

# Strong self-regulation and widespread facilitative interactions between groups of phytoplankton

December 21, 2018

## Abstract

XXX(introductory paragraph: 150 words)

## Introduction

How species or related genera can coexist together in spite of likely competition is still one of the main puzzles of community ecology, especially for primary producers that seemingly share the same basic resources [REFS]. Many theoretical studies of competition models have shown that competitive exclusion is likely in those circumstances [REFs], unless some special mechanisms involving spatial or temporal variation are at play<sup>1,2,3</sup> [More refs?]. Neutral theory, that in its most basic formulation assumes equal competitive abilities [REFs] and non-equilibrium coexistence sustained by immigration from a larger pool (but see ref.<sup>4,5</sup>) has been proposed as a solution to explain highly diverse communities<sup>6</sup> [More refs?]. However, the evidence gathered from plant communities starts to suggest that, in fact, niche rather than neutral processes may be paramount, with intraspecific competition dwarfing interspecific competition in most cases<sup>7,8</sup>. Whether these conclusions apply to other systems is currently unknown.

Moreover, the meta-analysis by Adler et al.<sup>8</sup> reported a large number of facilitative interactions (26%) and reviews<sup>9</sup> [More refs] have highlighted that in fact that facilitation may be much more widespread than ecologists usually tend to think. Although some theoretical studies suggest that facilitative interactions can be destabilizing (sensu resilience) and therefore undermine coexistence<sup>10</sup>, many modelling studies<sup>11</sup> [More refs] have suggested that facilitative interactions can to a large degree benefit coexistence.

Here, we analyse a large dataset of several long-term time series of phytoplankton dynamics along the French coastline , using statistical time series models allowing for interactions between groups.

## Results and discussion

Using MAR(1) autoregressive models, we have produced interaction matrices – Jacobian community matrices on the logarithmic scale<sup>12</sup> – that have two main features. First, we observe a strong self-regulation for all sites (Fig. 1, diagonal elements of all matrices), a feature that we have highlighted on a previous, more detailed analysis on one of the considered study sites only<sup>13</sup>. Second, although the percentage of facilitative interactions seem to vary among sites (between 40% and 71%), we have found that it is substantial for all sites: contrary to our initial expectations, commensalism (the interactions are rarely two-way mutualism [check][Will do]) seems to be the rule rather than the exception. This high percentage of positive interactions remains present within a whole region: consensus on positive interactions varies between 30% and 53% and is higher than the percentage of negative interactions, except for sites in the Mediterranean Sea.

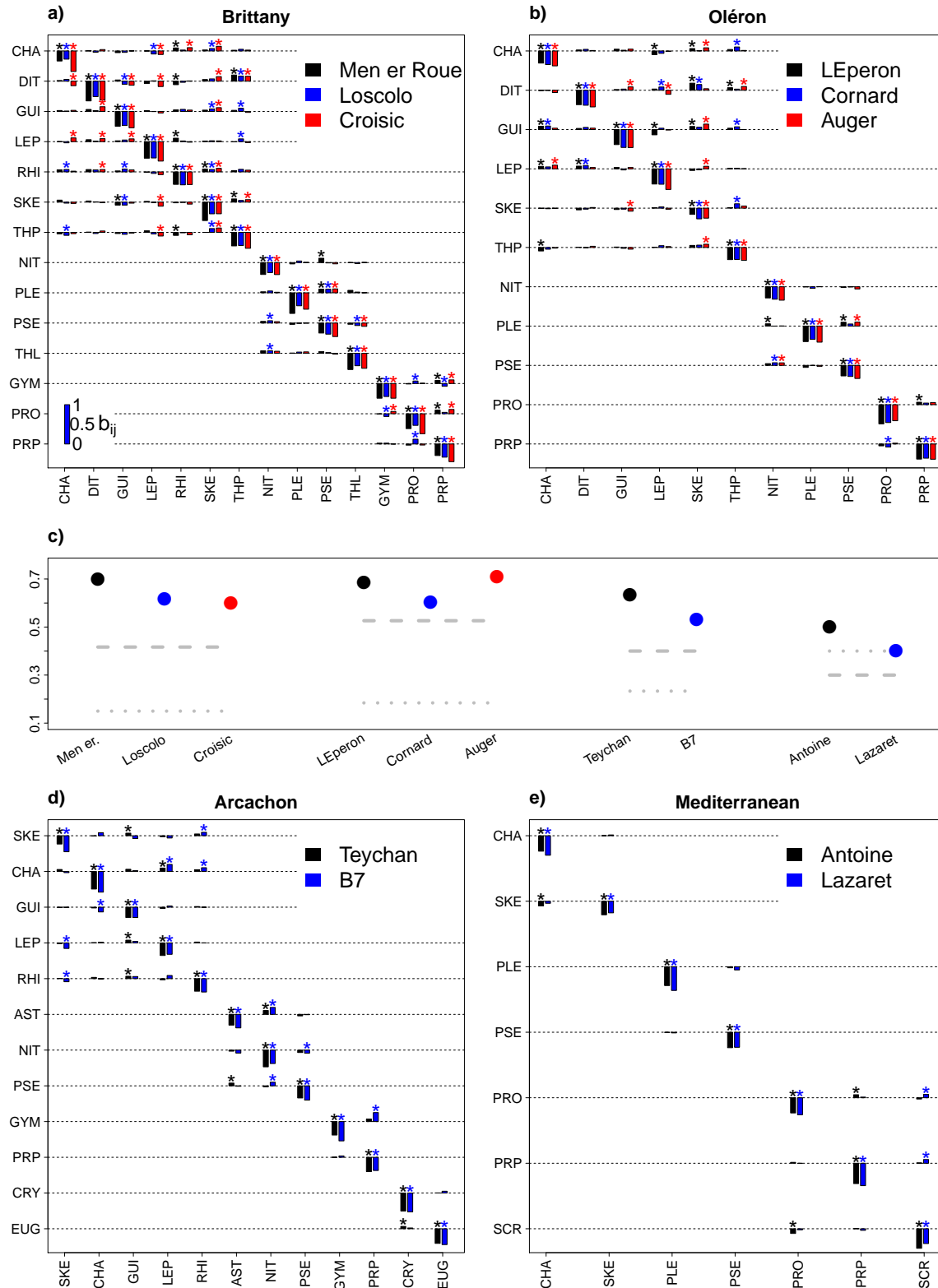


Figure 1: **Interaction matrices estimated in 10 sites along the French coastline.** Four regions are distinguished: Brittany (a), Marennes-Oléron (b), Arcachon (d) and the Mediterranean Sea (e). Only interactions between clades (pennate and centric diatoms, dinoflagellates, other planktonic taxa) are allowed. The figure should be read as taxon  $i$  having effect  $e_{ji}$  on taxon  $j$ . We present the community matrix minus the identity matrix ( $\mathbf{B} - \mathbf{I}$ ) for comparability reasons between intra- and inter-group interaction strengths. The scale for the coefficient values is given at the bottom left of panel a). 95% significance of coefficients was determined by bootstrapping and is marked by asterisks (\*). The composition of planktonic groups is given in the Supplementary Information. The ratio of positive interactions in each matrix is given by points in c) while the dashed (respectively, dotted) line represents the ratio of interactions remaining positive (respectively, negative) in all sites of a given region.

We then asked whether the stability (resilience) of those matrices was affected by the percentage of mutualistic interactions or their connectivity properties (Fig. 2).

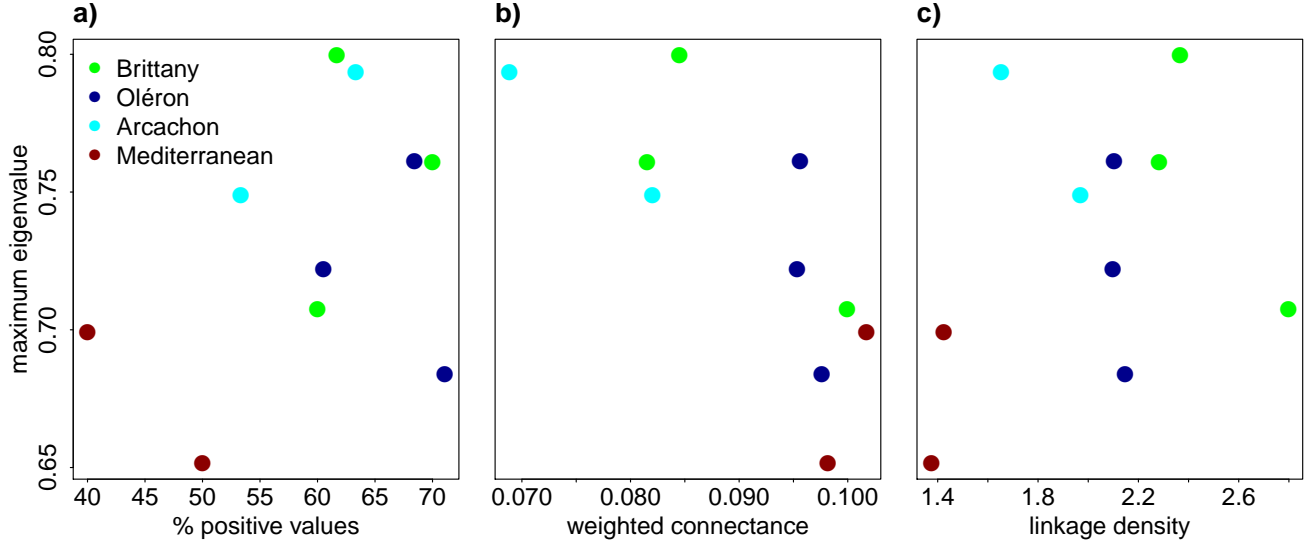


Figure 2: **Relation between stability and complexity of the interaction networks.** The maximum modulus of the community matrix eigenvalues approximates the stability while complexity is indicated by different network metrics. Each point color corresponds to a given region. Metrics formula for weighted connectance and linkage density are given in the Supplementary Information.

More information than sign and the magnitude of interspecific / intergroup interactions can be gathered from interaction matrices. We investigated whether the matrix has some particular structure that could help theoretical ecology to make better null models of joint community dynamics and interactions<sup>14</sup>. We defined two scores, vulnerability (summed effect of others on the focal species growth rate) and generality (summed effect of the focal species onto other species growth rates). Species that were more self-regulating also had also a higher vulnerability score, and those two influences are likely to trade-off: a high degree of self-regulation somehow buffers outside influences. Species that were less self-regulating were also more likely to have a broad range of effects onto other species [to interpret this we may also need to check that they were more abundant, if so I would say that species varying broadly have a higher likelihood to generate effects in other species][Will do]. It is important to note, however, that these patterns are weak and there is therefore a considerable amount of randomness dominating the interaction matrix: many scenarios of self-regulation vs limitation by others are therefore possible.

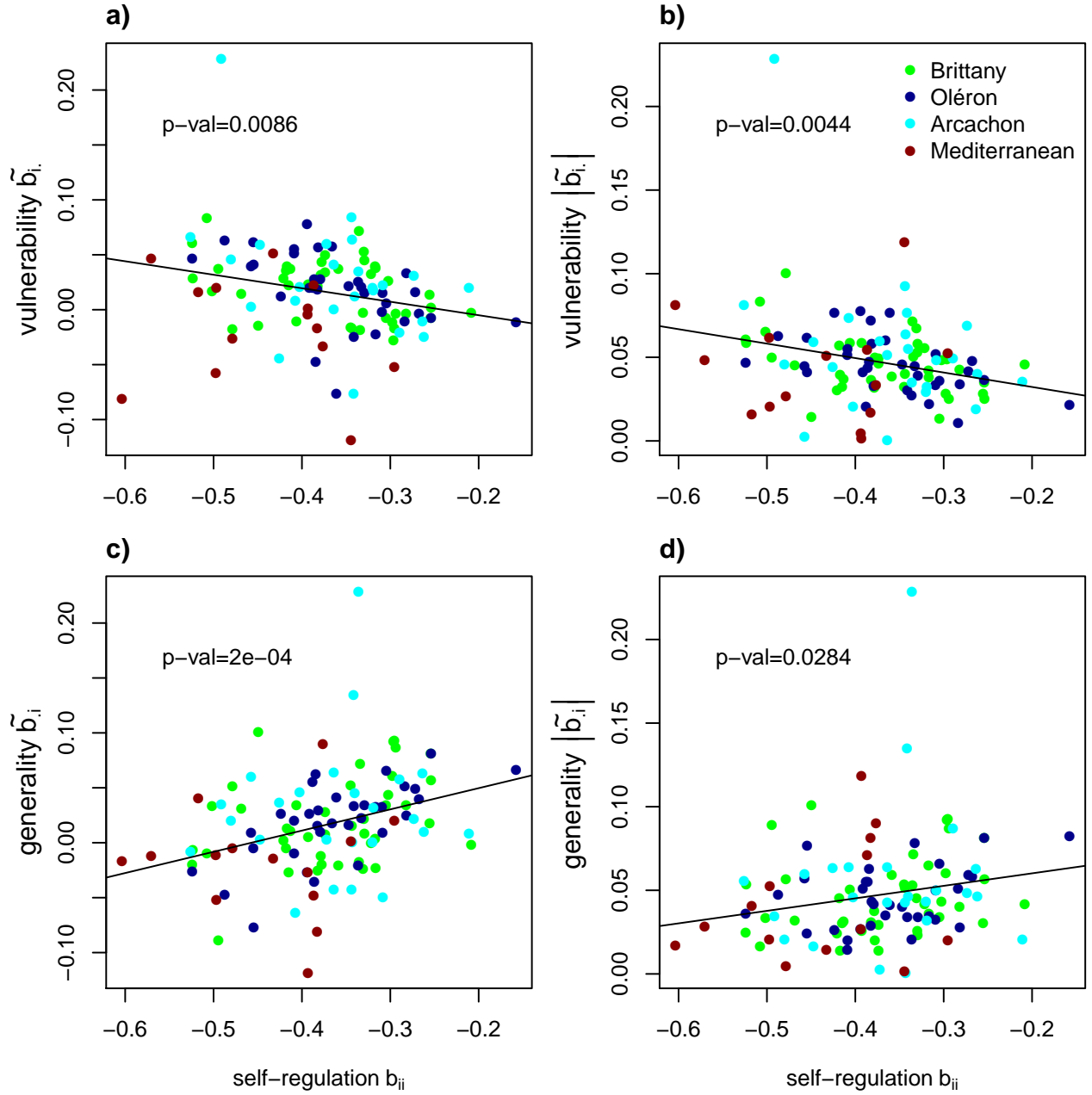


Figure 3: **Relation between vulnerability/generalizability and self-regulation.** Average vulnerability (a-b) and generality (c-d), as well as self-regulation, are computed on raw (a-c) or absolute (b-d) values of the coefficients in the community matrix minus the identity matrix ( $\mathbf{B} - \mathbf{I}$ ) for 10 different sites. Each color corresponds to a given region. The p-value of the Pearson correlation between vulnerability (respectively generality) and self-regulation is given in the top left of each panel and the corresponding slope is given by the black line across scatter plots.

Aside from these trade-offs, some of which promote some stability (sensu invariability), we found no remarkable patterns of covariation between matrix elements (other than a mean-variance scaling of interaction coefficients, see SI XXX Fig. XXX). The big finding of our study seems really to be the presence of very large niche differences between the species, translating into a much higher intragroup competition, together with a high degree of facilitative net interactions. In our previous publication<sup>13</sup>, we have argued that those could arise from predation effects that could

well create apparent mutualism between prey species and strong intragroup density-dependence [REFs]. This is still very much likely here given our new study regions (Oléron, Brittany, Mediterranean) could have similar predators or parasites, though there are other possibilities [name them, will do].

Finally, we have sought to put these results in a broader context by compiling the intra vs. inter group estimates of previous MAR(1) studies (detailed in Appendix XXX). We have found that the order of magnitude of intra/inter interaction strengths considered here ( $\approx 10$ ) not particularly below those found for most planktonic systems. We include in Fig. 4 not only plankton studies but also a couple of vertebrate or insect studies on smaller (less diverse) communities where interactions are stronger. The conclusion from this comparison seems to be that any diverse system of competitors and facilitators seem to have evolved large niche differences making intragroup competition much higher than intergroup interactions.

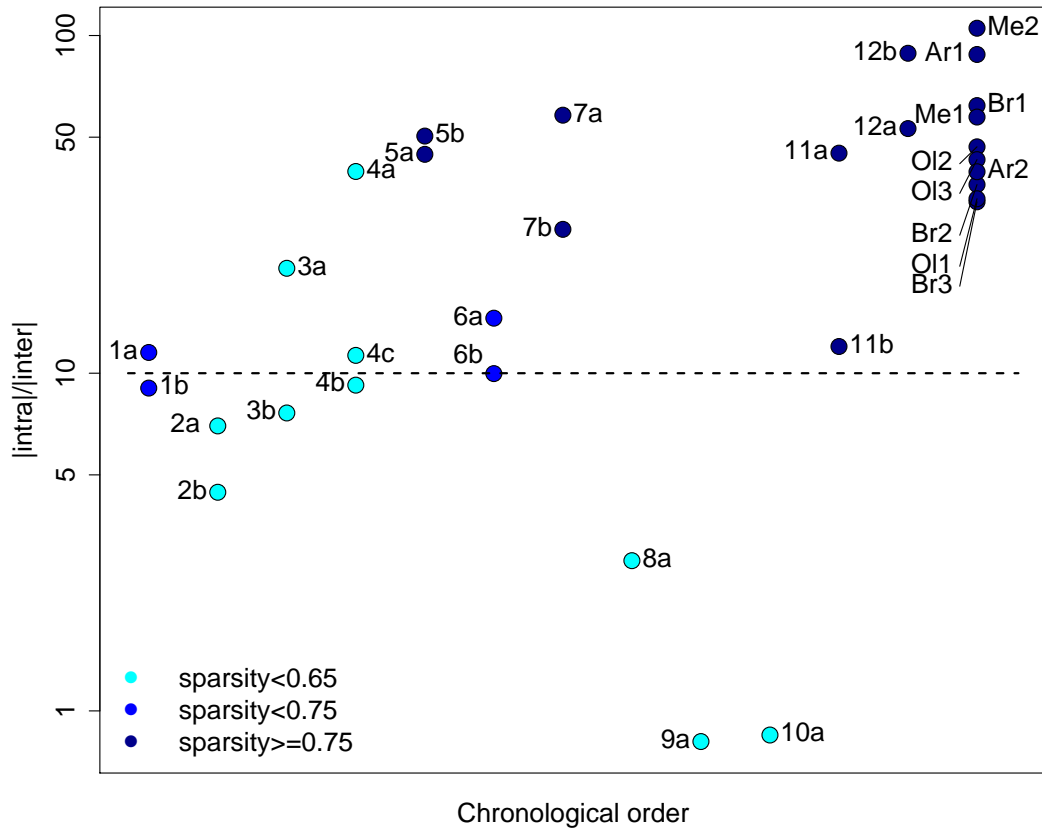


Figure 4: **Ratio of intra- and inter-group interaction strengths in Multivariate AutoRegressive (MAR) models in the literature.** Ratio is given on the log scale and the name of each studies is given in the Supplementary Information. The shape of each point corresponds to the number of species in the interaction matrix and its color corresponds to the sparsity of the community matrix (e.g., the proportion of null interactions in the matrix). Intergroup interactions were set to 0 when they were not given in the studies (most of the time, because they were not deemed significant at the 95% threshold). Last column corresponds to the present study.

This dominance of niche differentiation is similar to what has been found in plant community studies<sup>5,8</sup>, and might be due to the ghost of the competition past, i.e., selection and evolution leading to progressive niche separation. Species have evolved niches that allow them not to compete or to interact only weakly (very strong facilitative

effects might be likewise destabilizing [je crois qu’il y a des refs sur les effets non-monotone des interactions +, je rechercherai]. The predator effects that we highlighted above can be comprised within such niche differentiation *sensu largo*: specialized predators can make strong conspecific density-dependence emerge [REFs], while switching generalists can also promote diversity [REFs, same from Oikos]. Both predators and resources have often symmetrical effects and can contribute to niche differentiation<sup>15</sup>

An intriguing new possibility, dubbed the “ghost of competition” present<sup>16</sup>, suggests that spatial patterns in relation to abiotic factors might have a large impact on the patterns inferred from interaction models. Recent combinations of model fitting and removal experiments have shown that the model fitting usually underestimate the effect of competitors that are uncovered by removal experiments<sup>16,17</sup>. This could occur for instance if species are spatially segregated (at a small scale) because each species only exists within a domain where it is relatively competitive, while a focal species could spread out if competitors were removed. This means that a species can be limited by competitors, but act so as to minimize competition (a little like avoidance behaviour in animals) so that competition is in effect hard to detect when all species are present. This would require some fine-scale segregation between phytoplankton species at the scale of interactions, i.e., at the microscale. At the moment, it is known that turbulence generates inhomogeneities at the microscale<sup>18,19</sup> but it is quite unclear how this affects multiple interacting species and which multivariate spatial patterns (as in ref.<sup>20</sup>) could be expected. Also, even if turbulence generates spatial structure with segregation between species, it is not quite clear that the “ghost of competition present” mechanism could work for plankton, where the turbulence may somehow dictates where the plankton patches can or cannot appear.

## Methods

### Data collection

Data were collected for the National Phytoplankton and Phycotoxin Monitoring Network (REPHY<sup>21</sup>). Since 1987, this monitoring program has required 26 sites along the French coastline to be sampled every 2 weeks within 2 hours of high tide to determine both biotic (phytoplankton counts) and abiotic (water temperature, salinity) parameters. We chose to focus on sites which had been sampled at least 10 years after 1996, when methods were standardized among sites. We also excluded time series which had missing data for over 6 months or more than 20 days, in average, between sampling dates. This reduced the number of sampling points to 10 sites in 4 regions (Brittany, Marennes-Oléron, Arcachon and the Mediterranean Sea; a map is shown in the Supplementary Information).

Abiotic parameters were measured directly from the boat during the sampling process while water samples for biotic analyses were fixed with a Lugol’s solution and examined later. Phytoplankton cells above 20  $\mu\text{m}$  were identified at the lowest possible taxonomic level and counted with the Utermöhl method using an optical microscope<sup>22</sup>. Throughout the years and sites, more than 600 taxa were identified at different taxonomic levels. We aggregated

them at the genus, or group of genera, level based on previous work<sup>23</sup>, except for cryptophytes and euglenophytes in Arcachon, which could not be identified below the family level. For each region, we focused on the most abundant and most frequently observed genera to avoid most of the gaps in the time series. When gaps did not exceed a month, missing values were linearly interpolated and remaining missing values were replaced by a random number between 0 and half of the lowest observed abundance<sup>24</sup>.

## MAR(1) model

Multivariate autoregressive (MAR) models are often used to determine the interspecific interactions and abiotic parameters shaping a community<sup>12</sup>. They are based on a stochastic, discrete-time Gompertz model which relates log-abundance of  $S$  species at time  $t+1$  to interactions with the rest of the community at time  $t$ , and effects of  $V$  abiotic variables at time  $t+1$ , following eq. 1.

$$\mathbf{n}_{t+1} = \mathbf{B}\mathbf{n}_t + \mathbf{C}\mathbf{u}_{t+1} + \mathbf{e}_t, \mathbf{e}_t \sim \mathcal{N}_S(0, \mathbf{Q}) \quad (1)$$

where  $\mathbf{n}_t$  is the  $1 \times S$  log abundance vector of abundance of phytoplankton groups,  $\mathbf{B}$  is the  $S \times S$  community (interaction) matrix,  $\mathbf{C}$  is the  $S \times V$  environment matrix describing the effects of  $V$  variables  $\mathbf{u}_{t+1}$  on species growth, and  $\mathbf{e}_t$  is a  $1 \times S$  noise vector which covers both process and observation error, following a multivariate normal distribution with a variance-covariance matrix  $\mathbf{Q}$ .  $\mathbf{Q}$  is diagonal and we have previously showed that this parsimonious choice did not affect qualitative results<sup>13</sup>.

We used the MARSS package<sup>25</sup>, in R, to estimate parameters with a maximum likelihood procedure.

We have previously published a very detailed analysis of one of the dataset (Arcachon) for which more covariables were available<sup>13</sup>. The analysis of real data was complemented by that of simulated data for a plankton time series set-up and confirmed the ability of the technique to infer biotic interactions and abiotic forcing. This analysis revealed that the MAR(1) approach was satisfactory (i.e., no need for extra non-linearities to model the storage effect, which was found to be weak as in previous analyses of plant data for which strong-self regulation was observed<sup>7,26</sup>). Furthermore, the use of a slightly less exhaustive data set in the current study led to similar results to the ones obtained previously in Arcachon<sup>13</sup>. We are therefore confident that the MAR(1) models presented here are appropriate models. In general, MAR(1) models tend to be fairly robust to small deviations of the underlying data-generating model, provided that one asks mainly order of magnitude of coefficients values (rather than very precise estimates) and sign of interaction coefficients<sup>27</sup>. [for ease of interpretation we also present the connexion between MAR(1) model coefficients and a Beverton-Holt model in SI XX [see also ref.<sup>27</sup>]]

In this study, the number of phytoplankton groups,  $S$ , varies between region but we keep the same 2 covariates, i.e. water temperature and salinity. Therefore, the dimension of the equation system only depends on the (square of the) number of phytoplankton groups we study, which ranges between 7 (Mediterranean Sea) and 14 (Brittany).



The smallest system still requires 70 parameters to be estimated if we consider all possible interactions between species. Before analyzing these parameters, we wondered if such parameterization was necessary and compared different 'interaction scenarios' based on BIC. Our null model assumed no interaction between groups of species (diagonal interaction matrix) and was compared to four other models. The first model assumed that interactions could only occur between phylogenetically close organisms, i.e., within a class (namely, diatoms, dinoflagellates, and other phytoplanktonic organisms) while the second model further differentiated pennate and centric diatoms. The third model considered the inverse hypothesis that only unrelated organisms could interact (i.e., a diatom could only interact with a dinoflagellate or a cryptophyte, but not with another diatom), and the last model did not constrain the interactions at all (full interaction matrix). Based on convergent results in all sites (see Supplementary Information), we further investigated the second model, hereafter called the pennate-centric model.

## Interaction matrix analysis

The community matrix obtained from MAR(1) analyses can be used to determine the stability of a system<sup>12</sup>. We compared the maximum modulus of the eigenvalues of the pennate/centric matrices in each site, as a proxy of stability, to network metrics which could be related to complexity, such as weighted connectance and linkage density<sup>19</sup>. Weighted connectance is a measure of the proportion of realized links, taking into account the shape of the flux distribution, while link density measures the average proportion and strength of interactions for a given species. These metrics are adapted to weighted interaction matrix but cannot accommodate for both positive and negative coefficients: we therefore chose to focus on the absolute values of these coefficients, which can be linked to their strength even though we do not necessarily know the kind of interaction it describes.

In addition to these network-level metrics, we also computed the average vulnerability (average effect of other taxa on a focal taxon) and generality (average effect of a focal taxon on other taxa) and compared these to the regulation a focal species exerted on itself.

Finally, we compared our results on self regulation/intraspecific interactions v. interspecific interactions to other published studies based on the MAR model. A list of references is given in the Supplementary Information. Authors usually reported only coefficients that were significant at the 95% threshold, ignoring weak effects. We therefore computed the mean interspecific interactions as either the mean value of significant coefficients only (see Supplementary Information), or the mean value of all coefficients outside of the matrix diagonal, replacing missing values by 0 (Fig. 4). We should nevertheless mention two biases in this comparison: low-dimension matrices tended to be more complete (less sparse) than high-dimension matrices as these small community matrix were used to study known interaction phenomena (observed predation between organisms, for instance). There is therefore a positive correlation between sparsity and dimensionality (see SI). The number of parameters to estimate increase as the square of the number of interacting groups, leading authors to reduce this set before the estimation process. While we based our model choice on phylogeny (see above), the typical way of reducing the number of estimated parameters

is automated and usually based on the comparison of 100 randomly chosen interaction matrices, optimizing BIC<sup>12</sup>. [This means that, even for high-dimension matrices, we are not comparing exactly the same interaction models.]

## References

- [1] P. Chesson and N. Huntly. The roles of harsh and fluctuating conditions in the dynamics of ecological communities. *The american naturalist*, 150(5):519–553, 1997.
- [2] J. Huisman and F.J. Weissing. Biological conditions for oscillations and chaos generated by multispecies competition. *Ecology*, 82(10):2682–2695, 2001.
- [3] L. Li and P. Chesson. The Effects of Dynamical Rates on Species Coexistence in a Variable Environment: The Paradox of the Plankton Revisited. *The american naturalist*, 188(2):E46–E58, 2016.
- [4] I. Volkov, J.R. Banavar, S.P. Hubbell, and A. Maritan. Neutral theory and relative species abundance in ecology. *Nature*, 424:1035–1037, 2003.
- [5] I. Volkov, J.R. Banavar, S.P. Hubbell, and A. Maritan. Patterns of relative species abundance in rainforests and coral reefs. *Nature*, 450(7166):45–49, 2007.
- [6] S.P. Hubbell. *The Unified Neutral Theory of Biodiversity and Biogeography (MPB-32)*. Princeton University Press, 2001. ISBN: 978-0-691-02128-7.
- [7] P.B. Adler, S.P. Ellner, and J.M. Levine. Coexistence of perennial plants: an embarrassment of niches. *Ecology letters*, 13(8):1019–1029, 2010.
- [8] P.B. Adler, D. Smull, K.H. Beard, R.T. Choi, T. Furniss, A. Kulmatiski, J.M. Meiners, A.T. Tredennick, and K.E. Veblen. Competition and coexistence in plant communities: intraspecific competition is stronger than interspecific competition. *Ecology letters*, 21(9):1319–1329, 2018.
- [9] R.W. Brooker, F.T. Maestre, R.M. Callaway, C.L. Lortie, L.A. Cavieres, G. Kunstler, P. Liancourt, K. Tielbörger, J.M.J. Travis, F. Anthelme, C. Armas, L. Coll, E. Corcket, S. Delzon, E. Forey, Z. Kikvidze, J. Olofsson, F. Pugnaire, C.L. Quiroz, P. Saccone, K. Schiffers, M. Seifan, B. Touzard, and R. Michalet. Facilitation in plant communities: the past, the present, and the future. *Journal of ecology*, 96(1):18–34, 2008.
- [10] K.Z. Coyte, J. Schluter, and K.R. Foster. The ecology of the microbiome: Networks, competition, and stability. *Science*, 350(6261):663–666, 2015.
- [11] K. Gross. Positive interactions among competitors can produce species-rich communities. *Ecology letters*, 11(9):929–936, 2008.
- [12] A. R. Ives, B. Dennis, K. L. Cottingham, and S. R. Carpenter. Estimating community stability and ecological interactions from time-series data. *Ecological monographs*, 73(2):301–330, 2003.

- [13] F. Barraquand, C. Picoche, D. Maurer, L. Carassou, and I. Auby. Coastal phytoplankton community dynamics and coexistence driven by intragroup density-dependence, light and hydrodynamics. *Oikos*, 127(12):1834–1852, 2018.
- [14] A. James, M.J. Plank, A.G. Rossberg, J. Beecham, M. Emmerson, and J.W. Pitchford. Constructing Random Matrices to Represent Real Ecosystems. *The american naturalist*, 185(5):680–692, 2015.
- [15] P. Chesson. Updates on mechanisms of maintenance of species diversity. *Journal of ecology*, 106(5):1773–1794, 2018.
- [16] S.L. Tuck, J. Porter, M. Rees, and L.A. Turnbull. Strong responses from weakly interacting species. *Ecology letters*, 21(12):1845–1852, 2018.
- [17] P.B. Adler, A. Kleinhesselink, G. Hooker, J.B. Taylor, B. Teller, and S.P. Ellner. Weak interspecific interactions in a sagebrush steppe? Conflicting evidence from observations and experiments. *Ecology*, 99(7):1621–1632, 2018.
- [18] A.D. Barton, B.A. Ward, R.G. Williams, and M.J. Follows. The impact of fine-scale turbulence on phytoplankton community structure: Phytoplankton and turbulence. *Limnology and oceanography: fluids and environments*, 4(1):34–49, 2014.
- [19] R.E. Breier, C.C. Lalescu, D. Waas, M. Wilczek, and M.G. Mazza. Emergence of phytoplankton patchiness at small scales in mild turbulence. *Proceedings of the national academy of sciences*, 115(48):12112–12117, 2018.
- [20] D.J. Murrell and R. Law. Heteromyopia and the spatial coexistence of similar competitors. *Ecology letters*, 6(1):48–59, 2003.
- [21] REPHY French Observation, Monitoring program for Phytoplankton, and Hydrology in coastal waters. Rephy dataset - french observation and monitoring program for phytoplankton and hydrology in coastal waters. 1987-2016 metropolitan data. 2017. DOI: 10.17882/47248.
- [22] H Utermöhl. Zur Vervollkommnung der quantitativen Phytoplankton-Methodik. *Mitt. int. ver. theor. angew. limnol.*, 9, 1958.
- [23] T. Hernández Fariñas, C. Bacher, D. Soudant, C. Belin, and L. Barillé. Assessing phytoplankton realized niches using a French national phytoplankton monitoring network. *Estuarine, coastal and shelf science*, 159:15–27, 2015.
- [24] S.E. Hampton, M.D. Scheuerell, and D.E. Schindler. Coalescence in the lake washington story: Interaction strengths in a planktonic food web. *Limnology and oceanography*, 51(5):2042–2051, 2006.
- [25] E. E. Holmes, E. J. Ward, and M. D. Scheuerell. Analysis of multivariate time-series using the MARSS package. *User guide: <http://cran.r-project.org/web/packages/marss/vignettes/userguide.pdf>*, 2014.

- [26] S.P. Ellner, R.E. Snyder, and P.B. Adler. How to quantify the temporal storage effect using simulations instead of math. *Ecology letters*, 19(11):1333–1342, 2016.
- [27] G. Certain, F. Barraquand, and A. Gårdmark. How do MAR(1) models cope with hidden nonlinearities in ecological dynamics? *Methods in ecology and evolution*, 9(9):1975–1995, 2018.