

On the Emergent Behavior of the 2-Choices Dynamics

Emilio Cruciani¹, Emanuele Natale²,
André Nusser³, and Giacomo Scornavacca⁴

¹ Gran Sasso Science Institute, L'Aquila, Italy. emilio.cruciani@gssi.it

² Max Planck Institute for Informatics, Saarland Informatics Campus, Germany
& Simons Institute, Berkeley, California. enatale@mpi-inf.mpg.de

³ Max Planck Institute for Informatics, Saarland Informatics Campus, Germany.
anusser@mpi-inf.mpg.de

⁴ University of L'Aquila, L'Aquila, Italy. giacomo.scornavacca@graduate.univaq.it

Abstract. This short communication presents two recent results [13,14] that study the emergent behavior of the 2-CHOICES dynamics, a simple, local, and non-linear stochastic process on networks. We show that the 2-CHOICES dynamics can be exploited as an efficient distributed algorithm for graph clustering and as a model to explain phenomena in the contexts of sociology, biology, and neuroscience.

Keywords: 2-Choices Dynamics; Emergent Behavior; Consensus; Metastability; Distributed Computing; Label Propagation Algorithm

Introduction

Dynamics are simple stochastic processes on graphs in which nodes update their own state according to a symmetric function of the state of their neighbors and of their current state, with no dependency on time or on the topology of the graph [27,26]. In previous decades the computational power of this kind of systems has been investigated by computer scientists, mathematicians, and physicists. Recently it got a renewed interest from the theoretical computer science community, as new algorithmic techniques make it possible to analyze dynamics as distributed algorithms [15,3,10,11,5,4].

In the context of distributed computing one fundamental task is that of reaching a *consensus*, i.e., a configuration where all nodes have the same state. In this regard, the “trivial” example of dynamics is the so-called VOTER dynamics, in which in each round each node copies the opinion of a random neighbor. Despite its apparent simplicity, some properties of this model have been proven only recently [23,20,7]. The VOTER dynamics is characterized by the property of reaching a *proportionate consensus*¹ in polynomial time independently from the network topology [25]. However, the convergence time is $\Omega(n)$ even on networks with small diameter such as expander and complete graphs.

¹ The probability to converge to a state is proportional to the initial volume of that state, i.e., to the sum of the degrees of the nodes supporting that state [20].

The simplest non-trivial example of dynamics is arguably the 2-CHOICES dynamics, where at each discrete time step each node samples two random neighbors with replacement and, if they have the same state, the node adopts that state. Prior to our works [13,14], the 2-CHOICES dynamics has been only analyzed on networks with good expansion properties, where it exhibits the same emergent behavior of the VOTER dynamics, but converges very quickly [10,11].

Phase Transition on Core-Periphery Networks

We study the behavior of the 2-CHOICES dynamics theoretically and empirically on *core-periphery networks* [13]. These networks, well-known in social network analysis, sociology, and economics [9,17], are characterized by a bipartition of the nodes into a *core*, a small dense set of nodes that dominates the network, and a *periphery*, the loosely connected remaining nodes. We consider an axiomatic framework [1] which is based on two parameters only, *dominance* and *robustness*, to express the influence of the core on the periphery.² We consider a *natural* scenario, where agents in core and periphery initially have different states, and show the following *phase transition* phenomenon:

Theorem 1. *There exists a universal constant c^* such that: if the dominance is greater than c^* and the core is robust enough, then a configuration of almost-consensus is reached in $\mathcal{O}(\log n)$ rounds, with high probability;³ if the dominance is less than c^* , then a metastable phase takes place where most of the nodes retain their initial opinion for $n^{\omega(1)}$ rounds, with high probability.*

We validate our theoretical results with experiments on real-world networks. The experiments show some weaknesses in the heuristic used to identify the core in [1]. We design a new heuristic able to identify a core with a volume approximately equal to that of the periphery by repeatedly calculating densest-subgraph approximations; with the VOTER dynamics this is sufficient to have equal probability of consensus for both states and prior results on the 2-CHOICES dynamics [10,11] require an initial biased configuration of the states to give results on the consensus; our experiments show that the consensus is almost always reached on the core's state. We believe these results may be useful for understanding the principles that drive social and economic agents into forming core-periphery networks.

Metastability on Clustered Graphs and its Biological Implications

The long metastable phase that often occurs can be exploited in the context of *Label Propagation Algorithms* (LPAs, started by [28]), a widely used class of heuristics for graph clustering that matches the following general pattern:

1. Initially, each node is assigned a label, typically independent from the others and uniformly at random.
2. Then, each node starts updating its current label according to a simple *local* majority-based rule based on the labels of its neighbors.

² We refer the reader to [13] for a formal definition of these parameters.

³ Our analysis shows that the process does not deviate significantly from its expected evolution *with high probability*, i.e., with probability at least $1 - \mathcal{O}(n^{-c})$.

3. After a relatively short number of rounds, the labeling of the graph is expected to stabilize and is declared to be the clustering of the graph.

Most of the work on label propagation algorithms has been empirical, lacking formal arguments that show their ability in the detection of clusters. The only rigorous analysis present in literature is the one of MAX-LPA [21], where each node updates its label to the most frequent label among all its neighbors. Compared to the 2-CHOICES dynamics, MAX-LPA has a higher cost in the total number of messages exchanged in each round. For the first time we analyze a *sparsified* LPA, where every node chooses a random initial state out of two possible ones and executes the 2-CHOICES dynamics [14]. In the following we report the definition of the class of graphs considered for the analysis and our main theorem.

Definition 1 ([6]). A $(2n, d, b)$ -clustered regular graph is defined as a graph $G = (V_1 \cup V_2, E)$ such that: $|V_1| = |V_2| = n$; every node has degree d ; every node in V_1 has b neighbors in V_2 and every node in V_2 has b neighbors in V_1 .

Theorem 2. If the 2-CHOICES dynamics is initialized in a slightly asymmetric way w.r.t. the clusters,⁴ then with constant probability the network rapidly converges to a metastable phase in which almost all nodes within each cluster share the same state and the predominant states of the clusters are different.

The states of the nodes in the metastable phase constitute a labeling which reveals the clustered structure of the network. The probability to converge to such a phase can be amplified via Community-Sensitive Labeling [2]. Our analysis exploits some spectral properties of the transition matrices of simple random walks on clustered regular graphs and, compared to that of MAX-LPA, it considers much sparser communities at the price of a stricter condition on the cut.

Evolutionary Biology. From our theoretical results in [14], it is possible to derive some biological implications. Evolutionary dynamics is the branch of genetics which studies how populations evolve genetically as a result of the individuals' interactions [16]. Lieberman et al. initiated the study of evolutionary dynamics on graphs by investigating the *fixation probability* of the *Moran process*, namely the probability that a new mutation with increased fitness eventually spreads across all individuals in the population [22]. However, no simple dynamics has been proposed in the context of evolutionary graph theory to explain one of evolution's fundamental phenomena, namely *speciation* [12].

Two fundamental classes of driving forces for speciation can be distinguished: *allopatric* speciation and *sympatric/parapatric* speciation. The former, which refers to the divergence of species resulting from geographical isolation, is well understood [29]; the latter, namely divergence of species without complete geographical isolation, is still controversial [29,8]. In several evolutionary settings, the spread of a mutation appears nonlinear w.r.t. the number of interacting individuals carrying the mutation, exhibiting a drift towards the most frequent

⁴ For example, if every node choose the initial state according to the flip of a fair coin.

phenotypes [19,12]. We look at the 2-CHOICES dynamics on clustered graphs as a quadratic evolutionary dynamics on sympatric and parapatric scenarios. The random initialization of the 2-CHOICES dynamics can be regarded as two inter-mixed populations of individuals with different genetic pools. The interactions for reproduction between the two populations can be categorized in frequent interactions among individuals within an equal-size bipartition of the populations (the clusters) and less frequent interactions between the two populations. This can be interpreted as genetic admixture, i.e., interbreeding between two genetically-diverging populations [24]. Within such framework, our theorem provides an analytical evolutionary graph-theoretic proof of concept on how speciation can emerge from the simple nonlinear underlying dynamics of the evolutionary process on the population level.

Neuroscience. During mammalian development, neuromuscular junctions and some other postsynaptic cells transition from having multiple neurons innervating onto them into having a single neuron only [18,30]. This process takes place as synaptic sites are exchanged between different axons: Locations on the surface of the cell on which neurons innervate transition from being freed by the current innervating neuron to be occupied again by a new one [30]. In [30], a simplistic model for the aforementioned process has been provided, which is equivalent to the VOTER when the underlying graph is regular. In [14], we argue that our analysis provides evidence for the fact that, in order for a model based on dynamics to comply with experimental evidence on the outcome of the innervation process, either the innervation sites do not exhibit spatial bottlenecks or the dynamics cannot be based on majority-like mechanisms.

References

1. Avin, C., Lotker, Z., Peleg, D., Pignolet, Y.A., Turkel, I.: Core-Periphery in Networks: An Axiomatic Approach. arXiv:1411.2242 [physics] (2014)
2. Becchetti, L., Clementi, A., Manurangsi, P., Natale, E., Pasquale, F., Raghavendra, P., Trevisan, L.: Average whenever you meet: Opportunistic protocols for community detection. arXiv preprint arXiv:1703.05045 (2017)
3. Becchetti, L., Clementi, A., Natale, E., Pasquale, F., Silvestri, R.: Plurality consensus in the gossip model. In: 26th Symposium On Discrete Algorithms (2015)
4. Becchetti, L., Clementi, A., Natale, E., Pasquale, F., Silvestri, R., Trevisan, L.: Simple dynamics for plurality consensus. Distributed Computing (2017)
5. Becchetti, L., Clementi, A., Natale, E., Pasquale, F., Trevisan, L.: Stabilizing consensus with many opinions. In: 27th Symposium On Discrete Algorithms (2016)
6. Becchetti, L., Clementi, A., Natale, E., Pasquale, F., Trevisan, L.: Find your place: Simple distributed algorithms for community detection. In: 28th Symposium On Discrete Algorithms (2017)
7. Berenbrink, P., Giakkoupis, G., Kermarrec, A.M., Mallmann-Trenn, F.: Bounds on the Voter Model in Dynamic Networks. In: 43rd International Colloquium on Automata, Languages, and Programming (2016)
8. Bolnick, D.I., Nosil, P., Servedio, M.: Natural selection in populations subject to a migration load. *Evolution* (2007)
9. Borgatti, S., Everett, M.: Models of core/periphery structures. *Social Networks* (2000)

10. Cooper, C., Elsässer, R., Radzik, T., Rivera, N., Shiraga, T.: Fast consensus for voting on general expander graphs. In: 29th International Symposium on Distributed Computing, DISC 2015. pp. 248–262 (2015)
11. Cooper, C., Radzik, T., Rivera, N., Shiraga, T.: Fast plurality consensus in regular expanders. In: 31st International Symposium on Distributed Computing, DISC 2017. pp. 13:1–13:16 (2017)
12. Coyne, J.A., Orr, H.A.: Speciation. Sinauer Associates is an imprint of Oxford University Press (2004)
13. Cruciani, E., Natale, E., Nusser, A., Scornavacca, G.: Phase transition of the 2-choices dynamics on core-periphery networks. In: 17th Conference on Autonomous Agents and MultiAgent Systems (2018)
14. Cruciani, E., Natale, E., Scornavacca, G.: Rigorous analysis of a label propagation algorithm for distributed community detection. In: 33rd AAAI Conference on Artificial Intelligence (2019)
15. Doerr, B., Goldberg, L.A., Minder, L., Sauerwald, T., Scheideler, C.: Stabilizing consensus with the power of two choices. In: 23rd ACM Symposium on Parallelism in Algorithms and Architectures (2011)
16. Durrett, R.: By Richard Durrett - Probability Models for DNA Sequence Evolution: 2nd (second) Edition. Springer-Verlag New York, LLC (2011)
17. Easley, D., Kleinberg, J.: Networks, Crowds, and Markets: Reasoning About a Highly Connected World. Cambridge University Press (2010)
18. Gan, W.B., Lichtman, J.W.: Synaptic Segregation at the Developing Neuromuscular Junction. Science (1998)
19. Gavrillets, S.: Perspective: models of speciation: what have we learned in 40 years? Evolution; International Journal of Organic Evolution (2003)
20. Hassin, Y., Peleg, D.: Distributed Probabilistic Polling and Applications to Proportionate Agreement. Information and Computation **171**(2), 248–268 (2001)
21. Kothapalli, K., Pemmaraju, S.V., Sardeshmukh, V.: On the analysis of a label propagation algorithm for community detection. In: 14th International Conference on Distributed Computing and Networking (2013)
22. Lieberman, E., Hauert, C., Nowak, M.A.: Evolutionary dynamics on graphs. Nature (2005)
23. Liggett, T.M.: Stochastic Interacting Systems: Contact, Voter and Exclusion Processes. Springer Science & Business Media (2013)
24. Martin, S.H., Dasmahapatra, K.K., Nadeau, N.J., Salazar, C., Walters, J.R., Simpson, F., Blaxter, M., Manica, A., Mallet, J., Jiggins, C.D.: Genome-wide evidence for speciation with gene flow in *Heliconius* butterflies. Genome Research (2013)
25. Mertzios, G.B., Nikolettseas, S.E., Raptopoulos, C.L., Spirakis, P.G.: Determining majority in networks with local interactions and very small local memory. Distributed Computing **30**(1) (2016)
26. Mossel, E., Tamuz, O.: Opinion exchange dynamics. Probability Surveys (2017)
27. Natale, Emanuele: On the Computational Power of Simple Dynamics. PhD Thesis, Sapienza University of Rome (2017)
28. Raghavan, U.N., Albert, R., Kumara, S.: Near linear time algorithm to detect community structures in large-scale networks. Physical review E (2007)
29. Savolainen, V., Anstett, M.C., Lexer, C., Hutton, I., Clarkson, J.J., Norup, M.V., Powell, M.P., Springate, D., Salamin, N., Baker, W.J.: Sympatric speciation in palms on an oceanic island. Nature (2006)
30. Turney, S.G., Lichtman, J.W.: Reversing the Outcome of Synapse Elimination at Developing Neuromuscular Junctions In Vivo: Evidence for Synaptic Competition and Its Mechanism. PLOS Biol (2012)