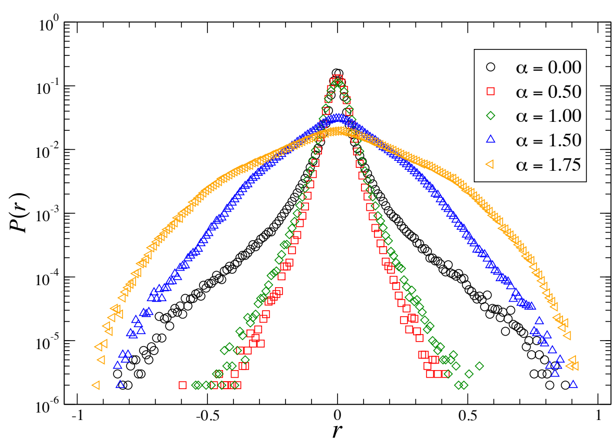

Figure 7: Return distributions for variations in the preferential attachment exponent (α).

qualitative and quantitative change in the return distribution in the case when $\alpha = 0$, as can be seen from Fig. (7).

We define the kurtosis of the distribution of returns

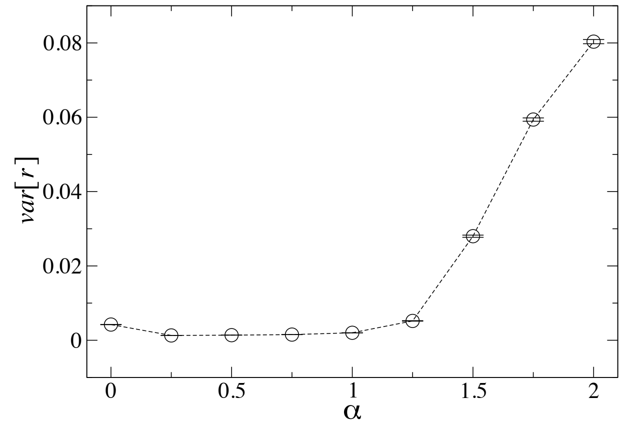

Figure 8: Variance of return distributions for variations in the preferential attachment exponent (α).

as:[8,12]

$$K_r = \frac{\langle r_\tau^4 \rangle}{\langle r_\tau^2 \rangle^2} \tag{19}$$

This is the fourth cumulant (also known as the *Binder cumulant*) of the distribution and it measures the weights of its tails. A Gaussian distribution has no heavy tails and its kurtosis is equal to 3. Distributions with heavier tails have greater kurtosis.

In Fig. (9)) we can see the kurtosis of the distribution plotted for various values of alpha. Again, we see that the return distribution is insensitive in most of the sublinear regime ($\alpha = 0.25$ and $\alpha = 1$). In the superlinear regime, the kurtosis drops to the level of a Gaussian, even though the variance increases. So, even though the probability of high fluctuations is much bigger than in the linear and sublinear regime, the tails decay exponentially fast. Consistent with what was said before, the $\alpha = 0$ case seems to be very strikingly different than the rest of the sublinear regime as far as the kurtosis is concerned.

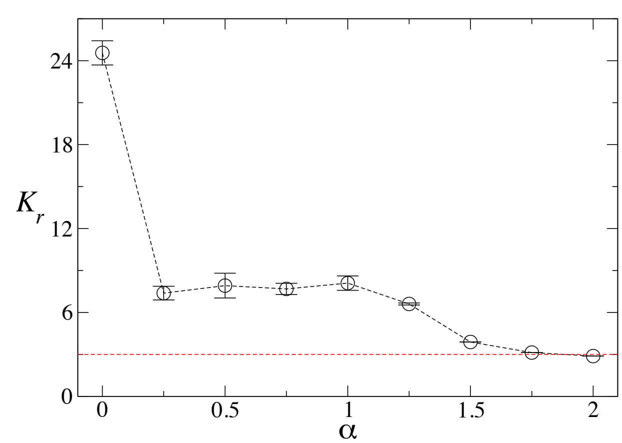

Figure 9: Kurtosis of return distributions for variations in the preferential attachment exponent (α). The red line corresponds to the evalue of the kurtosis of a Gaussian distribution.

These features can be appreciated if we plot the histogram of the absolute values of the normalized returns, shown in Fig. (10).

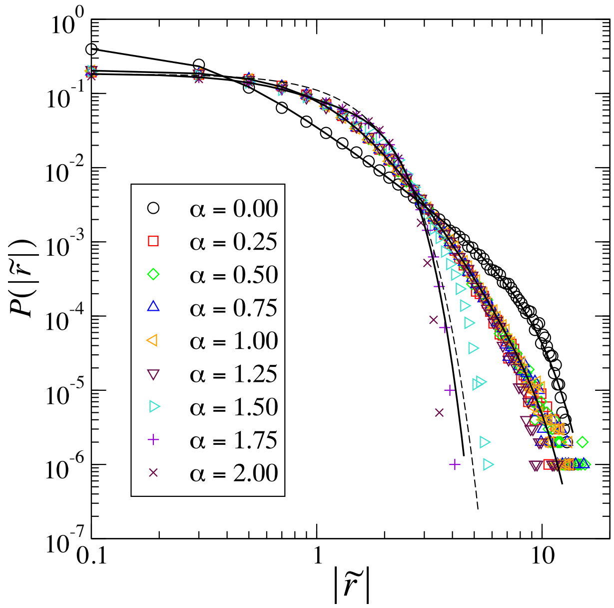

Figure 10: Distribution of absolute values of normalized returns for different values of (α). It is clear the distributions in the sublinear regime $\alpha = 0.25$ to $\alpha = 1$ are all very much insensitive to α. The $\alpha = 0$, however, is decidedly different. The superlinear regime shows distributions which resemble a Gaussian distribution. A Normal distribution is shown, for reference in dashed line.

VII. THE WEIGHTING FACTOR

The fields of *behavioral finance* and *social psychology* use the term *herding behavior* to describe tendencies of agents to imitate actions or decisions taken by other agents. This social phenomenon has been identified as one of the important factors that intervene in financial dynamics to explain the existence of bubbles and crashes in financial markets[14,17,18]. Individual investors are prone to buy stock for the simple reason of observing rising prices and noting that other investors that posses such assets are obtaining big returns. For these investors, it is not important that these price increments are justified or not by fundamental information. In fact, this *herding behavior* is not solely observed in the behaviors of individual non-professional investors.

It is also present in the world of professional investors, where Hedge-Fund managers are prone to imitate other fund managers and apply the same investment strategies in the same stock. According to Hong, Kubick and Stein, mutual fund managers are more inclined to invest in certain portfolios if other mutual fund managers, located in a certain geographic neighborhood, maintain similar portfolios[14,18]. These behaviors are consistent with an epidemic model in which investors spread or disseminate information regarding stocks and investment strategies in their close environment, often by spoken word.

According to studies made by Hong, Kubik and Stein, the most sociable individuals (those who are more prone to interaction with their neighbors and friends), are much more prone to invest in stock markets and maintain investment portfolios than non-sociable individuals. Studies by Shiller and Pound, regarding the behavior of individual investors and their stock-picking strategies, reveal that personal contacts such as friendships and family members play an important role in the decision making of such individuals. Recommendations of these personal contacts are often sufficient to motivate investors to imitate their investment strategies.

In addition to the effects of first-neighbors within the social networks, *behavioral finance* states that non-professional investors (whose number is rapidly increasing due to the accessibility of the Internet) are prone to the effects of certain specially influential agents in market dynamics, such as important personalities of investment institutions, known experts of technical and fundamental analysis, written or audiovisual media in their spreading of market news and of investment recommendations to the general public of investors[14,17,18].

These special agents possess great influence potential over numerous masses or groups of investors. The existence of these specially influential market agents provides the opportunity of introducing modifications or variants to the microscopic model (previously defined in section V) that take into account the inherent asymmetry of agents regarding their powers of influence. An agent with a lower capacity of influence must necessarily be taken into account with a different statistical weight than that of another agent of less influential power.

The specific way of quantifying these asymmetries of influence in the model of stochastic dynamics of opinion formation can be formalised by introducing, in the interaction fields, weighting factors based on the degrees of the nodes. We will define the weighting factors in such a way that greater importance is given to the opinions or orientations of those neighbors that posses greater degrees. The *weighting factor* of the opinion of the j-th agent, over the i-th, is defined as:

$$\omega_{ij}(\beta) = \frac{k_j^\beta}{\sum_{j'}^{N_i} k_{j'}^\beta} \tag{20}$$

where k_i is the degree of connectivity of the i-th node. The summation is taken over the N_i neighbors (links) of the i-th agent and the *ponderation exponent* (β) is a real positive number. In this variant of the original microscopic model, the interaction fields of financial agents are no longer the ones given by equation (13), which takes assumes uniform weights for the interaction with all the first neighbors of a given agent. Instead, we will take the interaction fields to be:

$$I_i(t) = a\xi(t)\sum_{j=1}^{N_i}\omega_{ij}(\beta)\sigma_j(t) + h_i\eta_i(t)r(t) \qquad (21)$$

where the summations are taken over the first neighbors of the i-th agent. Do note that the distribution of weights for the i-th node is locally normalized, so that $\sum_{j=1}^{N_i}\omega_{ij} = 1$. Consequently, for $\beta > 0$, the influence of high degree neighbors is greater than that of neighbors of low degree.

This is the modification that we introduce to the original model (which is recovered for $\beta = 0$).

We study the dependence of the return and volatility distributions with respect to the ponderation exponent. Specifically, we study the dependence of the exponent of the tails of the return distributions and the kurtosis and variance of the return distributions with respect to β.

For each value of β, we performed 100 Monte Carlo simulations of 10000 time steps each, which produced 100 different time series for returns, volatilities and prices. Each of the individual simulations utilized a different Bárabasi-Albert network with a total of $N = 2000$ nodes. The return and volatility histograms were constructed by taking into account all of the individual time series that were generated by the simulations for a same group of parameters. Since the scale-free network of each individual simulation is unique, it is reasonable to interpret each return and volatility histogram as the result of a unique financial system, whose network topology (with its respective adjacency matrix of connections between agents) is variable in time. The Barabási-Albert networks were constructed by choosing $N = 2000$ $m = m_0 = 5$. The model's parameters were fixed by choosing $a = \kappa = 2.1$.

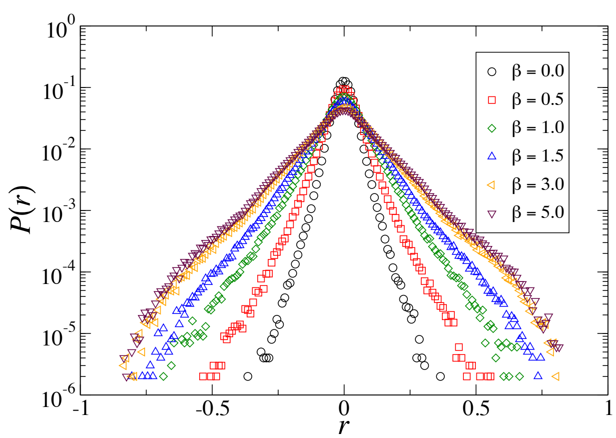

Figure 11: Return distributions for variations in the ponderation exponent (β).

The return distributions of the simulations are presented in Fig. 11. It is visually clear that the dispersion of the distributions increases as β increases. Thus, the introduction of preferential opinions within the financial network in terms of the degrees of the agents has a clear effect on the stochastic dynamics. This qualitative observation is quantified by the variance of the distributions, presented in Fig. 12, which increases correspondingly with the increments of β.

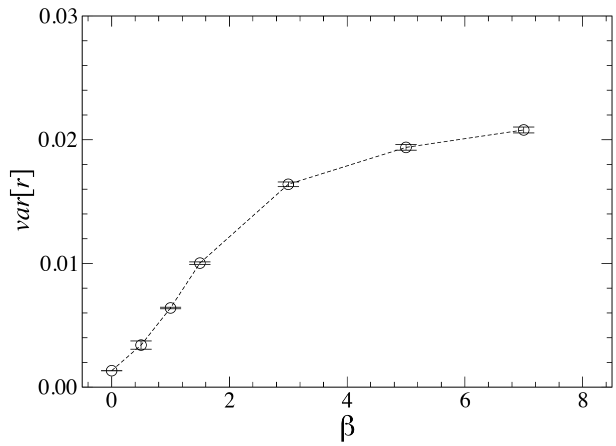

Figure 12: Variance of return distributions for variations in the ponderation exponent (β).

The weighting of opinions of agents within the network can be interpreted, in the context of systems of spin 1/2 particles, as the existence of different magnetic dipole moments for each particle. Awarding greater importance and influence to those nodes that have greater degree increases the probability of large scale synchronizations of opinions within the network and, consequently, increases the probability of occurrence of extreme events and the volatility of the time series of returns. This is a clear manifestation of *leadership behavior*, since the existence of special agents of great influence facilitates the tendencies of imitation of their behavior by their neighbors, leading to collective states large scale synchronization (i.e., big fluctuations).

If the returns are normalized according to equation (3), so that their time series exhibit zero mean and unit variance, their corresponding distributions show deviations with respect to a Student's t distribution in their tails, as can be seen in Fig. 13. These deviations in the tails, in the histograms for $\beta \neq 0$, are better quantified by measuring their respective kurtosis. It is easy to see that the kurtosis of the return distributions diminish with corresponding increments of the ponderation exponent, as can be seen in Fig. 14.

Do note that, although there is an important reduction of the kurtosis with increasing β, the corresponding return distributions do not show kurtosis near 3 and, consequently, they still remain in the leptokurtic regime. Recall that a Gaussian distribution has a kurtosis of 3; hence, the distributions obtained from the numerical simulations of the stochastic dynamics show heavy tails. This is understood visually in Fig. 13, where a normal distribution of zero mean and unit variance is shown.

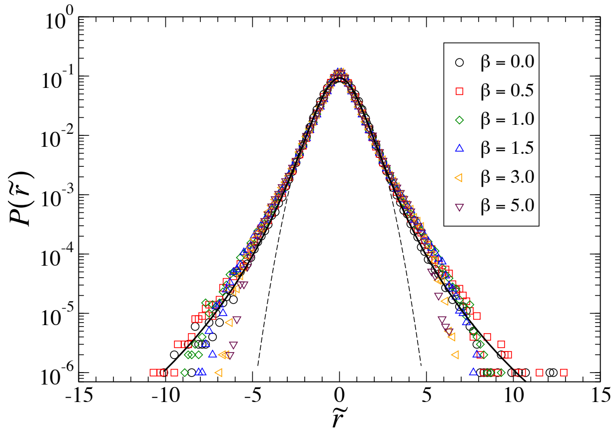

Figure 13: Distributions of normalized returns for variations in the ponderation exponent (β). As a reference, the dashed line represents a Gaussian distribution of unit variance and zero mean, corresponding to the prediction of Geometric Brownian Motion .

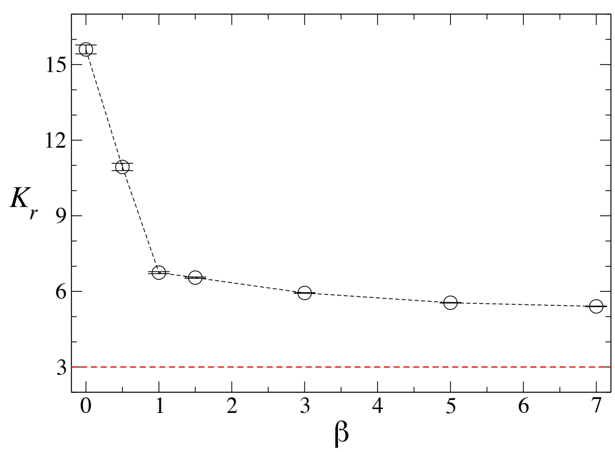

Figure 14: Kurtosis of return distributions for variations in the ponderation exponent (β). The red line shows the value of the kurtosis of a normal distribution (corresponding to the standard model of finance).

VIII. MEMORY AND VOLATILITY CLUSTERING

An aspect that we have not yet discussed about the model is the behavior of the autocorrelation functions of the time series of returns and the time series of the volatility. We shall now address this issue by looking at the autocorrelation functions of returns and volatilities, as defined in equations (5) and (6). It is clear from Fig. (4) that the typical time series produced by the original model produces returns that are not correlated in time, matching the behavior of empirical data and in accordance with the *Efficient Market Hypothesis*. This is a good feature of the model.

However, it is also clear from Fig. (5), that the volatility time series produced typically by the model do not match the long memory phenomena associated with volatility clustering that real financial data clearly show. Indeed, even though the model produces uncorrelated returns and long tails in the distribution of returns, it consistently fails to produce time series that exhibit volatility clustering and, therefore, a non-vanishing correlation time for the volatility.

We shall now introduce a change in the model with the purpose of having it produce volatility clusters. Several different attempts and ideas were implemented with this objective in mind. Of all the attempts made at changing the dynamics of the model, the only one that was capable of achieving volatility clustering consisted in providing each agent in the network with a memory. This memory allows them to remember the last global market changes (returns) within a certain finite time window ΔT in the immediate past (i.e., it allows the agents to remember the last ΔT returns of the asset in question) and it enters the model through the ponderation exponent β, previously defined in eqs. (20) and (21). Specifically, β now becomes a time dependent function, the dynamics of which are given by:

$$\beta(t) = \beta_{max} \frac{1}{\Delta T} \sum_{\tilde{t}=1}^{\Delta T} |r(t - \tilde{t})| \tag{22}$$

Equations (20) and (21) still define the dynamics of the system, but with a time dependent β. β_{max} is a tunable constant parameter.

As is clear from this new change to the dynamics of the model, the weights of the influences of neighboring agents will be functions not only of their degrees, but also of time. Therefore, the leaderships of the agents change in time in response to what happens in the market in the recent history. Thus, in a period of extreme fluctuations (i.e, a ‘crisis’ or a ‘bubble’), the leadership capabilities of the nodes with highest degree will be much greater than in calm periods, where influence capabilities are more likely to be democratically distributed.

The effects of these changes to the dynamics are now shown. Fig. (15) shows a time series of returns generated with this new model, compared to those obtained in the original model. It is clear that the introduction of this time dependence in β is capable of producing periods of volatility clusters, as claimed before.

Let us see how the system responds to different values of β_{max} for a time window of a fixed size. Fig. (16) shows that, as expected, the autocorrelation function of the volatility now exhibits memory. As is reasonable to expect, the bigger the value of β_{max}, the more pronounced the effect of the clusters. In the limit $\beta_{max} = 0$, we recover the memory absent model of the previous sections. As seen from the graphs, one can achieve typical correlation times of the order of

Now, if the β_{max} remains fixed and we change the size of the time window ΔT, we see the intuitively reasonable effect that the autocorrelation time of the system

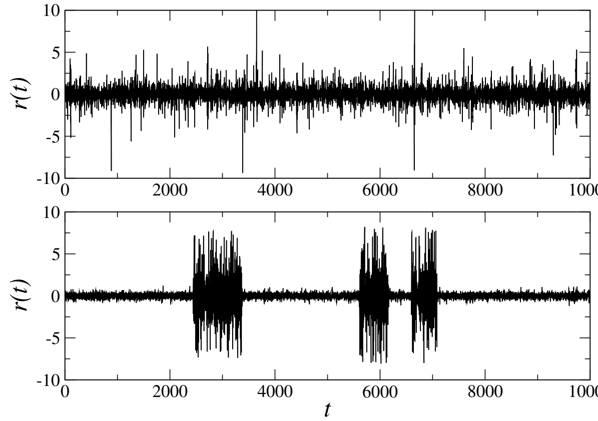

Figure 15: Time series of returns generated by original and modified model. The two graphs show time series of normalized returns generated by Monte Carlo simulations of the model with a 2000 node network, and $a = \kappa = 3$. The upper graph shows a time series that corresponds to the original model where opinions are not weighted. Clearly, large events of up to 10 standard deviations are observed, a clear indication of a heavy tailed return distribution, but these volatile instances are not clustered. The lower graph shows a time series that corresponds to the modified model where opinions are weighted and variable in time, with $\beta_{max} = 5$ and $\Delta T = 20$. Clearly, large fluctuations are still present, but they now appear clustered.

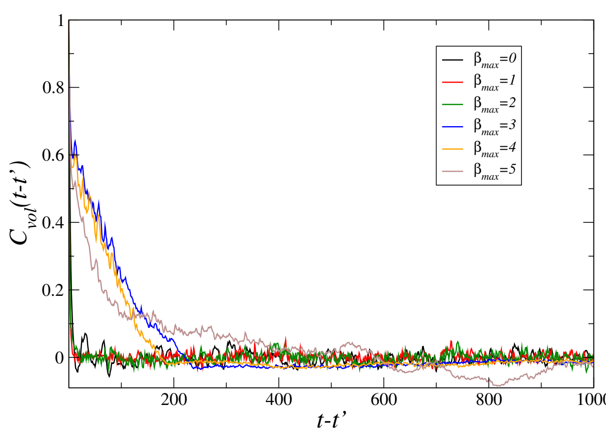

Figure 16: Autocorrelation functions of volatility time series generated by Monte Carlo simulations. Each curve is the ACF of the volatility time series produced by a particular simulation of a 2000 node system, with $a = \kappa = 4$ and a memory time window $\Delta T = 10$ for different values of β_{max}. When $\beta_{max} = 0$, we recover the original model, where opinions are not weighted. It is clear that, as expected, a non-vanishing ponderation exponent β_{max} produces memory in the volatility time series of the model.

increases, owing to a more pronounced of effect of volatility clustering. This can be seen in Fig. (17).

A reasonable idea for further exploration is to change the simple moving average approach used here and trying other kinds of dependencies on the past, giving more importance to the most recent events than to older events.

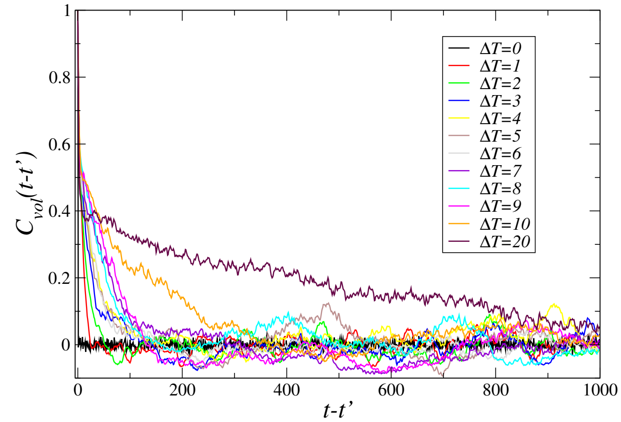

Figure 17: Autocorrelation functions of volatility time series generated by Monte Carlo simulations. Each curve is the ACF of the volatility time series produced by a particular simulation of a 2000 node system, with $a = \kappa = 4$ and $\beta_{max} = 5$, for different memory time windows ΔT. Intuitively, as the agents memory capabilities increase, so does the memory of the volatility.

This possible exploration of different time dependencies on the past history of the market is also necessary for the purpose of finding the one that is most appropriate in retaining the power law tails of the distribution of returns whilst conserving volatility clustering. This is important since the particular time series generated thus far, with the type of explicit time dependence on the past explained before, exhibit volatility clustering and return distributions with heavy tails, but said tails do not meet the requirement of having the particular form of a power law.

In any case, it is clear that providing each individual agent in the market (society) with an individual finite memory of the macroscopic state of affairs of the market (price history) is sufficient to produce memory in the volatility time series of the system by means of the presence volatility clusters.

IX. CONCLUDING REMARKS

In this project we have explored the effects of opinion weighting and network topology on the stochastic opinion formation model proposed by Bartolozzi, Leinweber and Thomas. The stochastic dynamics that rule the model's behaviour are sensible to the existence of asymmetries in the influences of agents, caused by weighting factors that affect the local fields of interactions by introducing differences in the powers of influence of each individual agent. Such weighting factors are defined in terms of the degrees of the nodes within the networks in such a way that nodes of high degree have greater influence than those of low degree.

As the value of the ponderation exponent (β) increases, the variance of the time series of returns increases correspondingly, suggesting that the existence of nodes of high influence serve the role of *leaders*, due to their abilities of increasing the frequency of occurrence of extreme events, associated with large scale synchronizations of opinions in the social network.

Allowing the agents to have memory of the market behavior within a certain time window is a mechanism that enables volatility clustering to appear in the time series of the volatilities, while retaining a lack of correlations in the returns. This has worked by not only allowing asymmetries in influences of agents, but also letting these asymmetries vary in time, becoming more relevant and pronounced in 'turbulent' phases.

The dynamics are sensitive to network topology as well. In the superlinear regime of pref. attachment, the variance increases, but the kurtosis decreases and the long tails disappear. In the sublinear regime, the distribution of returns remains insensitive over a very wide range of values of α. In the absence of preferential attachment, the distribution of returns changes very strongly and this merits further investigation.

[1] A. L. Barabási and R. Albert, Rev. Mod. Phys. **74**, 47 (2002).
[2] M. E. Newman, SIAM Rev. **45**, 167 (2003).
[3] S. N. Dorogotsev and J. F. Mendes, Advances in Physics **51**, 1079 (2002).
[4] R. P. Satorras, M. Rubi, and A. D. Guilera, *Statistical Mechanics of Complex Networks* (Springer-Verlag, Berlin (Germany), 2003).
[5] A. L. Barabási and E. Bonabeau, Sci. Am. **288**, 60 (2003).
[6] A. Chatterjee and B. K. Chakrabarti, *Econophysics of Markets and Business Networks* (Springer-Verlag, Milan (Italy), 2007).
[7] P. Sen, in *Econophysics and Sociophysics: Trends and Perspectives*, edited by B. K. Chakrabarti, A. Chakraborti, and A. Chatterjee (Wiley-VCH Verlag, Weinheim (Germany), 2006).
[8] J. Voit, *The Statistical Mechanics of Financial Markets* (Springer-Verlag, Berlin (Germany), 2005), 3rd ed.
[9] M. Bartolozzi, D. B. Leinweber, and A. W. Thomas, Physical Review E **72** (2005).
[10] A. Krawiecki, J. A. Holyst, and D. Helbing, Physical Review Letters **89** (2002).
[11] C. Castellano, S. Fortunato, and V. Loreto, Rev. Mod. Phys. **81** (2009).
[12] R. N. Mantegna and H. E. Stanley, *An Introduction to Econophysics: Correlations and Complexity in Finance* (Cambridge University Press, Cambridge (United Kingdom), 2000), 1st ed.
[13] J. P. Bouchaud and M. Potters, *Theory of Financial Risks: From Statistical Physics to Risk Management* (Cambridge University Press, Cambridge (United Kingdom), 2000), 1st ed.
[14] B. G. Malkiel, *A Random Walk Down Wall Street: The Time-Tested Strategy for Successful Investing* (W. W. Norton & Company, Inc., New York (United States), 2007).
[15] F. Reif, *Fundamentals of Statistical and Thermal Physics* (McGraw-Hill, Tokyo (Japan), 1965).
[16] M. E. Newman and G. T. Barkema, *Monte Carlo Methods in Statistical Physics* (Clarendon Press, Oxford (United Kingdom), 1999).
[17] R. J. Shiller, *Irrational Exuberance* (Princeton University Press, Princeton (United States), 2000), 1st ed.
[18] H. Hong, J. D. Kubik, and J. C. Stein, The Journal of Finance **60**, 2801 (2005).