

# **Master thesis title**

Pablo Lechón Alonso

Imperial College London

Word count: 10

## **Abstract**

# Contents

<b>1</b>	<b>Introduction notes</b>	<b>2</b>
<b>2</b>	<b>Introduction</b>	<b>4</b>
<b>3</b>	<b>Methods and Results</b>	<b>5</b>
3.1	Consumer resource model with cross feeding interactions . . . . .	5
3.2	Community Assembly . . . . .	8
3.3	A metric of community cohesion . . . . .	12
3.4	$\Theta$ predicts the outcome of community coalescence experiments. . .	13
<b>4</b>	<b>Discussion</b>	<b>14</b>
<b>5</b>	<b>Things to do in the future</b>	<b>16</b>
<b>6</b>	<b>Appendix</b>	<b>18</b>
6.1	Reversible enzyme kinetics . . . . .	18
6.2	Table of parameter values and meaning . . . . .	20

# 1 Introduction notes

- State my aims/hypotheses/questions by the end of the introduction.
- Make sure I explain everything adequately. Provide more background in the introduction and methods that
- What is the problem I am tackling? you end up looking forward to read more
- the general way the problem has been approached.
- Clearly define my aims of the research project.
- Why is it interesting? Why don't we know the answer?
- Build from the most general and fundamental hypotheses to the most refined or tenuous ones.
- How will I go about testing my hypothesis.

Following, what I want to talk about

- Microbial Community Coalescence in communities with mutualisms ie, cross-feeding.
- Thermodynamic constraints
- What drives a community to be successful in a coalescence event?
- Will a coalesced community be more persistent than a naive one upon a event if it has a history of coalescence?
- How cohesive a community is when we analyze it in terms of cohort and dominant species?
- Experimental studies report cohesiveness in community coalescence events.

- 23     • talk about why this model is good with methanogenic communities, strong  
24       cross feeding
- 25     • say at some point how many possible networks there are.
- 26     • talk about coselection
- 27     • measure community productivity and check that the correlation breaks  
28       down because there is co-selection.
- 29     • stress importance of dominant community interaction types
- 30     • microbial ensembles
- 31     • Theoretical efforts for community coalescence (Tikhonov 2016, Livingston  
32       et al. 2013, Toquenaga 1997, Gilpin 1994)
- 33     • Experimental efforts to understand community coalescence (Lu et al. 2018,  
34       Sierocinski et al. 2017)
- 35     • Community coalescence is just another expression of the fight competition  
36       vs facilitation...

## 37 2 Introduction

38 Microbial communities are widespread throughout our planet, from the deep  
39 ocean to the human gut, and they play a critical role in natural processes ranging  
40 from animal development and host health (Huttenhower et al. 2012) to biogeo-  
41 chemical cycles (Falkowski et al. 2008). These communities are very complex,  
42 often harboring hundreds of species (Gilbert et al. 2014), making them hard to  
43 characterize. Recently, DNA sequencing has facilitated a high-resolution map-  
44 ping of these consortia, opening a niche for ambitious theorists and experimen-  
45 talists to collaboratively disentangle the complexity of these systems (Marsland  
46 et al. 2019, Goldford et al. 2018, Goyal & Maslov 2018, Friedman et al. 2017,  
47 Costello et al. 2012). One of the problems yet to solved is community assembly  
48 – the process by which species come together and interact to establish a commu-  
49 nity. Contrary to what is found in the macroscopic world, in microbial ecology it  
50 is usual that whole communities move to a region where they encounter another  
51 community. The process by which two or more communities that were previously  
52 separated join and reassemble into a new community has been termed "commu-  
53 nity coalescence" (Rillig et al. 2015). This type of event happens repeatedly in  
54 nature due to abiotic (wind, tides or river flow), biotic (animal courtship, parent-  
55 offspring interactions or leaves falling), and anthropogenic (industrial anaerobic  
56 digestion, agriculture, between-human contact) factors (Castledine et al. 2020).  
57 Despite the frequency and importance of microbial community coalescence, the  
58 mechanisms responsible for the community structure and function resulting from  
59 coalescence events remain poorly understood.

60 Early mathematical models of community-community invasion revealed that  
61 when two communities previously separated by a barrier merge due to its re-  
62 moval, asymmetrical dominance of one community over the other one is likely to  
63 occur (Gilpin 1994, Toquenaga 1997). As an explanation for this observation, it  
64 was argued that, because communities have been assembled through a history of

65 competitive exclusion, they are likely to compete with each other as coordinated  
 66 entities, rather than as a random collection of species. New theoretical work uses  
 67 consumer-resource models to show that coalescing microbial communities exhibit  
 68 an emergent cohesiveness (Tikhonov 2016, Tikhonov & Monasson 2017). These  
 69 communities have been termed Metabolically Cohesive [microbial] Consortium  
 70 (MeCoCos) by Pascual-García et al. (2020), and suggested to be pervasive in  
 71 microbial ecology. Recent results from coalescence experiments of methanogenic  
 72 communities suggest that during a coalescence event between two communities,  
 73 multiple taxa from the same community act as cohesive units and are selected  
 74 together (ecological co-selection) (Sierocinski et al. 2017). Further experimen-  
 75 tal evidence of co-selection in community coalescence has been reported in Lu  
 76 et al. (2018), where it was shown that successful collective invasions are ac-  
 77 companied by strong community-level interactions. The microbial communities  
 78 used in these experiments are characterized by complex cross-feeding interactions  
 79 (Hansen et al. 2007, Lawrence et al. 2012, Embree et al. 2015). Furthermore,  
 80 the type of trophic interactions present in a community has been suggested as a  
 81 factor that might affect the outcome of community coalescence (Castledine et al.  
 82 2020). Yet, theoretical models used in community coalescence studies so far  
 83 have considered competition between species as the only force driving commu-  
 84 nity assembly.

85 In this work, I explore the role of other types of interactions, which appear to be  
 86 ubiquitous in microbial communities. Specifically, I propose a metric of commu-  
 87 nity cohesion that accounts for both competitive and mutualistic interactions. I  
 88 then use a consumer-resource model that includes both facilitation of metabo-  
 89 lites via by-product secretion, and competition for substrates, to simulate many  
 90 instances of community assembly. Finally, I measure the cohesion level on the  
 91 simulated communities and use it to predict the outcome of microbial community  
 92 coalescence events.

### 93 3 Methods and Results

#### 94 3.1 Consumer resource model with cross feeding interactions

##### 95 Simpler description propositions:

- 96 • The availability  $h_i$  of each resource  $i$  determines the dynamics of  $n_\mu$ . The  
97 changes in species abundance translate into changes in the total demand  
98 for resources, denoted  $T_i$ . This total demand, in turn, depletes resource  
99 availability  $h_i$ .
- 100 • The cost of the model This cost model corresponds to the assumption of  
101 approximate neutrality.
- 102 • the proportionality constant is not important because I am only concerned  
103 with the equilibrium state where  $(dN_\alpha/dt) = 0$

104 In order to simulate communities with cross-feeding interactions, I use a consumer-  
105 resource model inspired in the the work of (Marsland et al. 2019) that PhD stu-  
106 dent Jacob Cook has developed. I modify parts of this model to make it more  
107 suitable for my purposes.

108 I consider the population dynamics of  $s$  consumers (eg. bacterial strains) that  
109 feed on  $m$  resources. In this model, a species is defined by its metabolic stratetgy  
110 to harvest energy from the environment. Let  $G_\alpha(\mathcal{M}, \mathcal{N})$  be the metabolic net-  
111 work of species  $\alpha$ , where  $\mathcal{M}$  is a set of nodes  $\mathcal{M} = \{x : x \text{ is an integer from}$   
112  $\text{the interval } [1, m] \text{ labeling the metabolite}\}$  and  $\mathcal{N}$  a set of uni-directed edges  $\mathcal{N}$   
113  $= \{(x, y) : x \in \mathcal{M}, y \in \mathcal{M} \text{ and } x < y \text{ (x and y are product and substrate of}$   
114  $\text{a chemical reaction})\}$ . The growth power of species  $\alpha$ ,  $J_\alpha^{grow}$  will be given by  
115 the product of the amount of generated energy  $\eta_i$  and rate  $q_i$  of each reaction,  
116 summed across all reactions in  $\mathcal{N}$ .

$$117 \quad J_\alpha^{grow} = \sum_{i=1}^{|\mathcal{N}|} q_i \eta_i \quad (1)$$



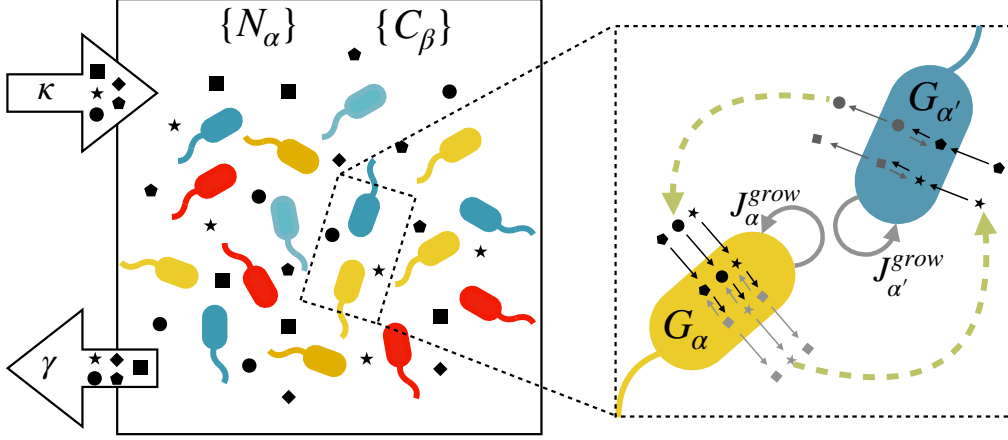


Figure 1: **Schematic of the model.** (left) All  $m$  metabolites are steadily supplied into a chemostat with different bacterial strains at rate  $\kappa$  and diluted at rate  $\gamma$ . (right) Bacteria use the metabolites in the environment,  $\{C_\beta\}$  through their reaction networks  $G_\alpha$ , to obtain the necessary power  $J_\alpha^{grow}$  to increase their abundance  $\{N_\alpha\}$  by replicating. The green arrows in the magnified portion emphasize that species  $\alpha$  (yellow) facilitates metabolites to species  $\alpha'$  (blue) and viceversa. The duble arrows in the reactions happening inside the cells stand for the reversible enzyme kinetics considered by this model.

118 where  $|\cdot|$  denotes cardinality of a set. Refer to subsection 6.1 for specifications  
 119 on  $q$  and  $\eta$ .

120 Every species has a maintenance cost  $\chi_\alpha$  that represents the required energy to  
 121 sustain life, and I model as

$$122 \quad \chi_\alpha = \chi_0 \sum_{\mathcal{N}} (y - x) \quad (2)$$

123 where  $\chi_0$  is the average cost per reaction, and the summatory term adds up  
 124 the metabolite gap of all reactions. Therefore, the more reactions a species  
 125 has or/and the more energetically they are, the higher is the maintenance cost.  
 126 The cost function (Eq. 2) ensures that neither generalists, nor specialists, are  
 127 systematically favored during the community assembly.

128 Under this parametrization, the time evolution of the population of species  $\alpha$

129 can be writtne as

$$130 \quad \frac{dN_\alpha}{dt} = g_\alpha N_\alpha [J_\alpha^{grow} - \chi_\alpha] \quad (3)$$

131 where  $g_\alpha$  is a proportionality constant relating energy to abundance of strain  $\alpha$   
 132 The dynamics of the resources depend on the incoming and outgoing resource  
 133 fluxes due to the biochemical reactions taking place inside bacteria, as well as  
 134 the resource extrenal dynamics. The incoming resource flux of metabolite  $\beta$   
 135 generated by strain  $\alpha$  is its rate of consumption due to all the biochemical reac-  
 136 tions possesed by  $\alpha$  in which  $\beta$  is a substrate. The outgoing flux is that due to  
 137 reactions in which  $\beta$  is a product.

$$138 \quad v_{\alpha\beta}^{in} = \sum_S q \quad \text{with} \quad \mathcal{S} \equiv N \cap \{(x = \beta, y)\}$$

$$139 \quad v_{\alpha\beta}^{out} = \sum_P q, \quad \text{with} \quad \mathcal{P} \equiv N \cap \{(x, y = \beta)\}$$

140 The external resource dynamics are modelled as a supply rate minus a dilution  
 141 rate that depends on the resource concentration to ensure convergent dynamics.

$$142 \quad h_\beta = \kappa - \gamma C_\beta \quad (5)$$

143 Therefore, the variation with time of the concentration of metabolite  $\beta$  has the  
 144 form

$$145 \quad \frac{dC_\beta}{dt} = h_\beta + \sum_{\alpha=1}^s (v_{\alpha\beta}^{in} - v_{\alpha\beta}^{out}) N_\alpha \quad (6)$$

146 Thus, the model consists of  $s + m$  coupled differential equations completely  
 147 specified by Eqs. 3 & 6.

### 148 **3.2 Community Assembly**

149 Armed with this model I now simulate  $n_s = 2 \cdot 10^3$  community assembly events  
 150 of  $s = 10$  species in an environment with  $m = 15$  substrates that are steadily  
 151 supplied.

152 The values of the parameters of the model (subsection 6.2, table 1) remain  
 153 constant throughout all simulations. The reason for this is that, my aim is not  
 154 to parametrize the model to reveal large-scale patterns found in experiments  
 155 (although that would be a fruitful endeavour because of the rich parameter space  
 156 of this model). Rather, I use it to simulate a set of microbial communities with  
 157 cross-feeding interactions that will be later used in the community coalescence  
 158 experiments.

159 In order to do so, I first create  $s \cdot n_s$  random reaction networks,  $G_\alpha(\mathcal{M}, \mathcal{N})$  using  
 160 the following procedure. Consider, the  $m \times m$  adjacency matrix  $A_{ij}^\alpha$ , whose ele-  
 161 ments represent the edges  $(i, j)$  of  $G_\alpha$ . Since the reaction network is hierarchical  
 162 ( $i < j$ , subsection 6.1), the adjacency matrix is an upper triangular matrix with  
 163 zeros in the main diagonal, and the reactions possessed by strain  $\alpha$  can be ex-  
 164 pressed as  $(i, i + k)$ , where  $k$  represents the  $k^{th}$  diagonal of  $A$  ( $k = 1, \dots, m - 1$   
 165 with  $k = 0$  being the main diagonal), and  $i$  is the row number of one of its  
 166 elements ( $i = 1 \dots m$ ). The reaction network  $G_\alpha$  is constructed by sampling  $n_r$   
 167 pairs  $(i, k)$  according to the algorithm summarized below.

- 168 1. Choose  $n_r$  by sampling it from a uniform distribution  $U(1, m)$
- 169 2. Choose  $k$  by sampling one value from truncated normal distribution  
 170  $N(1, \sqrt{m-1})$  with limits  $[1, m-1]$ , and rounding it to the closest in-  
 171 teger.
- 172 3. Sample  $i$  from a uniform distribution of integers  $U(0, m-k)$ .
- 173 4. The reaction  $(i, i+k)$  is stored, and the process is repeated until  $n_r$  reac-  
 174 tions have been sampled.

175 Some notes about this algorithm are, first, sampling  $k$  from a truncated normal  
 176 distribution ensures that high metabolite gaps (very energetic reactions) are not  
 177 likely to happen. This introduces a bias against the presence of super-organisms  
 178 with few and very energetic reactions, which are rare in microbial communities.

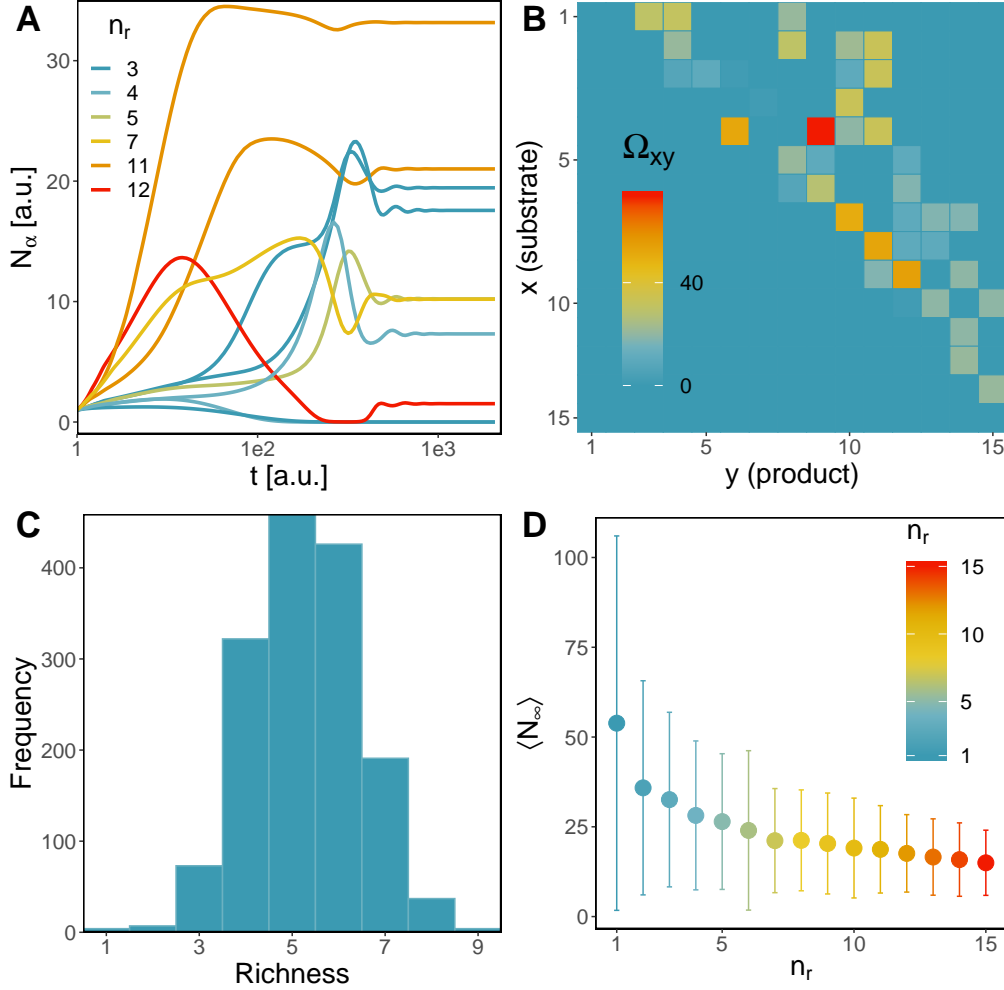


Figure 2: **Results from community assembly simulations.** Plots (A) and (B) exemplify one community assembly event and (C) and (D) convey results across simulations. (A) Time variation of species' abundance for one instance of community assembly with  $m = 15$ ,  $s = 10$ , and a set of  $s$  randomly generated reaction networks. Time (x-axis) and population (y-axis) are measured in arbitrary units. Each time series is coloured according to  $n_r$ , the number of reactions possessed by the reaction network of each strain. (B) Community reaction network, obtained by summing the adjacency matrices of all species weighted by their respective carrying capacity:  $\Omega = \sum_{k=1}^s N_\infty^k A_k$ . (C) Histogram of richness of the  $n_s$  simulations. (D). Mean value of carrying capacity  $\pm \sigma$  (error bars) against the number of reactions  $n_r$ : species with less number of reactions (specialists) are more abundant than those with higher  $n_r$  (generalists); several specialists deplete all resources through their combined action more efficiently than one generalist.

179 Second, the truncation limits in step 2 have been chosen to respect the hierarchi-  
180 cal character of the network:  $k \neq 0$  to avoid reactions of the form  $(i, i)$ . Third,  
181 the upper limit of the uniform distribution from which  $i$  is sampled is bounded  
182 by  $k$ , the diagonal we are sampling from.

183 When the sampling of reaction networks is completed, equations 6 and 3 are  
184 integrated using a Runge Kutta method (Dormand & Prince 1980) with ini-  
185 tial conditions  $N_\alpha(0) = 2$  and  $C_\beta(0) = 0$ . Relevant results stemming from the  
186 simulations of community assembly events are plotted in figure 2

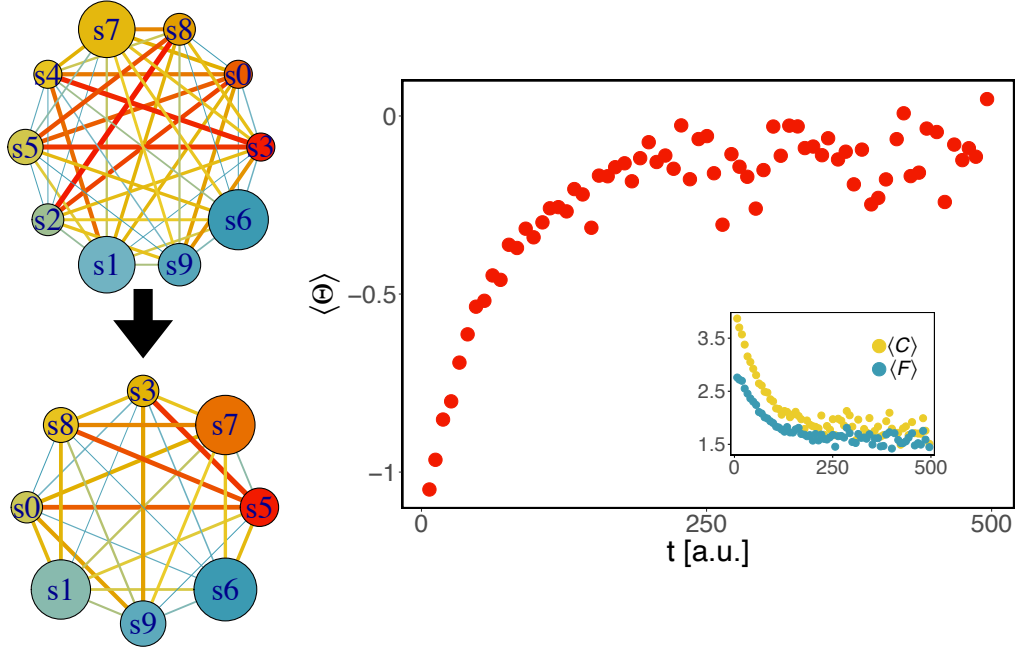


Figure 3

### 187 3.3 A metric of community cohesion

188 Following the ensemble of many synthetic communities, I postulate a metric  
 189 of community cohesion that is later used to predict the outcome of community  
 190 coalescence.

191 Let  $s_1$  and  $s_2$  be two sequences of integers labeling metabolites. I am interested  
 192 in measuring their *overlapping degree*  $\xi(s_1, s_2)$ , ie, the proportion of metabolites  
 193 of  $s_1$  that intersect with  $s_2$  summed with the proportion of metabolites of  $s_2$   
 194 that intersect with  $s_1$ , normalized to 1.

$$195 \quad \xi(s_1, s_2) = \frac{1}{2} \sum_{k \in s_1 \cap s_2} \left( \frac{D_{s_1}(k)}{|s_1|} + \frac{D_{s_2}(k)}{|s_2|} \right) \quad (7)$$

196 Here,  $k$  takes the values in the set that results from intersecting  $s_1$  and  $s_2$ .  $D_s(k)$   
 197 is the number of elements from the sequence  $s$  that are equal to  $k$ . The purpose  
 198 of all denominators in equation 7 is to normalize  $\xi$  to 1. Note that fractions  
 199 inside the summation term are equal only if the number of elements in both se-

200 quences  $s_1$  and  $s_2$  are the same.

201 The cohesion of a community – the degree to which species in a community  
 202 collaborate with each other, can be thought of as the difference  $cohesion = facil-$   
 203  $itation - competition$ . One way to capture the facilitation of a community is by  
 204 calculating its facilitation matrix  $F$ , which is composed of the facilitation indices  
 205 of all possible ordered pairs  $i, j$  of species in the community. Precisely, the facili-  
 206 tation index  $f_{ij}$  of species  $i$  towards species  $j$ , is given by the overlapping degree  
 207 of the sequence of products of species  $i$ ,  $y_i$ , with the sequence of substrates of  
 208 species  $j$ ,  $x_j$ . Equivalently, the competition matrix  $C$  gathers the competition  
 209 level of the community. The competition index between species  $i$  and  $j$ ,  $c_{ij}$  is  
 210 given by the overlapping degree of the sequence of substrates of species  $i$ ,  $x_i$  and  
 211 the sequence of substrates of species  $j$ ,  $x_j$ . Thus,

$$212 \quad f_{ij} = \begin{cases} \xi(y_i, x_j) & \text{if } i \neq j \\ 0 & \text{if } i = j \end{cases} \quad c_{ij} = \begin{cases} \xi(x_i, x_j) & \text{if } i \neq j \\ 0 & \text{if } i = j \end{cases} \quad (8)$$

213 Note that facilitation is directional but competition is not. This implies that  
 214  $f_{ij} \neq f_{ji}$  and  $F$  is not symmetric, but  $c_{ij} = c_{ji}$  and  $C$  is symmetric.

215 Community-level cohesion can now be formally defined as the mean value of the  
 216 matrix  $F - C$

$$217 \quad \Theta = \langle F - C \rangle \quad (9)$$

### 218 **3.4 $\Theta$ predicts the outcome of community coalescence experi-** 219 **ments.**

220 Following the community assembly and consequent measure of cohesion in all  
 221 communities, I perform community coalescence experiments to test the predic-  
 222 tive power of my metric.

223 First, I select the  $N$  communities with 5 species, and perform all  $\binom{N}{2}$  community  
 224 coalescence events in which a resident community  $C_R$  is mixed with an invading

225 one  $\mathcal{C}_I$ . At each event, calculate the similarity of between post-coalescence and  
 226 resident communities as the normalized scalar product of their species abundance  
 227 vector at stable state.

$$228 \quad S(\mathcal{C}_R, \mathcal{C}_P) = \frac{\vec{N}_\infty^R \cdot \vec{N}_\infty^P}{\sqrt{|\vec{N}_\infty^R|} \sqrt{|\vec{N}_\infty^P|}} \quad (10)$$

229

## 230 4 Discussion

- 231 • Although it certainly seems an exaggeration to view the communiity as  
 232 such a tightly regimented entity, it is also perilous to ignore the fadct that  
 233 coevolutionary processes can play an important role in communities.
- 234 • Cooperation reduces community stability, tho increases community fit-  
 235 ness... (Coyte et al. 2015)
- 236 • This thesis addresses the question of what are truly the mechanisms ex-  
 237 plaining what experiments show? An alternative measure of cohesivenes  
 238 that stems from a more realistic modeling of microbial ecosystems is able  
 239 to reproduce this results, and thus is closer to uncover what are the real  
 240 mechanisms behind community cohesion.
- 241 • community coalescence is a way to explicitly show and test the cohesiveness  
 242 of microbial communities while asking questions about how these commu-  
 243 nities came to be.
- 244 • use Pascual-García et al. (2020) for the discussion
- 245 • discuss why I chose  $m$  as the upper limit for the number of reactions that  
 246 a strain can posses.
- 247 • Discuss possible refinements of the measure of cohesion: instead of aver-  
 248 aging, I could count the number of closed loops.



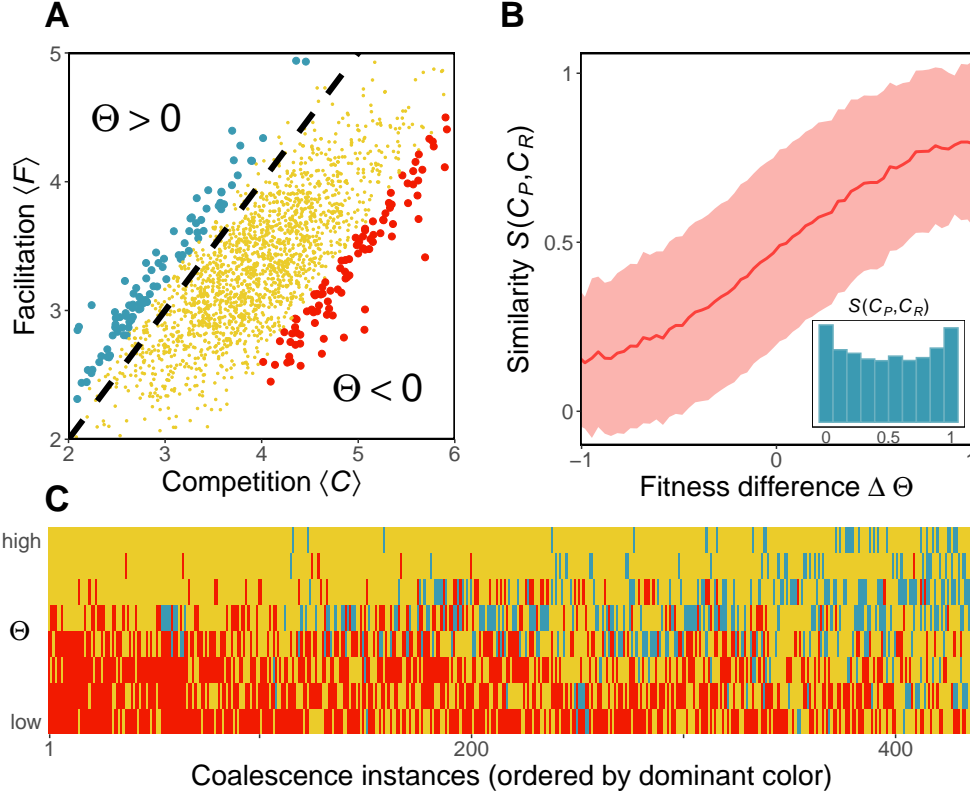


Figure 4: **Results from community coalescence experiments.** (A) Each simulated community is plotted in a competition-facilitation diagram. Communities above the dashed line  $\langle C \rangle = \langle F \rangle$  have  $\Theta > 0$ ; they are in the facilitation-dominated regime. Communities below the dashed line have  $\Theta < 0$ ; they belong to the competition-dominated regime. The extremes of each regimes are selected (blue and red dots), and coalescence experiments where one community from the blue group mixes with one community from the red group are performed, only for communities of richness 4. (C) Altruistic communities ( $\Theta > 0$ ) outperform competitive communities ( $\Theta < 0$ ) in the latter experiments. In this elimination assay, each column represents one coalescence instance, and each element in a column is a species. Extinctions are coloured to match the group in plot A to which the extinct species belonged. There is a higher proportion of extinct species from the red group (more red tiles than blue tiles). (B) The outcome of community coalescence is predicted by community-level cohesion. The similarity between the post-coalescence community and the resident community,  $S(C_P, C_R)$  is plotted as a function of the community cohesion difference  $\Delta \Theta$  between them, for all possible coalescence events between 2 communities of richness 5. Shown is binned mean (100 bins) over communities with similar  $\Delta \Theta$  (solid line)  $\pm \sigma$  (shaded) (B, inset) Histogram of similarity showing that monodominance of one community after coalescence ( $S = 0$ ,  $S = 1$ ) is more frequent than a perfect mixing ( $S = 0.5$ )

- 249 • Discuss why the traditional fitness (how fast resources are consumed) doesn't  
250 correlate with what I call fitness: community cohesion. Show that in the  
251 case of pure competition, it does (Tikhonov 2016), but in the case of purely  
252 facilitation, it doesn't
- 253 • My measure of cohesion is an approximate one. Does facilitation help the  
254 same degree that competition bothers?
- 255 • talk about environment engineering, and reference
- 256 • Maybe facilitation is actually not that important, But competition, and  
257 functional groups, the ones that drive cohesion..
- 258 • There is no allusion to individual species fitness, because here it's more  
259 important the cohesion between them.
- 260 • Talk about innovation rather than improvement when it comes to facilitation.  
261

## 262 5 Things to do in the future

- 263 • Ask Emma about papers of hierarchy of metabolites
- 264 • Find a paper that says that organisms with few and very energetic reactions  
265 are rare.
- 266 • Should I include a page at the end specifying what things I did, and  
267 what things I didn't do, and that way I don't have to do it during the  
268 paper?
- 269 • make a nice looking table of the parameters of the model.
- 270 • plot mean abundance as a function of number of reactions for both assembled  
271 communities

- 272 • List of plots I want to make: 1. The s plot with richness == 5. 2. The  
 273 histogram of similarity with richness == 4. 3. The elimination asssay. 4.  
 274 The cloud. 5. The evolution of cooperation. 6. The community reaction  
 275 network.
- 276 • Revise the cohesion of my thesis as a whole: are sections well separated?  
 277 Do they link well with each other?
- 278 • When I talk about stable equilibrium, say that A stable (non- invade-  
 279 able) equilibrium is characterized by an extra condition that all the absent  
 280 species, if re-introduced, would be driven back to extinction
- 281 • Turn to my dictionary of cool words, and use them.
- 282 • In figure 3, why  $jF_i - jC_i \neq j\Theta_i$ ? Or is it?

## 283 6 Appendix

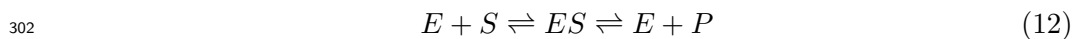
### 284 6.1 Reversible enzyme kinetics

285 Outside the the bacterial cell, the energy resides in the form of chemical potential  
 286  $\mu$  held by the metabolites, and biochemical reactions inside the cell produce  
 287 energy due to a difference in the chemical potentials of substrate and product. I  
 288 assigned chemical potentials to each metabolite according to

$$289 \mu_{\beta} = E \left( 1 - \sqrt{\frac{\beta - 1}{m - 1}} \right) \quad (11)$$

290 where  $\beta = 1, \dots, m$  and  $E$  is the energy of the most energetic metabolite. I have  
 291 chosen this chemical potential function because I hope to find papers where they  
 292 explain that there is a hierarchy on the metabolite energetic spectrum. This  
 293 means that the energy produced by a reaction of the type  $(\beta, \beta + 1)$  decreases as  
 294 you go down the hierarchy. Reactions involving metabolites situated higher in  
 295 the hierarchy are more energetic than reactions that involve those lower in the  
 296 hierarchy.

297 The rate at which a given chemical reaction transforms substrate into product  
 298 is modeled using reversible Michaelis-Menten enzyme kinetics. Thus, the model  
 299 considers chemical reactions where a substrate  $S$  binds to an enzyme  $E$  to form  
 300 an enzyme-substrate complex  $ES$ , which in turn produces a product  $P$ , and  
 301 recovers enzyme  $E$ .



303 The choice of fully reversible enzyme kinetics, instead of the traditional assump-  
 304 tion of irreversibility in the second reaction, aims to capture more accurately the  
 305 nature of biochemical reactions taking place in microbial communities. In these  
 306 reactions the Gibbs energy change  $\Delta G$  is not always high, which implies that  
 307 the reaction of product formation can reach equilibrium at a similar time scale  
 308 as the formation of the complex (Keener & Sneyd 2008). In this case, the tradi-

309 tional irreversible Michaelis-Menten scheme breaks down, and more elaborated  
 310 frameworks, like the fully reversible one that this model offers, need to be used.  
 311 To comply with  $2^{nd}$  law of thermodynamics, the network  $G_\alpha$  is completely hier-  
 312 archical, ie. the edges are unidirectional ( $x < y$ ), going from the more energetic,  
 313 to the less energetic metabolite. Thus, for the reaction scheme in 12 and the  
 314 imposed thermodynamic constraint only reactions where  $\Delta G^0 = \mu_P - \mu_S < 0$   
 315 can take place.  
 316 With all the above considerations, the expression for the rate of reaction  $i$  poss-  
 317 esed by strain  $\alpha$  is given below. A formal derivation of equation 13 can be found  
 318 in Hoh & Cord-Ruwisch (2000)

$$319 \quad q_{\alpha i} = \frac{q_m^{\alpha i} S_\alpha (1 - \theta_\alpha)}{K_S^{\alpha i} + S_\alpha (1 + k_R^{\alpha i} \theta_\alpha)} \quad (13)$$

320 Here,  $\theta_\alpha$  measures how far is the reaction from equilibrium (0 being the furthest,  
 321 and 1 being equilibrium).

$$322 \quad \theta = \frac{[P]}{[S]K_{eq}} \quad (14)$$

323 where  $[ ]$  denote concentration and  $K_{eq}$  is the equilibrium constant

$$324 \quad K_{eq} = \exp \left( \frac{-\Delta G^0 - \eta \Delta G_{ATP}}{RT} \right) \quad (15)$$

325 The energy produced by the reaction is then stored in the form of ATP molecules.  
 326 In the model,  $\eta$  represents the moles of ATP molecules produced per mole of  
 327 reaction. For a given reaction  $(x, y)$  eta I calculate eta as

$$328 \quad \eta = \frac{y - x}{m} \quad (16)$$

329 which represents the normalized metabolite gap between substrate and product  
 330 of the reaction. Therefore, the higher the gap, the more energy will be stored.

331

332 **6.2 Table of parameter values and meaning**

Parameter	Meaning	Value
m	Number of metabolites	100
s	Number of strains	10
$\Delta G_{ATP}$	ATP Gibbs energy	$7.5 \cdot 10^4$
$\mu_0$	Most energetic metabolite	$3 \cdot 10^4$
nATP	$\max\left(\frac{\Delta G_{S \rightarrow P}^0}{\Delta G_{ATP}}\right)$	4
$\eta$	Moles of ATP energy per reaction	0.5
$q_m$	Maximum reaction rate	1
$K_S$	Saturation constant	0.1
$k_r$	Reversibility constant	10
g	Growth factor	1
m	Maintenance factor	$0.2 \cdot J_{grow}$
$\kappa$	Externally supplied resource	1
$\gamma$	Dilution rate	0.5
$N_0$	Populations initial conditions	(1, 1, ..., 1)
$C_0$	Concentrations initial condition	(0, 0, ..., 0)

Table 1: Parameter meanings and their values

## References

- Castledine, M., Sierocinski, P., Padfield, D. & Buckling, A. (2020), ‘Community coalescence: An eco-evolutionary perspective’.
- Costello, E. K., Stagaman, K., Dethlefsen, L., Bohannan, B. J. & Relman, D. A. (2012), ‘The application of ecological theory toward an understanding of the human microbiome’.
- Coyte, K. Z., Schluter, J. & Foster, K. R. (2015), ‘The ecology of the microbiome: Networks, competition, and stability’, *Science* .
- Dormand, J. R. & Prince, P. J. (1980), ‘A family of embedded Runge-Kutta formulae’, *Journal of Computational and Applied Mathematics* .
- Embree, M., Liu, J. K., Al-Bassam, M. M. & Zengler, K. (2015), ‘Networks of energetic and metabolic interactions define dynamics in microbial communities’, *Proceedings of the National Academy of Sciences of the United States of America* .
- Falkowski, P. G., Fenchel, T. & Delong, E. F. (2008), ‘The microbial engines that drive earth’s biogeochemical cycles’.
- Friedman, J., Higgins, L. M. & Gore, J. (2017), ‘Community structure follows simple assembly rules in microbial microcosms’, *Nature Ecology and Evolution* .
- Gilbert, J. A., Jansson, J. K. & Knight, R. (2014), ‘The Earth Microbiome project: Successes and aspirations’.
- Gilpin, M. (1994), ‘Community-level competition: Asymmetrical dominance’, *Proceedings of the National Academy of Sciences of the United States of America* .

- Goldford, J. E., Lu, N., Bajić, D., Estrela, S., Tikhonov, M., Sanchez-Gorostiaga, A., Segrè, D., Mehta, P. & Sanchez, A. (2018), ‘Emergent simplicity in microbial community assembly’, *Science* .
- Goyal, A. & Maslov, S. (2018), ‘Diversity, Stability, and Reproducibility in Stochastically Assembled Microbial Ecosystems’, *Physical Review Letters* .
- Hansen, S. K., Rainey, P. B., Haagenzen, J. A. & Molin, S. (2007), ‘Evolution of species interactions in a biofilm community’, *Nature* .
- Hoh, C. Y. & Cord-Ruwisch, R. (2000), ‘A practical kinetic model that considers endproduct inhibition in anaerobic digestion processes by including the equilibrium constant’, *Biotechnology and Bioengineering* .
- Huttenhower, C., Gevers, D., Knight, R. & Al., E. (2012), ‘Structure, function and diversity of the healthy human microbiome’, *Nature* .
- Keener, J. P. & Sneyd, J. (2008), ‘Mathematical Physiology’, *Book* .
- Lawrence, D., Fiegna, F., Behrends, V., Bundy, J. G., Phillimore, A. B., Bell, T. & Barraclough, T. G. (2012), ‘Species interactions alter evolutionary responses to a novel environment’, *PLoS Biology* .
- Livingston, G., Jiang, Y., Fox, J. W. & Leibold, M. A. (2013), ‘The dynamics of community assembly under sudden mixing in experimental microcosms’, *Ecology* .
- Lu, N., Sanchez-gorostiaga, A., Tikhonov, M. & Sanchez, A. (2018), ‘Cohesiveness in microbial community coalescence’, *bioRxiv* .
- Marsland, R., Cui, W., Goldford, J., Sanchez, A., Korolev, K. & Mehta, P. (2019), ‘Available energy fluxes drive a transition in the diversity, stability, and functional structure of microbial communities’, *PLoS Computational Biology* .



- Pascual-García, A., Bonhoeffer, S. & Bell, T. (2020), ‘, which can be found in subsection 3.3 (page 11).’.
- Rillig, M. C., Antonovics, J., Caruso, T., Lehmann, A., Powell, J. R., Veresoglou, S. D. & Verbruggen, E. (2015), ‘Interchange of entire communities: Microbial community coalescence’.
- Sierocinski, P., Milferstedt, K., Bayer, F., Großkopf, T., Alston, M., Bastkowski, S., Swarbreck, D., Hobbs, P. J., Soyer, O. S., Hamelin, J. & Buckling, A. (2017), ‘A Single Community Dominates Structure and Function of a Mixture of Multiple Methanogenic Communities’, *Current Biology* .
- Tikhonov, M. (2016), ‘Community-level cohesion without cooperation’, *eLife* .
- Tikhonov, M. & Monasson, R. (2017), ‘Collective Phase in Resource Competition in a Highly Diverse Ecosystem’, *Physical Review Letters* .
- Toquenaga, Y. (1997), ‘Historicity of a simple competition model’, *Journal of Theoretical Biology* .