

¹ A game-theory modeling approach to utility and
² strength of interactions dynamics in biomedical
³ research social networks

⁴ J. Mario Siqueiros-García Rodrigo García-Herrera
⁵ Enrique Hernández-Lemus Sergio A. Alcalá-Corona

⁶ November 19, 2016

⁷ **Abstract**

⁸ What would happen if researchers in a community were given the
⁹ same amount of resources and then set free to interact among each
¹⁰ other? How would these resources be distributed in the population
¹¹ after many interactions? The final result may be determined by the
¹² researchers' social and political power, and by their prestige. Never-
¹³ theless, all these attributes depend, to a certain degree, on the research
¹⁴ collaboration social network in which scientists are embedded. Here
¹⁵ we develop a model based on the Prisoner's Dilemma implemented
¹⁶ on a collaboration network. With this model we explore the distribu-
¹⁷ tion of resources and the dynamics of the strengthening or weakening
¹⁸ of collaboration interactions among partners, as scientists cooperate
¹⁹ or defect in order to gain access to resources shared by others. The
²⁰ network is based on data about projects that were awarded with a
²¹ grant. In this sense, we assume that resources are funds, equipment
²² and time. We define Fitness as how much access a researcher has to the
²³ resources of his colleagues. We tested our simulation on the real biomedical
²⁴ research network and compared the results with an Erdős-Renyí, a
²⁵ Watts-Strogatz small-world and Barabási-Albert topologies. Different
²⁶ topologies display different fitness and connections strength distribu-
²⁷ tions. Moreover, the distribution of fitness and connections strength
²⁸ in the researchers network is similar to that of Barabási-Albert and
²⁹ Watts-Strogatz topologies, respectively. We believe that fitness distri-
³⁰ bution in the researchers network suggests that there are socio-cultural
³¹ mechanisms governing the network that produce an asymmetric distri-
³² bution of resources. The high distribution of strong connections might
³³ reflect some sort of subordination among researchers by which they are
³⁴ morally obliged to cooperate by the same socio-cultural mechanisms.

35 The range around the threshold that regulates the decision to cooperate
36 or defect according to the agent's historical balance between utility
37 and strength of collaborative relationships and carrying capacity of the
38 system is small, suggesting that there is a region in which a phase transition
39 takes place from a population of cooperators to a population of
40 defectors. Simulations like this may help to develop science policies to
41 promote fair distribution of resources.

42 1 Introduction

43 Collaboration has become a cornerstone in biomedical research today. In
44 contrast to physics which has a long history and experience in collaborative
45 projects, biology is only recently becoming an evermore collaborative
46 discipline[1]. Biology has an interesting record in such matters because sci-
47 entific collaboration means something different to different branches of bi-
48 ology: molecular biology has traditionally been a research activity of small
49 laboratories[2, 3], whereas in natural history there has been data and sam-
50 ples exchange since the *XVIIth* century[4, 5]. Despite the differences in
51 culture and practices, the Human Genome Project made collaboration a
52 central feature of biology.

53
54 Nowadays it is widely acknowledged that collaboration takes many forms,
55 from sharing of biological samples and biobanking to international groups in
56 charge of helping research communities to harmonize and share their data.
57 Sharing resources such as equipment, funds, and time is critical; building
58 trust among scientists is fundamental. Also, resources are mobilized in order
59 to create strategic alliances.

60
61 The analysis of cooperation in scientific research has been the subject of
62 a number of studies [1, 7, 8, 9, 10, 3, 5]. This is not surprising since coopera-
63 tion and competition are quite important in today's academic success. How
64 does collaboration happen within a competitive academic environment and
65 what kind of payoff is present in these settings were questions considered
66 recently by Wardil and Hauert [20] in the context of cooperation in multi-
67 authored publications. Also, the role of game theory over complex scientific
68 information and collaboration networks has attracted attention, mainly fo-
69 cusing on how long-term strategies may shape different scenarios for Nash
70 equilibria [21]. Prisoner's Dilemma has been used in the study of impact
71 factor and collaboration [13, 14].

72

73 Even with all these research efforts, cooperation in the context of scientific
74 collaboration is still loosely defined and the long term dynamics of academic cooperation (and its consequences) are yet to be fully elucidated.
75 Furthermore, to our current knowledge, there has been no use of game theory and complex network analysis for understanding how the topology of scientific collaboration networks affects access to resources among individuals present in the network ¹. Our work aims to contribute to our current understanding on the matter, specially when agents have to maximize their access to resources while taking care of their collaboration links.

82

83 In this article we explore the network effect on the distribution of players having access to certain amount of resources shared by other players in the network and the distribution of the strength of interactions among them. Particularly, we implemented two games played simultaneously: one for maximizing individual utility based on the iterated Prisoner's Dilemma; the other, a coordination game for maximizing the connection strength between players. We are interested in how they affect each other in the context of a network of scientific collaboration under the idea that while researchers are interested in maximizing their utilities, they also know that it is important to invest in building collaborative relationships. These two behaviors are explored in a biomedical research community of México.

94

95 In the context of our paper, utility represents access to resources shared by others. The value of the Utility function for a player is the sum of the payoffs of playing with its neighbors. The opossing force comes from the other concurrent game: players trying to maximize the strength of their interactions with other players. In the coordinating game the best strategy is to adopt the same strategy as the other player, as it pays the most regardless of cooperation or defection in the utility game. When both cooperate the interaction gets a positive payoff, when both defect, the interaction doesn't get affected; but if they anti-coordinate, then the interaction looses. Finally, cooperation is a central feature of scientific work. For our biomedical network, cooperation can be thought of as sharing resources such as time, students, equipment, even money. Examples of defection to a cooperator are ghost authorship or prestige authorship.

108 The manuscript is structured in five sections. First we describe *FOS-ISS*, the main program for grants destined to biomedical applied research in México. This is the source of the database from which we created the

¹For an account of scientific collaboration and definitions, please refer to [6].

111 researchers collaboration network. Next we describe our model and the dif-
112 ferent network topologies on which we explored it. We then present our
113 results and discuss them. In the last section we draw some final remarks
114 and conclusions.

115 **2 Biomedical research: CONACyT and FOSISS**

116 CONACyT (National Council of Science and Technology) is the Mexican
117 government entity in charge of promoting the development of science and
118 technology. Among CONACyT's functions are to develop science and tech-
119 nology policies according to national needs and demands, to advise the dif-
120 ferent instances of government on scientific and technological topics, to pro-
121 mote the creation of research networks among the scientific community, to
122 grant scholarships for masters and doctoral studies, and to manage different
123 trusts intended to fund individuals and groups for scientific and technologi-
124 cal research.

125

126 In the year 2002 CONACyT, along with other government agencies and
127 entities, created sectoral funds to cover and equally promote research capac-
128 ities of different areas such as energy, agriculture and health. Technological
129 innovation is fostered by the generation of human resources and by helping
130 research groups to consolidate. It is expected that the knowledge generated
131 under the sponsorship of these funds will be the product of applied research
132 that attends national public needs, and promotes economic growth.

133

134 *FOSISS* or Sectoral Fund for Health and Social Security Research (*Fondo*
135 *Sectorial en Investigación en Salud y Seguridad Social*) is one of such funds.
136 FOSISS is constituted by CONACyT, SSA, IMSS and ISSSTE,² all of them
137 being the major public health providers and research institutions in the
138 country. Every year CONACyT opens a call for funds limited to a set of
139 health research areas previously defined by a group of experts. Such areas
140 range from public health issues to chronic and degenerative diseases.

141

142 Most applicants are public universities and research institutions, but el-
143 igibility is open to public and private health research sectors. From 2002

²SSA is the acronym for Secretariat of Health *Secretaría de Salud*; IMSS is the acronym for Social Security Mexican Institute (*Instituto Mexicano del Seguro Social*); ISSSTE stands for Institute for Social Security and Services for State Workers (*Instituto de Seguridad y Servicios Sociales de los Trabajadores del Estado*)

144 through 2013, there were 91 institutions funded that comprised 4988 re-
145 searchers.

146

147 From these data some important considerations should be made clear.
148 Scientists in the database take on the roles of principal investigators (PIs),
149 associate researchers, postdoctoral associates, postgraduate and undergrad-
150 uate students. Unfortunately, information on these roles is not specified in
151 the database. We acknowledge the importance of this deficiency because
152 researchers in our network act under different circumstances and we know
153 that this diversity has a real impact on the structure and eventually on the
154 dynamics of the network, as well as on the results of our model.

155

156 Our database includes the name of the project, the year it was approved
157 for funding and the research area to which it was assigned. It specifies the
158 names of PIs or the people responsible for the project and the names of
159 collaborators. Researchers can be PIs in one project and collaborators on a
160 different project. The institutional affiliation of all participants is included.
161 Through this affiliation we determine the principal institution behind every
162 project.

163

164 Even though curation and analysis work of this database is still going
165 on, some relevant facts about the biomedical research can be said. Over the
166 period of 12 years, 32 general research areas have been defined, the three
167 most funded research areas are *chronic and degenerative diseases, malignant*
168 *neoplasms, and infectious and parasitic diseases*. The least funded area is
169 *Ethics and medicine*. The area with the most researchers is *malignant neo-*
170 *plasmas*. Other areas of relevance for México are *diseases related to poverty*
171 and *Health and vulnerable groups*.

172

173 From the institutions that have participated in a protocol funded by
174 FOSISS, less than one fifth have been responsible for a project and more
175 than 95% of them are Mexican, public institutions. There is also an impor-
176 tant presence of foreign institutions as collaborators, most of them from the
177 United States, though institutions from the UK, France, Spain, Netherlands,
178 Colombia and Cuba are also in the database.

179

180 Besides the characteristics of the population there are some other bound-
181 ary conditions that play an important role on the network topology and dy-
182 namics, that motivated the development of our model. Biomedical research
183 in México constitutes a vibrant community and collaboration is part of ev-

184 everyday work. However, México does not have public biobanks for research
185 purposes (which are specially relevant for research in genomics, for exam-
186 ple), there is no regulation on the access to biological samples such as tissue,
187 cells, DNA, RNA, etc.³ Something similar happens with data. There have
188 been some attempts to create open data repositories for biomedical research,
189 but they have not been established yet. Regulation on these subjects is still
190 missing. Finally, technologically advanced equipment such as high through-
191 put sequencers are kept by institutions with the highest research profiles
192 and sometimes PIs manage them in a self serving way.

193

194 From our ethnographic work to date, we have been able to see that bio-
195 logical samples, data and technology can become instruments for negotiat-
196 ing collaboration. For example, among people involved in research projects,
197 there are researchers that do not have direct access to samples, simply be-
198 cause their parent institution does not offer clinical services. Many of them
199 are non medical doctors but chemists, biologists, physicists, and mathemati-
200 cians. There is another group of researchers that are placed on hospitals
201 which is able to do research and have access to biological samples from their
202 own patients. It seems that this group is the most privileged one, and the
203 one with the least pressure to establish collaboration at whatever cost. Fi-
204 nally, there is one more group formed by those who work as clinicians at
205 small hospitals with no research infrastructure whatsoever. This group may
206 have an interest in research and the way for them to become part of a project
207 and be listed as authors in scientific papers is by giving researchers who do
208 not have access to biological samples access to patients.

209

210 Due to these differences in the access to resources, researchers in general
211 are compelled to build strategic alliances through which samples, data, tech-
212 nology and authorship, among other assets, become part of a constant flow
213 through the network. Social and political capital, as well as concentrations of
214 resources become fundamental tools for establishing fruitful collaborations.

215 **3 Methodology**

216 Our model is based on the iterated version of the Prisoner's Dilemma (PD)
217 and a coordination game instantiated on networks. Implementing games
218 on networks is not new and it's an active area of research aimed to under-

³Regulation exists regarding researcher-subject relations based on legal and ethical grounds. Also, all projects need to be approved by the Ethics Committee and IRB.

stand the evolution of cooperation in networks populated by selfish agents [22, 23, 24, 28, 27]. In many network models on which some of game theory games are simulated, agents' decision to cooperate or defect depend on a specific strategy, such as the well known *tit-for-tat* [25, 26]. In some other cases, agents can modify the weight of the interactions with their neighbors [27]. From a different perspective others have explored the effect of different topologies on the emergence of cooperation [28, 29]. **In our model, an agent's decision to cooperate or defect depends on a balance between utilities and the current strength of its collaboration relationships. Such balance reflects the overall success or failure of its strategies.** We study the behavior of the system under different topologies, including a real-world network.

230

In our model, agents are embedded in a network with varying number of neighbors. Following the traditional PD game, the strategy chosen by an agent and the strategy chosen by its neighbors will produce a pay-off. Pay-off follows the traditional PD rule: $T > R > P > S$. T is for temptation to defect. It is the highest pay-off and it takes place when the player defects and the other cooperates. R is for reward for when both players cooperate. P is the punishment for when both players defect. And S is for suckers pay-off, the worst outcome that takes place when the player cooperates but its neighbor defects. Utility is a property of agents in which pay-off is accumulated.

241

PD utility pay-off matrix

243

	Cooperate	Defect
Cooperate	R, R	S, T
Defect	T, S	P, P

246

The strength of the interaction, represented by w , is a property of the link between two agents and gets updated according to an A_{ij} matrix of a coordination game. In the w matrix, the highest value goes to an edge when both agents cooperate, getting an R for reward, if one of them defects, the connection gets weaker getting P for the collaborative connection being punished. If both agents defect, the value w doesn't change, which means that agents didn't interact or that the interaction gets nullified N . In this game, the best action for any agent is to coordinate with its neighbor, either because it wins or because it doesn't lose.

256

257 **w pay-off matrix**

	Cooperate	Defect
Cooperate	R	P
Defect	P	N

261

262 After each game, the agent adds-up utility (u), which is the sum of the
263 pay-offs following the PD matrix. A pair of neighbors will add-up to the
264 strength of their interaction (w_{ji}) as they coordinate or anti-coordinate, be-
265 ing w also cumulative. We measure global utility and connection strength
266 for the whole network. Global utility U is the sum of all individual utilities
267 and global strength of connections or W is the sum of every pair of agents'
268 links w . The strength of interaction can be thought as some sort of “trust”.

269

270 It should be noted that the same actions or behaviors work for both u
271 and w . There are two reasons for this decision in the design of the model.
272 The most general one is that we believe that in the real world, actions such
273 as cooperating and defecting affect the strength of the interaction among
274 people. The second one is that we think that selfishly maximizing access
275 to resources and strengthening relationships are *opposing forces* acting on
276 the same set of behaviors. The actions of an agent imply a trade-off in
277 which defecting may increase its utility at the expense of its collaborative
278 relationships. If collaborators have nourished their relationships, they might
279 be strong enough to endure occasional defection. Cooperating may build up
280 relationships but it can be expensive for the player.

281 **3.1 Network initialization and agent state update**

282 All networks are initialized equally. The number of nodes for every network
283 is 4122, the same as in the FOSISS network. The same utility is given to
284 every agent and all edges are assigned the same weight. In the case of the FO-
285 SISS network, edge weight is given by the number of collaborations among
286 researchers, utility remains the same for all nodes as in the other networks.

287

288 The probability for an agent to cooperate or defect depends on a num-
289 ber (η) that refers to a historical balance between average utility and the
290 average strength of the connections with its neighbors. This is so because
291 we assume that whatever the result in utility or strength of connection, as
292 long as one of them increases, the player will be confident in the strategy
293 followed so far.

294
295 η is calculated as:

296
297
$$\eta_i = \frac{\langle f_i \rangle + \langle w_{ij} \rangle_j}{2}$$

298

299 For the agent to decide whether to cooperate or not, η is compared to
300 a global threshold ν . If the agents' $\eta > \nu$, then the agent will cooperate,
301 otherwise he will be suspicious and will defect. ν is a global parameter that
302 establishes a threshold that an agents' η must cross in order to decide to
303 cooperate. In this way, η can limit the size of the population of cooperators.
304 Due to what the system and the game can offer to agents in terms of utilities
305 and the strength of collaboration relationships, η represents the carrying
306 capacity of the system for the population of cooperators.

307 Our simulation was tested on an Erdös-Renyí, a Watts-Strogatz small-
308 world and Barabási-Albert topologies, as well as on the real biomedical
309 research collaboration network. The simulation was run in a synchronous
310 manner, in which all agents update their behavior simultaneously.

311

312 We ran two different experiments. In the first we simulated different
313 values of carrying capacity ν . With this experiment we were able to see how
314 the number of cooperators, utility, strength of connections among agents
315 and the ratio of shifting state population would change in the range of the
316 carrying capacity. The states of the agents were the same at initialization,
317 for all values of the carrying capacity. Since the model is deterministic, it
318 will return the same result if run under the same conditions.

319

320 The second experiment consisted in running the simulation under the
321 same degree of carrying capacity ν but randomizing the initial states of the
322 agents. This would show that the system converges to a global state. For
323 every network, the simulation was run 100 times and results were averaged.

324 **4 Implementation of the model in different topolo-**
325 **gies**

326 We built three classical topologies for networks besides the FOSISS network,
327 their parameters are shown in the following table.

328

Topology	m	$\langle k \rangle$	$\langle C \rangle$	$\langle l \rangle$
Erdős-Rényi	25591	12.4	0.003	3.6
Watts-Strogatz	206100	100	0.7	3.4
Barbási-Albert	183465	89	0.06	2.13
FOSISS	23391	11.39	0.87	5.49

4.1 Erdős-Renyí

Erdős-Renyí networks [30] (random networks) are constructed by randomly selecting a pair of N possible nodes and attaching them with an edge, given a probability p , as long as there is no edge between them. The result is a Poisson distribution for connectivity of nodes $P(k)$, where each node has a degree quite close to the average $\langle k \rangle$. Also for this type of network, average clustering coefficient $\langle C \rangle$ is small, actually it is equal to p (the probability of connecting two nodes) and the average shortest path length $\langle l \rangle = \frac{\ln N}{\ln \langle k \rangle}$.

4.2 Small-World

Watts-Strogatz networks [31] (small-world networks) are in a regime between a fully regular grid (lattice) and a random network (Erdős-Renyí). In order to build them, a node is chosen from a lattice (a ring) and the edge that connects it to nearest neighbor in a clockwise sense. With probability p , this edge is reconnected to a node chosen uniformly at random over the entire ring, with duplicate edges forbidden; otherwise the edge is left in place. This process is repeated by moving clockwise around the ring, considering each node in turn until one lap is completed. Next, the edges connect nodes to their second-nearest neighbors clockwise. As before, each of these edges is randomly rewired with probability p . This process continues, circulating around the ring and proceeding outward to more distant neighbors after each lap, until each edge in the original lattice has been considered once. The main characteristic of these networks is that the average shortest path length is small and grows as $\log(N)$ ($\langle l \rangle \sim \log(N)$). Also, the average clustering coefficient $\langle C \rangle$ remains large in terms of p . For $p < 0.1$, $\langle C \rangle \sim 1$.

4.3 Barbási-Albert

Barbási-Albert networks [33] (scale-free networks) are generated by adding new nodes to a network. Each new node is added connecting it to an existing node with a probability proportional to the degree k (connectivity) of each node (*preferential attachment*). The result is a power law distribution for

360 connectivity of nodes $P(k)$ where few nodes have many connections and the
361 most have very few connections. Furthermore these networks are also small
362 world networks, showing a quite small $\langle l \rangle$.

363 **4.4 FOSISS: Biomedical research community network**

364 The biomedical research network on which we are running our model was
365 generated with data from collaborative projects. Our data was obtained
366 from CONACyT and includes information for twelve years of *FOSISS* grants.
367 Data included names of Principal Investigators, collaborators, research top-
368 ics, etc. The network we are using here has researchers as nodes and edges
369 represent the connection of two scientists when they collaborate on the same
370 project. Edges are also weighted according to the number of projects shared
371 by any pair of scientists.

372

373 **5 Results**

374 In this section we present the main results of the study, namely the topolog-
375 ical structure of the underlying network models, the dynamics of the games
376 under different parameters and network topologies and the distribution of
377 utility and of the strength of interaction resulting of playing the games in
378 all the different scenarios considered, including the real FOSISS network.

379

380 FOSISS network summed-up a total of 145 components or subnetworks,
381 but we ran the model on the giant component made-up of 4122 researchers,
382 and 23391 edges. The giant component was analyzed using *Cytoscape* **Fig-**
383 **ure 1.** Results show that it is a well integrated network, with a clustering
384 coefficient $\langle C \rangle = 0.870$, an average shortest path length of $\langle l \rangle = 5.493$ and
385 a density of $p = 0.003$. Such properties recall a small-world topology [31],
386 and a great deal of self-organization when compared to a random network
387 with the same density and number of nodes. Network centralization is 0.023,
388 since there are no visible researchers that play as hubs in the network. Nev-
389 ertheless the network heterogeneity is 0.873, which means that the network
390 is highly hierarchical. When the degree distribution is analyzed, degree
391 decreases as a power-law with an exponent of 1.7, similar to other social
392 networks described as scale-free topology networks [33]. Finally, the aver-
393 age number of neighbors of each node is 11.39 [34].

394

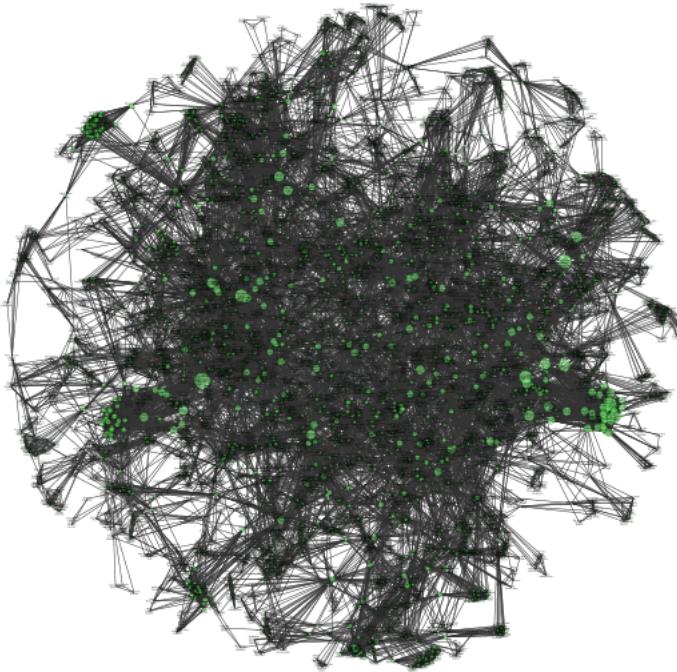


Figure 1: Bioemdical research collaboration network (FOSISS) giant component.

395 Other results to report are those of the dynamics of different variables
 396 as the carrying capacity ν changes. The most salient result is that for ν
 397 between 0.19 and 0.24, there is an apparent phase transition in all different
 398 topologies and for all the different variables. It is worth noting that the
 399 shape of the phase transition is different according to the topology of the
 400 network under study. When ν is between 0.0 and 0.2, that is, when there is
 401 no space or a very short one for suspiciousness, all agents cooperate, when
 402 *carrying capacity* is above 0.25, all agents defect. *Utility, strength of inter-*
 403 *actions, and changing state population* replicate that same behavior for the
 404 same limits.

405

406 In **Figure 2**, we present how the number of cooperators in the popula-
 407 tion change as ν changes. In the Erdős-Renyí, network, between 0 and 0.18
 408 approximately, all agents converge to a cooperative behavior, from 18 to 20,
 409 convergence to cooperative state takes longer but eventually all agents are
 410 cooperating. Close to $\nu \approx 0.21$ there is a sharp fall to a point in which

411 around half the population is cooperating and the rest is defecting. Reaching
412 $\nu \approx 0.25$ there is another sharp fall of cooperators and all agents turn
413 into a defecting state.

414

415 In the case of the Watts-Strogatz, small-world network, the whole popu-
416 lation remains cooperating for ranges between 0 and 0.2 but as it gets closer
417 to 0.2, more time is needed for the population to become full of coopera-
418 tors. In $\nu \approx 0.2$ the cooperators will represent only half of the population
419 and such number of them will be constant up to $\nu \approx 0.25$ forming a short
420 plateau. From $\nu \approx 0.25$ to $\nu \approx 0.6$ cooperators will be present at the be-
421 ginning of the simulation but will diminish as time goes on. In the case of
422 the Barabási-Albert network, crossing the threshold of $\nu \approx 0.2$, there is a
423 sharp decrease in the number of cooperators, but stays constant over time.
424 Such behavior is present for a very short range of ν , and before $\nu \approx 0.24$,
425 cooperators disappear for the rest of values of ν . Finally, FOSISS network
426 behaves similarly to the other networks in that there is a fall in the num-
427 ber of cooperators close to $\nu \approx 0.2$. In contrast with the other networks,
428 the FOSISS network lacks the sharp reduction of cooperators, instead this
429 population declines smoothly and progressively; specially, when it reaches
430 a $\nu \approx 0.25$ cooperators decrease in a less dramatic manner all the way to
431 $\nu \approx 0.5$. It is also noteworthy that from $\nu = 0$ to $\nu \approx 0.5$ the number of
432 cooperators converge to a certain degree and stays constant for the rest of
433 the simulation.

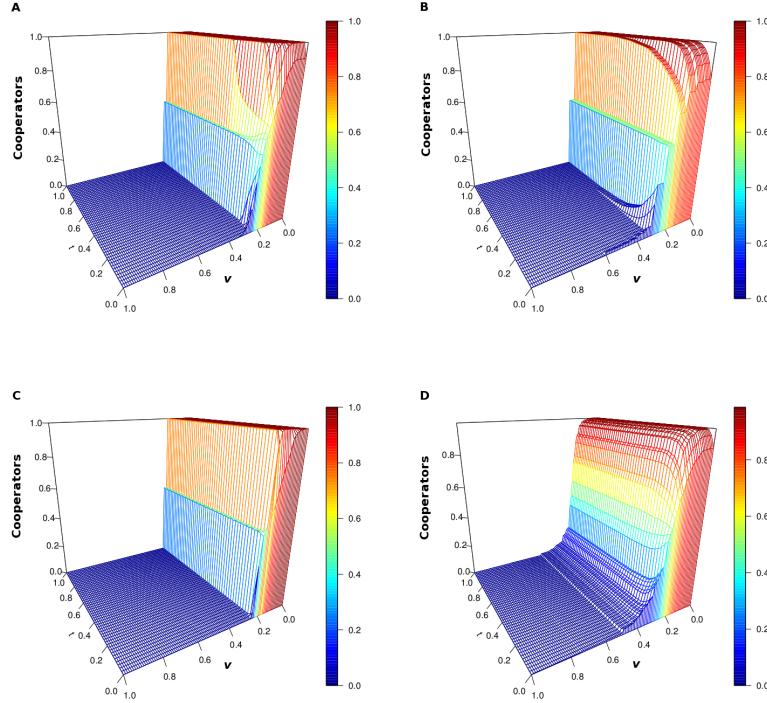


Figure 2: Ratio of cooperators as a function of carrying capacity and time. Topologies are: A. Random, Erdős-Rényi. B. Watts-Strogatz. C. Barabási-Albert. D. FOSISS.

434 Utility and strength of interactions dynamics under different ν are similar
 435 to the phase transition found before. **Figures 3 and 4** show that there is
 436 a drop in utility and strength of interactions according to the drop in the
 437 number of cooperators. Erdős-Renyí and Barabási-Albert networks are quite
 438 similar in the way these variables fall in two steps, the first one at $\nu \approx 0.2$
 439 and the next one at $\nu \approx 0.23$. The fall is sharper still in the Barabási-Albert
 440 topology. Utility and strength of interactions phase transition in Watts-
 441 Strogatz network is significantly more staggered compared to the former
 442 networks. In the case of utility, there is a region in the limits of $\nu \approx 0.25$ and
 443 $\nu \approx 0.3$, before utility goes to 0, in which it remains low but stable over time.
 444 In general, strength of interactions follows the same pattern as utility but in
 445 the same $\nu \approx 0.25$ and $\nu \approx 0.3$, strength of interactions grows to a value that
 446 is higher than the one given by default but soon starts to decrease as the

447 simulation runs. For the FOSISS network, utility and interaction strength
 448 fall quite steeply but smoothly, without sharp cuts. Between $\nu \approx 0.23$ and
 449 $\nu \approx 0.28$, utility and strength of interactions start at their lowest, but there
 450 is a slight increase in both of them.

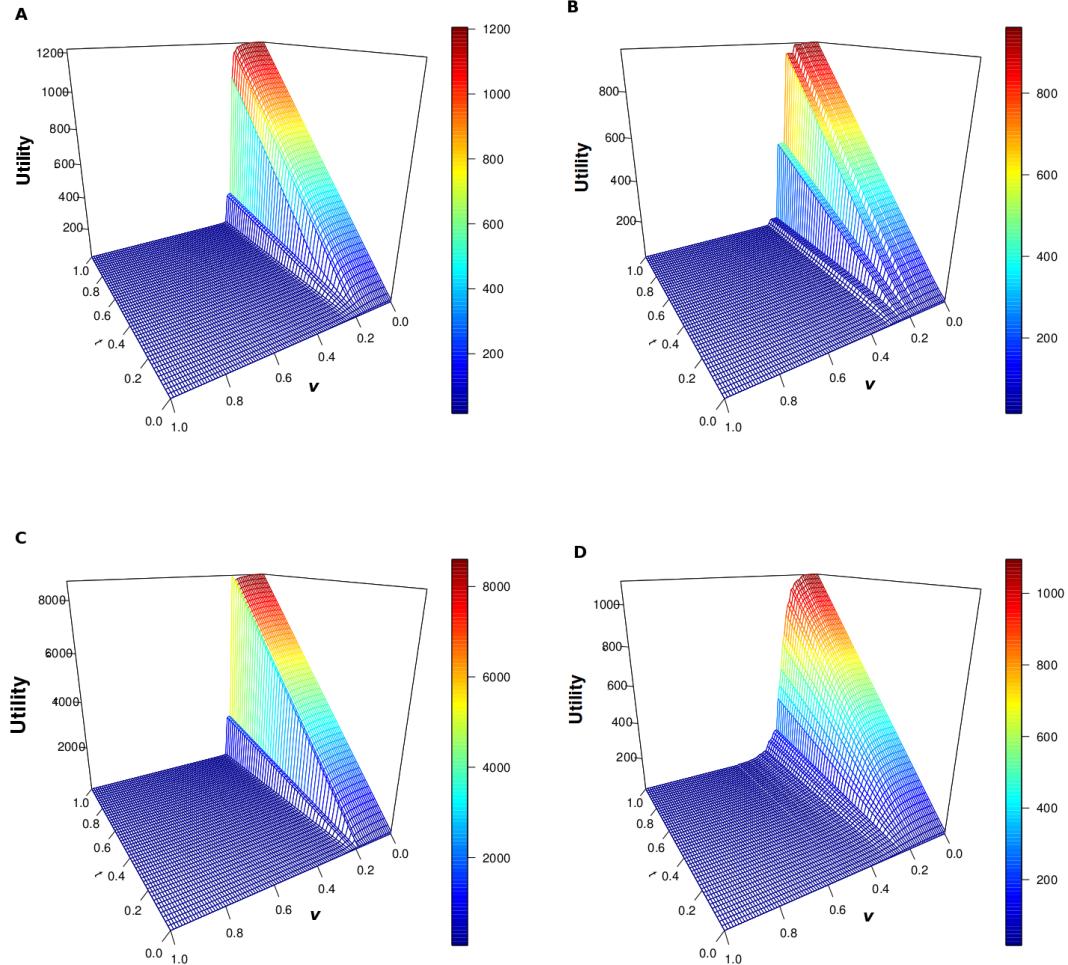


Figure 3: Utility U dynamics as a function of carrying capacity ν and time. Topologies are: A. Random, Erdős-Rényi. B. Watts-Strogatz. C. Barabási-Albert. D. FOSISS.

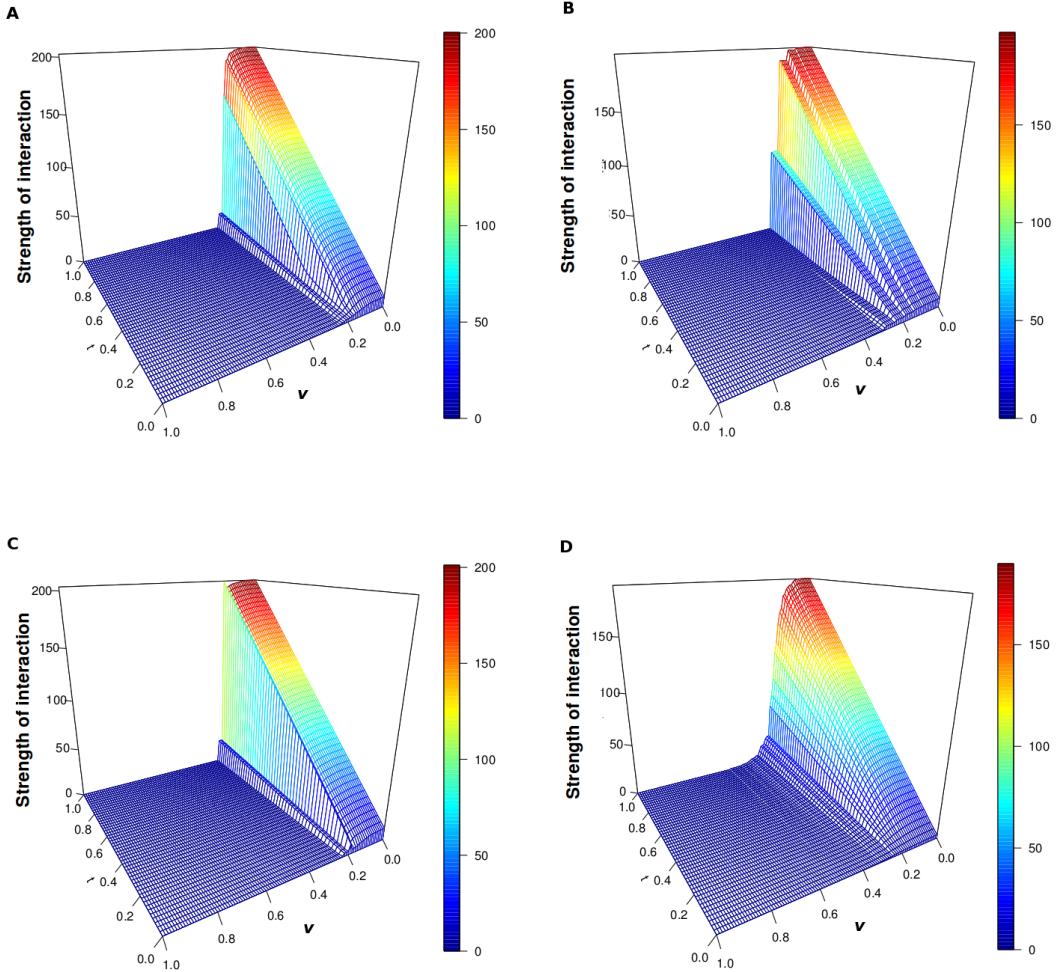


Figure 4: Strength of interactions w dynamics as a function of carrying capacity ν and time. Topologies are: A. Random, Erdős-Rényi. B. Watts-Strogatz. C. Barabási-Albert. D. FOSISS.

451 We also measured the number of agents shifting states –between cooperating
 452 and defecting– under different ν values. We found that for all networks
 453 there is a critical point around $\nu \approx 0.2$ in which all agents are shifting states.
 454 For Erdős-Renyí and Barabási-Albert networks, for this region, agents never
 455 settle to a single state. Contrary to the former cases, the number of shifting

agents decreases considerably for the Watts-Strogatz and FOSISS networks, and find an equilibrium state. Once the the limits of this region are crossed as ν increases, the number of agents shifting states falls to 0 and all nodes become defectors.

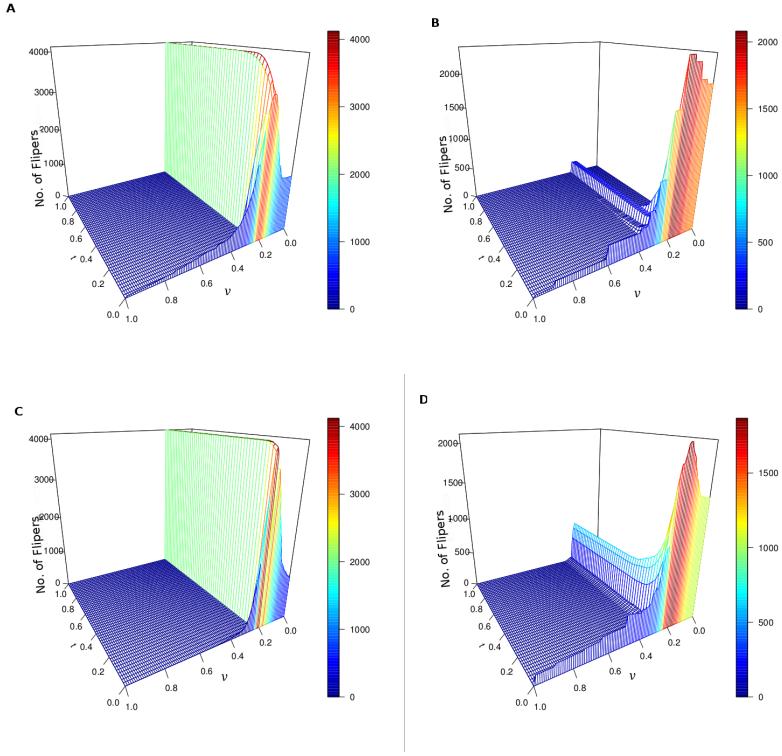


Figure 5: Shifting population between cooperators and defectors as a function of carrying capacity ν and time. Topologies are: A. Random, Erdős-Rényi. B. Watts-Strogatz. C. Barabási-Albert. D. FOSISS.

Central to our argument is the differences in utility and strength of interactions distribution at the end of the simulation, for every topology. We found that utility distribution for the *FOSISS* network, resembles quite accurately to the distribution of utility in the Barabási-Albert network.

464

The distribution of utility on each topology is induced by the degree distribution. This is so, since a given agent (node) will interact with its neighbors to either cooperate or defect, in such a way that connectivity influences the number of events played and thus the likelihood of increasing

469 its corresponding utility. For instance, utility distribution in the random,
470 Erdős-Rényi network displays a normal-like curve. The algorithm that gen-
471 erates this kind of topology, assigns to every node the same probability of
472 connecting with any other node, which produces a *poissonian* degree dis-
473 tribution [30]. Since the Watts-Strogatz degree distribution is described by
474 a function that is midways between a random distribution and a scale-free
475 network [32] one may expect also an intermediate behavior of the utility
476 distribution. This assumption seems to be fulfilled by the distribution in
477 Figure 6B.

478 The resemblance of the degree-distribution and utility distribution also
479 holds for the Barabási-Albert network. As mentioned in the methods sec-
480 tion, the degree distribution of a Barabási-Albert topology follows a power-
481 law that describes the fact that there are a small number of nodes with
482 large k and most nodes have a small k [33]. As it is shown in the following
483 figure, most utility is concentrated in a few number of agents, while most
484 agents have a small amount of it. This is consistent with other research in
485 which concentration of resources, fame or citations in science decreases as a
486 power-law [37, 38, 39]. FOSISS network utility distribution is also skewed
487 to the left, similar to that of the Barabási-Albert network.

488

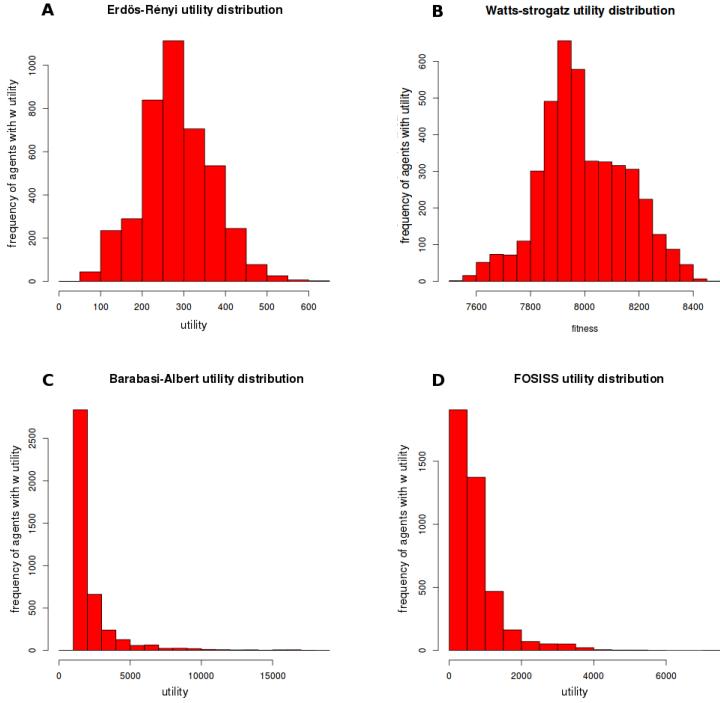


Figure 6: Utility distribution on different topologies. A. Random, Erdős-Rényi network displays a normal-like distribution. B. Watts-Strogatz network utility distribution is highly skewed to the right. C. Barabási-Albert network. Utility is distributed highly skewed to the left. D. FOSISS, biomedical researchers collaboration network distribution of utility resembles to Barabási-Albert network.

489 Regarding strength of interactions distribution, the Erdős-Rényi ran-
 490 dom network displays a normal distribution of strength of interactions, as
 491 expected. Again strength of interactions values are highly influenced by
 492 the corresponding degree distribution (Figure 7A). However, in the Watts-
 493 Strogatz network topology (Figure 7B), the strength of interactions distri-
 494 bution is a highly asymmetric bimodal, with a really-low frequency mode of
 495 low strength of interactions and a highly probability mode for high strength
 496 of interactions. A possible explanation for this phenomenon is that under
 497 network topologies maximizing inter-node communication (by minimizing
 498 the average distance between nodes) such as the Watts-Strogatz, strength
 499 of interactions is favored both among the cooperators (constituting the ma-

500 jority of players) and the defectors.

501

502 The Barabási-Albert network (7C) presents also a symmetric unimodal
 503 distribution with values higher (on average) than those of the Erdős-Rényi
 504 random network, this may be the effect of increased communication due to
 505 more efficient network navigability. Interestingly, the network corresponding
 506 to the real FOSISS collaborations (7D) is an asymmetric unimodal distribu-
 507 tion in which moderate to high values of strength of interactions are more
 508 likely. We hypothesize that this effect is also due to the communication
 509 properties of the network.

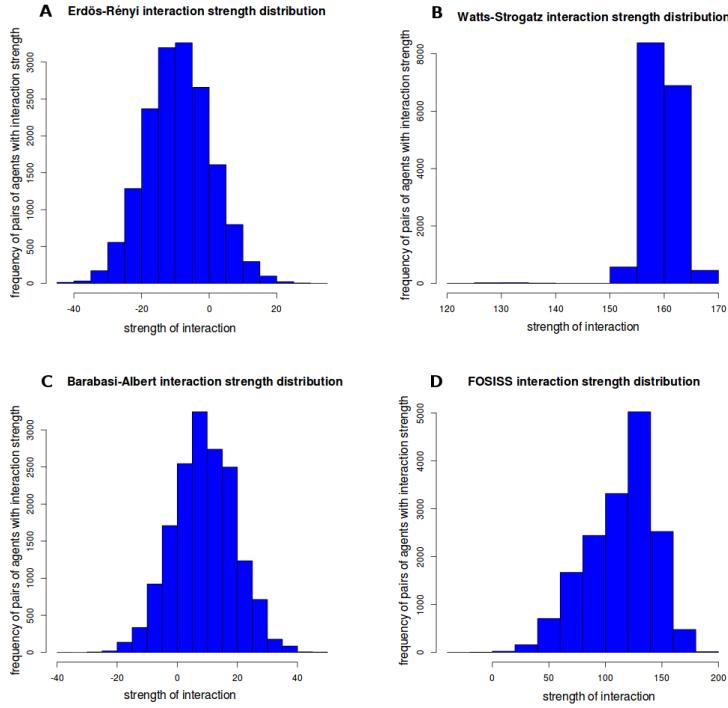


Figure 7: Strength of interactions distribution on different topologies. A. Random, Erdős-Rényi network displays a normal-like distribution. B. Watts-Strogatz network strength of interactions distribution is bimodal and highly skewed to the right. C. Barabási-Albert network strength of interactions distribution is also a normal-like curve. D. FOSISS, biomedical researchers collaboration network distribution of strength of interactions is skewed to the right.

510 An interesting feature of highly communicated networks (characterized
511 by high values of clustering coefficient) is the fact that certainty among play-
512 ers seems to be enhanced, that is, in such networks the rate of change of
513 strategy is significantly lower (and smoother) than in poorly connected net-
514 works. This is another instance in which easier communication (i.e. lower
515 average minimum path lengths) leads to better performance of the whole
516 collaborative research system.

517

518 6 Discussion and Conclusions

519 In this work we have analyzed the influence of parameters given by the
520 underlying social structure of a science collaborative network on collective
521 strength of interactions and utility dynamical behavior, based on a class of
522 iterated PD and coordination games. Such parameters include mainly local
523 and global connectivity like the degree centrality and average clustering co-
524 efficient, as well as communication patterns.

525

526 Under the assumptions given by the model, we were able to notice that,
527 in general, communication within the social collaborative networks has a
528 positive correlation with average strength of interactions between the indi-
529 viduals partaking in the games and also with the global collective utility
530 (given by the sum of the individual payoffs). The better the communication
531 among players, the higher the strength of interactions and the utility lead-
532 ing to an optimized functioning of the whole *scientific collaboration system*.
533 This is an important result that may be useful for scientific policy planning
534 and may set a foundation for the optimal use of social networks in scientific
535 collaboration as a means to improve the relationships among collaborating
536 peers and ultimately the performance of research systems. We obviously
537 need more qualitative work in order to validate these results from a socio-
538 logical and anthropological perspective.

539

540 To close this article, we would like to comment on three issues we con-
541 sider important. The first one is about confronting our results with data
542 available to us from the biomedical research community. Second, some re-
543 marks we believe are important about the role of computational simulations
544 in the social sciences. Finally we would like to comment on our future work.

545

546 The two main results of our model are the smoothness in the phase

547 transition-like behavior for different parameters and the distribution of util-
548 ity and strength of interactions in the FOSISS network compared to other
549 networks and their topologies. We are certain that the particular structure
550 of FOSISS network is playing a central role in the results and because of
551 that we would like to discuss it a little bit further.

552

553 FOSISS network is highly hierarchical according to its heterogeneity of
554 0.873. Because of it, one would expect to find the presence of important
555 hubs [35], that is, a few researchers control the whole network, as has been
556 reported in some other places [36]. Surprisingly, FOSISS network central-
557 ization is very low 0.023, i.e. there are no researchers that centralize the
558 majority of connections. Our guess is that the network is composed of
559 many small communities or groups with a central researcher or Principal
560 Investigator (PI). If this is the case in the FOSISS network, it means that
561 those groups have a very hierachic structure as well.

562

563 Under such structure, when the clustering coefficient is considered, it can
564 be said that groups are also well connected but that inter-group connections
565 are sparse. In other words, individual groups are strongly connected within
566 but the network as a whole is supported by a small number of links. We
567 came by this idea partially from another study about scientific collabora-
568 tions based on co-authorships in one of the research centers that is part of
569 the FOSISS network [10]. In the cited reference an apparently well inte-
570 grated community was found (high clustering coefficient and a very short
571 characteristic path length). Such integration was mostly superficial, since
572 it depended on the presence of external collaborators from other research
573 institutions (most of them from overseas). Removing these external collab-
574 orators broke down the network into small subgraphs that worked indepen-
575 dently. Remarkably those subgraphs corresponded to the real groups of that
576 research center. What is more, several groups had a hierarchical structure
577 as the one we suspect is common in FOSISS subgraphs. The results showed
578 that collaboration was poor between groups but strong among the members
579 of each group, and that collaboration among groups doesn't emerge bottom-
580 up, instead it seems to be promoted from the top, from the administrative
581 authorities.

582

583 We believe that the the situation just refered is also true for the whole
584 biomedical research community in México. The amount of PIs who have also
585 been collaborators in other projects is about one fifth of the total number of
586 participants. This is a number big enough to connect the whole network in

587 one giant component. Yet, due to the topological characteristics of the FOS-
588 ISS network, it appears to be the case that researchers can be the leaders in
589 one project and collaborators of different project of *its own research group*.
590 CONACyT's funding policies makes it impossible for a researcher to get a grant
591 from a fund if that researcher has an ongoing project with a grant from that same fund.
592 That is, a researcher can ask for a grant from FOSISS if and only if at the moment he doesn't already have one. This policy has
593 lead researchers from the same group to ask for grants from the same fund in order to rise their budgets.
594 A consequence of this behavior is that group interactions get reinforced, but intergroup connections not necessarily so.
595

597

598 As is common among scientific communities in biomedical research, PIs
599 play a central role in the network. Strong PIs and well connected groups
600 seem to be somehow responsible for the high levels of strength of interactions
601 and the centralization of utility in our simulation. As for what seems to be
602 phase transitions, in the case of FOSISS networks, these are smoother than
603 those in the other networks with different topologies, even for those with a
604 small-world topology. We think that this behavior is also the result of the
605 hierarchical structure already mentioned. If this is true, strength of interactions
606 is first lost in the edges that link diffrent goups and then in the edges
607 that connect members of the groups. Connections between groups would
608 not be as dense as those inside the groups, which means that there would
609 not be enough information of the behavior of one group regarding its neighbors
610 to constrain them as it seems to happen with individual researchers
611 inside their communities. Nevertheless, if values of the carrying capacity
612 ν keep increasing and it becomes more difficult to strengthen interactions,
613 then strength of interactions begins to diminish inside groups.

614

615 Another issue that we would like to mention about FOSISS network
616 topology is that it is not a robust collaboration network. At the level of
617 groups, these might be well connected and consolidated but at the level of
618 the network, this could be no more than an aggregate of individual groups.
619 A robust network would be resistant to changes in the connections between
620 groups but in the case of FOSISS, it seems that the network would brake
621 down into small research groups by cutting some edges, as it happened in our
622 co-authorship collaboration network [10]. The lack of robustness might be
623 indicative of the fact that resources stay inside the groups, that is, they do
624 not circulate through or articulate different communities. For example, one
625 may think of certain expensive technologies for genomic research that could
626 be bought once and shared among research groups, however, this doesn't

627 seem to happen very often. There are some other consequences, such as
628 low communication among groups, atomization of practices and know-hows,
629 redundancy in equipment tenancy, difficulties for implementing community-
630 wide infrastructures such as biobanks, etc.

631

632 On these grounds, part of our future work is based on the results pre-
633 sented here. We would like to identify researchers in our simulations and
634 corroborate their situation in the model and in the real world. We are also
635 interested in going back to the field and interviewing those groups with an
636 interesting behavior found in our simulations, probably we would follow a
637 similar strategy as the one developed in [42]. Finding communities beyond
638 the level of the groups is an important task. We think that there are many
639 possibilities that emerge from the integration of different methodologies.
640 Moreover, studying social processes in science is particularly attractive due
641 to the amounts of data already available that can be easily collected. This
642 is a privilege because simulations can be designed on real world data, some-
643 thing that only very recently has become possible [43].

644

645 No doubt, social sciences are becoming more “computational”. This
646 can be seen everywhere, but curiously enough, disciplines like physics and
647 computer science are moving towards the social sciences and not so much
648 the other way around. It might be the case that the pioneering disciplines
649 in the computational social sciences will set the agenda, an agenda that
650 will apparently be mostly based on taking advantage of big data and on
651 hypothesis-free approaches. We believe that the social sciences have impor-
652 tant questions that should be added to that agenda and those questions may
653 not be answered only by big data techniques but they may require creating
654 models and simulations in the style of the best hypothesis-driven research.

655 7 Acknowledgements

656 The authors gratefully acknowledge support by grants: CB-222220-R/2013
657 and CB-179431/2012 (Consejo Nacional de Ciencia y Tecnología). We would
658 also like to acknowledge CONACyT for letting us access their database.
659 Finally, we acknowledge the help of Francisco Allende for his work curating
660 databases of researchers.

661 References

- 662 [1] Vermeulen, Nikki; Parker, John N & Penders, Bart, Understanding
663 life together: A brief history of collaboration in biology, *Endeavour*,
664 37(3): 162-171 (2013)
- 665 [2] Knorr-Cetina, Karol; *Epistemic cultures: how the sciences make*
666 *knowledge*, MIT Press, Cambridge, MA (1999)
- 667 [3] Strasser, B.J. Collecting and Experimenting: The Moral Economies
668 of Biological Research, 1960s-1980s. *Preprint no. 310*. Berlin: Max
669 Planck Institute for the History of Science, (2006)
- 670 [4] Müller-Wille, S. & Charmantier. I. Natural history and information
671 overload: The case of Linnaeus. *Studies in History and Philosophy of*
672 *Biological and Biomedical Sciences*. 43:4-15, (2012)
- 673 [5] Strasser, BJ. Data-driven sciences: From wonder cabinets to elec-
674 tronic databases. *Studies in History and Philosophy of Biological and*
675 *Biomedical Sciences* 43:85-87, (2012)
- 676 [6] Sonnenwald, DH. Scientific Collaborations. *Annual Review of Infor-*
677 *mation Science and Technology*, 41.
- 678 [7] Newman, MEJ. Clustering and preferential attachment in growing net-
679 *works. Physical Review E*, 64:025102-1/025102-4, (2001)
- 680 [8] Newman, MEJ. Coauthorship networks and patterns of scientific col-
681 laboration. *PNAS*, 101, no. suppl 1, 5200-5205 (2004)
- 682 [9] Elango, B. & Rajendran, J. Authorship Trends and Collaboration Pat-
683 tern in the Marine Sciences Literature: A Scientometric Study. *Inter-*
684 *national Journal of Information Dissemination and Technology*, 2(3),
685 166-169, (2012)
- 686 [10] Hernández-Lemus, Enrique & Siqueiros-García JM, Information the-
687 oretical methods for complex network structure reconstruction. *Com-*
688 *plex Adaptive Systems Modeling*. 1: 8, doi:10.1186/2194-3206-1-8
689 (2013).
- 690 [11] Gilbert, N. A Simulation of the Structure of Aca-
691 demic Science. *Sociological Research Online*, 2(2) 3,
692 <http://www.socresonline.org.uk/2/2/3.html>

- 693 [12] Edmonds, B., Gilbert, N., Ahrweiler, P., & Scharnhorst, A. Simulating the Social Processes of Science. *Journal*
 694 *of Artificial Societies and Social Simulation*, 14(4) 14,
 695 <http://www.jasss.soc.surrey.ac.uk/14/4/14.html>
- 697 [13] Hara, N. Solomon, P. Kim SL. Sonnenwald DH. An Emerging View of
 698 Scientific Collaboration: Scientists Perspectives on Collaboration and
 699 Factors that Impact Collaboration. *Journal of the American Society*
 700 *for Information Science and Technology*, 54(10):952965, (2003)
- 701 [14] Lieberman, E. Hauert, C. & Nowak, MA. Evolutionary Dynamics on
 702 Graphs. *Nature*, 433, doi:0.1038/nature03204 (2005).
- 703 [15] Lo, T.S., Chan, H.Y., Hui, P.M., Johnson, N.F., Theory of networked
 704 minority games based on strategy pattern dynamics, *Physical Review*
 705 *E* 70, 056012 (2004)
- 706 [16] Buesser, P. and Tomassini, M., Evolution of cooperation on spatially
 707 embedded networks, *Physical Review E* 86, 066107 (2012)
- 708 [17] Ichinose, G., Tengnishi, Y., and Tanizawa, T., Robustness of coop-
 709 eration on scale-free networks under continuous topological change,
 710 *Physical Review E* 88, 052808 (2013)
- 711 [18] Ahmed, A., Karlapalem, K., Inequity aversion and the evolution of
 712 cooperation, *Physical Review E* 89, 022802, (2014)
- 713 [19] Vilone, D., Ramasco, J.J., Sánchez, A., San Miguel, M., Social imi-
 714 tation versus strategic choice, or consensus versus cooperation, in the
 715 networked Prisoners Dilemma, *Physical Review E* 90, 022810, (2014)
- 716 [20] Wardil, L., Hauert, C., Cooperation and coauthorship in scientific
 717 publishing, *Physical Review E* 91, 012825, (2015)
- 718 [21] Hanauske, M., Evolutionary Game Theory and Complex Networks
 719 of Scientific Information, in A. Scharnhorst et al. (eds.), *Models of*
 720 *Science Dynamics, Understanding Complex Systems Series*, Springer
 721 Berlin Heidelberg, (2012)
- 722 [22] Szabó, György & Fáth, Gábor, Evolutionary games on graphs. *Physics*
 723 *Reports*, 446: 97-216 (2007)
- 724 [23] Nowak, Martin A. & May, Robert M, Evolutionary Games and Spatial
 725 Chaos. *Nature*, 359: 826-829, (1992)

- 726 [24] Oshtuki, Hisashi, Hauert, Christoph, Lieberman, Erez & Nowak, Martin A, A simple rule for the evolution of cooperation on social graphs.
 727 *Nature*, 441: 502-505 (2006)
- 729 [25] Axelrod, Robert, *Evolution of cooperation: Revised edition*, Basic
 730 Books, Cambridge, MA (2006)
- 731 [26] Nowak, Martin A. & Highfield, Roger, *Supercooperators. Altruism,*
 732 *Evolution, and Why We Need Each Other to Succeed*. Free Press, New
 733 York, NY (2011)
- 734 [27] Santos, Francisco C., Pacheco, Jorge M., Lenaerts, Tom, Cooperation
 735 prevails when individuals adjust their social ties. *PLoS Computational*
 736 *Biology*. 2(10): 1284-1291 (2006)
- 737 [28] Santos, F.C. & Pacheco, J.M., Scale-Free Networks Provide a Unifying
 738 Framework for the Emergence of Cooperation. *PRL*, 95: 098104-1 to
 739 098104-4, (2005)
- 740 [29] Hauert, Christoph & Doebeli, Michael, Spatial structure often inhibits
 741 the evolution of cooperation in the snowdrift game. *Nature*, 428: 643-
 742 646, (2004)
- 743 [30] Erdős, P. & Renyí, A. On Random Graphs. *Publ. Math*, 6:290 (1959).
- 744 [31] Watts, Duncan J. & Strogatz, Steven H, Collective dynamics of 'small-
 745 world' networks. *Nature*, 393: 440-442 (1998)
- 746 [32] Barrat, A.; Weigt, M., On the properties of small-world network mod-
 747 els, *European Physical Journal B* 13 (3): 547560, (2000)
- 748 [33] Barabási, Albert-László. & Albert, Réka, Emergence of Scaling in Ran-
 749 dom Networks. *Science*, 286: 509-512 (1999)
- 750 [34] Shannon P, Markiel A, Ozier O, Baliga NS, Wang JT, Ramage D,
 751 Amin N, Schwikowski B, Ideker T. Cytoscape: a software environment
 752 for integrated models of biomolecular interaction networks. *Genome*
 753 *Research*, 13(11):2498-504, (2003)
- 754 [35] Jun Wu, Yue-Jin Tan, Hong-Zhong Deng, Da-Zhi Zhu. A new measure
 755 of heterogeneity of complex networks based on degree sequence. Eds.
 756 Ali Minai, Dan Braha, Yaneer Bar-Yam *Unifying Themes in Complex*
 757 *Systems. Proceedings of the Sixth International Conference on Com-*
 758 *plex Systems*, Springer Berlin Heidelberg, 66-73, (2008)

- 759 [36] Yousefi-Nooraie R, Akbari-Kamrani M, Hanneman RA, Etemadi A.
760 Association between co-authorship network and scientific productiv-
761 ity and impact indicators in academic medical research centers: A
762 case study in Iran. *Health Research Policy and Systems* 6:9 (2008)
763 doi:10.1186/1478-4505-6-9.
- 764 [37] Simon, HA. On a class of skew distribution functions. *Biometrika*, 42
765 (3-4): 425-440, (1955)
- 766 [38] Price, Derek JS. Networks of Scientific Papers. *Science*, 149(3683):
767 510-515, (1965)
- 768 [39] Merton, Robert K, The Matthew Effect in Science, *Science*, 159(3810):
769 56-63 (1968)
- 770 [40] Epstein JM & Axtell R. *Growing artificial societies. Social science*
771 *from the bottom up*. The Brookings Institution Press and The MIT
772 Press, Washington, DC. (1996)
- 773 [41] Lansing JS. "Artificial Societies" and the Social Sciences. *Artificial*
774 *Life*, 8:279-292, (2002)
- 775 [42] Hara N, Solomon P, Kim SL, Sonnenwald DH. An emerging view of
776 scientific collaboration: Scientists' perspectives on collaboration and
777 factors that impact collaboration. *Journal of the American Society for*
778 *Information Science and Technology*, 54(10):952965, 2003
- 779 [43] Barabási, AL. The network takeover. *Nature*, 8:1416, (2012)
- 780 [44] Conte R, Gilbert N, Bonelli G, Cioffi-Revilla C, Deffuant G, Kertesz
781 J, Loreto V, Moat S, Nadal J.-P, Sanchez A, Nowak A, Flache A, San
782 Miguel M, Helbing D. Manifesto of computational social science. *The*
783 *European Physical Journal, Special Topics* 214(1):325346, (2012)