# Stochastic Pair Approximation 

Antonio Fernández-Peralta ${ }^{1}$, Adrián Carro ${ }^{2}$, Raúl Toral ${ }^{1}$ and Maxi San Miguel ${ }^{1}$<br>${ }^{1}$ IFISC(CSIC-UIB), Instituto de Física Interdisciplinar y Sistemas Complejos, Campus Universitat de les Illes Balears, Palma de Mallorca, Spain<br>${ }^{2}$ Institute for New Economic Thinking at the Oxford Martin School, University of Oxford, Oxford, United Kingdom

Binary-state models on complex networks are used to describe the effect of interactions within a population of individuals. There is a broad spectrum of models that can depict competition between two opinions, language dynamics, the spreading of diseases, neural activity, etc. Analytical and numerical methods have been proposed capable of estimating the most important macroscopic quantities that portray the global state of the system [1]. However, most of the times, they are based on deterministic assumptions, where fluctuations and finite size effects are neglected. In this work, we relax these assumptions and we generalize the previous proposed methods, taking into account the full stochastic nature of the models. Specifically, we will focus on the noisy voter (Kirman) model, which is of particular relevance on the study of price fluctuations of financial markets. The model divides the population between optimistic (state 0) and pessimistic (state 1) agents. The agents are allowed to change state randomly (idiosyncratically) with probability rate $a$ or copying the state of a random neighbour agent in the network with probability rate $h$. The level of description of the method presented here includes as variables the number $L$ of active links (connecting a node in state 0 to another in state 1 ) and the total number $n_{k}$ of nodes that hold state 1 and have degree $k$ (number of neighbours). When an agent with $k$ neighbours, among which $q$ are in state 1 , changes its state, the global variables change like $n_{k} \rightarrow n_{k} \pm 1$ and $L \rightarrow L \pm(k-2 q)$. Using the pair approximation [2], we are able to obtain some effective rates for the processes mentioned above, that only depend on global variables $\left\{n_{k}\right\}, L$. The most problematic issue that one faces when dealing with the stochastic description of this type of models is the capability of closing the equations for the moments $\langle L\rangle,\left\langle n_{k}\right\rangle$, $\left\langle n_{k}^{2}\right\rangle, \ldots$ In order to overcome this issue, we propose a system size expansion $(N)$ like

$$
\begin{align*}
L & =N \phi_{L}(\lambda)+N \delta_{L}+N^{0} \nu_{L}  \tag{1}\\
n_{k} & =N \phi_{k}(\lambda)+N \delta_{k}+N^{1 / 2} \xi_{k}+N^{0} \nu_{k} \tag{2}
\end{align*}
$$

with a Van Kampen like part that involves $\delta_{k}$ (deterministic), $\xi_{k}$ and $\nu_{k}$ (stochastic), plus and additional non-trivial stochastic part $\phi_{k}(\lambda)$ that corresponds to a special deterministic trajectory which is an attractor of all the rest (where $\lambda$ is a stochastic variable that corresponds to the parametrization of the trajectory). The idea behind this peculiar expansion is that the usual Van Kampen expansion fails to reproduce correctly the results in the critical region of parameters, while Eqs.(1-2) perform well in the entire region. Within this approach, we are able to obtain analytical expressions for the stationary average of the order parameter $\langle L\rangle_{\mathrm{st}}$, the fluctuations of the total number of agents in state 1, i.e. $\sigma_{\mathrm{st}}^{2}[n]$, as well as the correlation function $\langle n(t+\tau) n(t)\rangle_{\mathrm{st}}$. The results are compared with the novel method [3], which makes use of an annealed network type of approximation, as well as,
with the pair approximation but neglecting fluctuations (deterministic), see Figure 1. The method presented here improves previous results and it is a step forward in the characterization of stochastic effects in binary-state models on complex networks.


Figure 1: Stationary average density of active links $\langle\rho\rangle_{\mathrm{st}}=$ $2\langle L\rangle_{\mathrm{st}} / N\langle k\rangle$ (upper panel) and variance of the total number of agents in state 1 (lower panel). Dots correspond to numerical simulations of three different networks, Erdös-Rényi (ER), Barabási-Albert (BA) and Dichotomous (DEL), with $N=2500$ and $\langle k\rangle=4$. Colored solid lines are the stochastic pair approximation prediction, the black solid line is the deterministic pair approximation, while dashed colored lines are the annealed network approximation [3].
[1] James P. Gleeson. Phys. Rev. X. 3, 021004 (2013).
[2] F. Vazquez and V. M. Eguíluz. New J. Phys. 10, 063011 (2008).
[3] A. Carro, R. Toral and M. San Miguel. Sci. Rep. 6, 24775 (2016).

