Responses to Reviewers' Comments for Manuscript

The distribution of distances to the edge of species coexistence

Addressed Comments for Publication to

PCI Journal

by

Mario Desallais, Michel Loreau et Jean-François Arnoldi

Dear Dr. De Laender,

Please find enclosed the revised version of our previous submission entitled “The distribution of distances to the edge of species coexistence” (previously entitled "On the robustness of species coexistence to environmental perturbations" on the first version). It has been reworked to suit the new PCI Ecology standards (regarding formatting). This includes the separation of the appendices, which are put online on a repository. We have also enclosed a document showing the substantive differences between the first and second versions of the article. We would like to thank you and the reviewers for the valuable comments which help improving the quality of our manuscript. In this revision, we have carefully addressed the reviewers’ comments. A summary of main modifications and a detailed point-by-point response to the comments from Reviewers 1 to 3 (following the reviewers’ order in the decision letter) are given below.

Sincerely,

Mario Desallais, in the name of the authors

Note: To enhance the legibility of this response letter, all the editor’s and reviewers’ comments are typeset in blue boxes. Our responses are below in standard text. Main changes are in grey boxes. Line numbers at the beggening of theses grey boxes refer to the version of the manuscript with track changes. Rephrased or added important sentences are typeset in blue and important sentences that have been removed are typeset in red. For sake of readability, citations from the first manuscript are not shown here, only new bibliographical additions are annotated. Finally, in case of unintentional differences between the documents, always refer to the PDF of the resubmitted manuscript.

Authors' Response to the Editor

General Comments. The manuscript "On the robustness of species coexistence to environmental perturbations" uses the geometrical concept of the feasibility domain to ask how much one can perturb intrinsic growth rates without losing species. The focus of the paper is on the smallest perturbations that would push a community dangerously close to species loss. A first analysis gains some intuition with a 3-species case, showing that the proportion of growth rate vectors within a distance z of the feasibility domain's edges is proportional to the radius of the largest disk one can fit in this feasibility domain. Next, this idea is extended to higher dimensions, showing how robustness (defined as the minimum amount of growth rate perturbation to cause extinction) depends on essentialy species interactions and species richness. Then, the probability of the distance to the feasibility domain's edge is approximated for two different scenarios, and insight is gained on the contribution of individual species to robustness. Finally, there is an application of the theory to experimental data.

The three reviewers are overall positive and see the merit of this work, but also highlight some room for improvement. Most of these comments revolve around a more complete appraisal of similar efforts in this field (which may have gone unnoticed by the authors), a more careful pronunciation of one of the study's premise (coexistence ensues if feasibility holds), and some numerical experiments to test how the theory predicts consequences of "real world" parameter perturbations.

Response: We very appreciate your handling of the review process. The first version of our article seems to have been very well understood and the resulting feedback has been a great support. According to the reviewers' comments, we have checked our manuscript and addressed them in the following way:

1. We have better contextualised our study with regard to recent literature on the feasibility domain. To this end, major changes have been made to the abstract, introduction and the general structure of the article. The previous manuscript gave

the impression that the study differed from the rest of the literature in its approach of including the shape of the feasibility domain and taking ecological disturbance into account. This was a mistake, as other recent studies have gone in this direction. In this new manuscript, we have sought to isolate and highlight what makes this study original: the $p(z)$ approximation of the edge distance distribution function and the definition of a contextual biotic role in coexistence.

2. We have greatly simplified the technical and mathematical aspects of our study by presenting the important results directly (starting with the $p(z)$ function) without going into the details of the demonstration. The latter is the subject of a new mathematical appendix, which is very complete and detailed. You can find it online, alongside the other appendices (URL and DOI are listed in the "Data, script, code, and supplementary information availability" section of the manuscript).
3. A new appendix section has been added to test, using L-V model simulations, the link between measures derived from $p(z)$ and the actual persistence of communities and species in the face of environmental disturbances. Note that this test does not have a very strong proof value, since it only concerns communities simulated via a generalized lotka-volterra model with random parameters. This, plus the fact that it does not constitute the core of the novelties presented by our study, have prompted us to place these results in an appendix. However, in the absence of empirical data, we believe it will provide a preliminary insight into the relevance of using $p(z)$ and the D_* and SV_i metrics derived from it.

Comment 1

I very much enjoyed this paper and agree with the reviewers it is an important step forward. However, I feel the section “Distribution of distances to the edge of feasibility” is quite technical and risks shooing away many generalist readers, appealing almost exclusively to scientists with the right mathematical chops. If this is the intended audience, fine for me, but I think the content of the paper merits a broader readership, as some of the implications of the results are potentially important (as explained very well in the excellent discussion). One approach could be to develop a more detailed/technical supp. document for people like me who want to understand every detail (which I was not able given only the main text) and then focus on the main equations and what they mean ecologically in the main text. Regarding literature that could be included in the intro or discussion, I spontaneously think of Cenci et al’s paper (10.1016/j.jtbi.2017.10.016) and our recent paper (https://onlinelibrary.wiley.com/doi/abs/10.1111/ele.14278), which both integrate environmental change into feasibility domain theory.

Response: Thank you for your criticism and compliments! After reading all the reviewers’ comments, we’ve decided to condense the technical part of the article by deleting a number of sections (and with them, some of the mathematical details). Namely, the sections "Perturbation intensity as a distance to the edge of feasibility", "The most robust state of coexistence and the radius D_* of the inscribed disc", and "Edge effects in high dimensions, and the role of species richness". The idea is to move directly from the explanation of $p(z)$ for standard triangles to the presentation of the $p(z)$ function applied to feasibility domains and ecological systems. A full explanation of this function is now provided in a full mathematical appendix, on your advice. The new section is entitled "Distribution of distances of the edge of a feasibility domain" and regroups five previous subsections of the original manuscript. This way of condensing the technical part of the manuscript leading to $p(z)$ had the side effect of highlighting this result. Since the latter is precisely one of the new elements introduced by our study and allows us, among other things, to distinguish ourselves from nearby and recent studies (e.g. Lepori et al. 2024, Allen-Perkins et al. 2023), this should support the relevance of these changes

Authors' Response to Reviewer 1

General Comments. This work quantitatively evaluates the fragility of ecological communities to loose species given the size and shape of the feasibility domain. The wide adoption of the study of structural stability in ecology focus first on the size of the feasibility domain, which determines the range of opportunities for species to coexist (Rohr et al. 2014, Saavedra et al. 2017, Godoy et al. 2018). Aftwerwards, researchers realized the importance of the shape of such feasiblity domain for understanding species coexistence as not all species have equal probabilities to persist (Medeiros et al. 2021, Allen-Perkins et al. 2023, Lepori et al. 2024). This study follows this line of research evaluating the robustness of species coexistence to environmental perturbations.

Response: Thanks for your feedback! We're glad to see that our article, in its first version, is understandable and has caught your interest.

Comment 1

The manuscript is well presented and the figures clarify the work, yet it does not reflect well in my opinion work that has done before. For instance, the abstract states that prior work has focused on understand the set of intrinsic growth rates compatible with species coexistence, while in this present work, the author ask the novel question: Given the fact that species coexist, which will be the minimal perturbation that changes this outcome. I am afraid other studies has answered already the same question including at least Allen-Perkins et al. 2023 ELE and Lepori et al. 2024 Proc Roy B. So the abstract needs to be reoriented in my opinion to better reflect the novelty of the work without diminish previous efforts.

Response: Thanks for this critic. We are very committed to not overshadowing anyone's work and we believe that the various papers cited (Allen-Perkins et al. 2023 ELE and

Lepori et. al 2024 Proc Roy B.) and our article all have their own specificity, all worthy of interest. In this new manuscript, we have gone back to one of my starting points, which was to emphasize the clarification of the consideration of domain shape and the notion of perturbation. Instead, and following the advice of all the reviewers, we have emphasized the characterization of $p(z)$, which allows a complete analytical understanding and provides a new object of study. And also the deepening of the study carried out at the species level with the defining of biotic role on robustness of coexistence. The abstract has therefore been modified accordingly, as follows:

> **ABSTRACT** Here we examine how biotic interactions determine the robustness of species coexistence in the face of environmental perturbations. For In Lotka-Volterra community models, given a set of biotic interactions, recent approaches characterized, and applied have analysed the probability of finding at set of species intrinsic features (e.g. intrinsic growth rates) a set of species intrinsic growth rates (representing intraspecific demographic features) that will allow coexistence. Here we ask instead: if species do coexist, given their interactions, how fragile this coexistence should be to variations in species demographic parameters ? This change of framing allows us to derive the essential features of interactions that determine the robustness of coexistence, while not reducing it to a single number. Several metrics have been used to quantify the fragility of coexistence in the face of variations in those intrinsic growth rates (representing environmental perturbations), thus probing a notion of 'distance' to the edge of coexistence of the community. Here, for any set of interacting species, we derive an analytical expression for the whole distribution of distances to the edge of their coexistence. Remarkably, this distribution is entirely driven by (at most) two characteristic distances that can be directly computed from the matrix of species interactions. We illustrate on data from experimental plant communities that our results offer new ways to study the contextual role of species in maintaining coexistence, and allow us to quantify the extent to which intraspecific features and biotic

interactions combine favorably (making coexistence more robust than expected), or unfavourably (making coexistence less robust than expected). Because it has both as central tenets, our work helps synthesize coexistence and ecological stability theories. Our work synthesizes different study of coexistence and proposes new, easily calculable metrics to enrich research on community persistence in the face of environmental disturbances.

Comment 2

Likewise, it is introduced that Allen-Perkins et al. (2023) has a promising direction yet not enough to test quantitative predictions regarding the robustness of species coexistence to actual perturbations (line 41). As coauthor of such paper I would like to express a different opinion. Actually, Allen-Perkins et al (2023) propose a metric called "Exclusion ratio" that actually measure distance to the edge of the feasbility domain according only to the interaction matrix that give an assymetric domain. In the exact same way as this work, it is considered that environmental perturbation affects species intrinsic growth rates considering an isometrix (non-directional) perturbation, and therefore, it evaluates what will be the minimal perturbation to reach the edge of the feasiblity domain and start to lose species. Note that such metric called exclusion ratio does not take into account a particular configuration according to an observed intrinsic growth rate, rather it focuses only in the structure of the matrix of interactions. See more detail in Box 1 and Fig. 2 of Allen-Perkins et al. (2023), and the isometric perturbation which is exactly the same approach as Fig. 2 of present work. Moreover, metrics of Allen-Perkins et al. 2023 (Box 1) also consider the difference in distance to the edge between an equilateral triangle and an non-equilateral one (Fig. 2 of present work as well as equations 7 and 8), and finally Allen-Perkins et al. (2023) also provide metrics to evaluate the robustness of species coexistence at both the individual species level and the entire community for an arbitraty number of species in the community. So unless I am missing something both approaches are alike or very simmilar, and a key aspect here is to clarify which part of the methodology represents a significant advance.

Response: The two studies are indeed similar and the first version of our manuscript does not sufficiently reflect the contribution of our work compared to that of the Allen-Perkins et al. paper (2023). Thank you for the clarification of the work you have provided and we do agree with the overall comment. The "exclusion ratio" is a measure of asymmetry that characterizes the shape of the feasibility domain in the same way as D or D^*. As for the measure used at species level, it too characterizes a property similar to what we call "species vulnerability" (SV_i, or SCV_i in the first manuscript version). De facto, the two studies go in the same direction and some conclusions are shared. However, the analytical demonstrations used are not the same and our approach allows a complete characterization of the $p(z)$ function, which perhaps paves the way for other studies. In this sense, your comment 4 is very interesting. Therefore, we have modified the introduction to take better account of our previous work, and to highlight the deeper insights we're adding :

LINE 47 Recently Allen-Perkins et al. (2023) showed that a notion of 'domain asymmetry' can be correlated with variations of population dynamics across species in experimental plant-communities. These results are promising, but do not yet test quantitative predictions regarding the robustness of species coexistence to actual perturbations. They show that theoretical rankings of species vulnerabilities, based on the shape of the feasibility domain, are consistent with observed variations in population dynamics. Here, to make feasibility theory more directly interpretable, we will include ecological perturbations in its formulation. The goal would then be to move beyond a purely geometric description of the feasibility domain, and explicitly characterize the robustness of coexistence to environmental perturbations. In line with these recent approaches, the aim of our study is to expand on the study of feasibility by proposing an explicit mathematical relationship between the robustness of coexistence in the face of environmental disturbances, and the shape and size of a feasibility domain.

LINE 71 As we hinted above, our the description of the distribution of distances to the edge of coexistence, is in line with recent work by Allen-Perkins et al. (2023). Using a similar logic to study the asymmetry of the feasibility domain (but different analytical calculations) these authors introduced different metrics related to the robustness of coexistence of the community. Remarkably, they used one of these metrics, the so-called "probability of exclusion", to characterize species vulnerability in grasslands, showing that theoretical predictions based on the shape of the feasibility domain are consistent with observed population dynamics. In a similar vein we show here how to use features of the function $p(z)$ to study the relative vulnerability of species. The idea is to address the biotic role played by each species in the robustness of coexistence, in the context of the community to which it belongs.

LINE 80 We apply our methods to data simulated ecological communities, either drawing parameters at random (See appendix B) or inferring them from experimental plant community experiments analyzed by Barbier et al. (2021), who computed the carrying capacities and interaction forces of species. The results (in line with Allen-Perkins et al. (2023)) confirm the link between the coexistence measures we derive from our work and the actual persistence of species through time in a changing environment.

Some changes have necessarily been made in the discussion to better fit the new orientation of the article. First at:

LINE 421 Unlike previous measures As in previous studies of asymmetry of the feasibility domain (See Appendix S7 of Grilli et al. (2017) and Saavedra et al. (2017) and later Tabi et al. (2020) and Allen-Perkins et al. (2023)), we included the notion of disturbance in the mathematical definition of z (see equations 5 and 6). It is this step that subsequently allows us to dispense with a purely geometric analysis of the feasibility domain, and instead use standard objects of the Lotka-Volterra model (the matrix A and its inverse, which commonly occur in stability analysis). However, our theory strongly depends on the way environmental disturbances are modeled. This highlights the importance of taking into account the type of disturbance when studying the stability of a community and suggests that different results could be obtained by considering other types of disturbance (ie. that vary through time, and/or scale with species standing biomass). Deepening our theory to account for more general types of disturbance could be an interesting direction.

And secondly, by removing the following paragraph:

LINE 433

It should be made clear that we did not directly test the relevance of our function $p \mapsto p(z)$ to predict actual species persistence under real environmental perturbations. To do so, we would need to compare these metrics on experimental datasets highlighting the persistence of species over time. This is precisely what Allen-Perkins et al. (2023) have done, and their "Asymmetry index" bares many similarities with our analysis (use of the incenter position, species-level contributions to coexistence). Their results show a good match between the species' actual persistence over time and the predictions made based on their indexes. Also, their recent measurement seems far less sensitive than others with small feasibility domains and large community sizes, and is therefore far more practical to use. We can see our analysis as an extension of theirs, and the results of their data analysis are a good indication that this extension relies on solid foundations, both theoretically and empirically, to study species coexistence and persistence.

Comment 3

In my opinion, a key novel aspect compared to prior work is to classify species according to their competitive roles. I do like a lot the section about “Contextual species contributions to the robustness of coexistence” which provide a tractable analysis to see species through their competitive behaviour, therefore it can be seen which species play a central role for the coexistence of communities and which species can be considered as “disruptors”. This is really novel and interesting because can reduce the complexity we often observe in horizontal network to point out the importance of particular species. I can envision the importance of this work for applied fields such as restoration ecology, ecosystem functioning or species conservation.

Response: Indeed, this result merits greater emphasis. To this end, in addition to the changes made in the abstract and introduction previously described, we have added the idea of particular species in the discussion section :

> **LINE 399** It is interesting to note that the species present in the dataset used in the study seem to retain relatively the same role regardless of community composition. It would be interesting to extend this analysis to larger datasets to study the consistency of species roles in maintaining robust coexistence. If we consider the contribution to the community-scale robustness of coexistence as a function rendered by a species within the community, it is likely that certain species correspond to "key species" [1], [2].

[1] M. E. Power, D. Tilman, J. A. Estes, et al., “Challenges in the Quest for Keystones: Identifying keystone species is difficult—but essential to understanding how loss of species will affect ecosystems,” BioScience, vol. 46, no. 8, pp. 609–620, Sep. 1996. DOI: `10.2307/1312990`. eprint: `https://academic.oup.com/bioscience/article-pdf/46/8/609/650270/46-8-609.pdf`.

[2] R. Whittaker and H. E. W. Cottee-Jones, “The keystone species concept: A critical appraisal,” Frontiers of Biogeography, vol. 4, pp. 117–127, Sep. 2012. DOI: `10.21425/F54312533`.

Significant changes have been made in the rest of the manuscript to match the new orientation of the article. Overall, I think that condensing the technical part of the article to present more directly how the function $p(z)$ is obtained and how it is applied to the data highlights theses results (See Comment 6 and the new mathematical appendix).

On the advice of another reviewer (Comment 7), we've also made the description of the use of SV_i and r_i^*/D_* clearer (in the "Contextual species vulnerability" section, previously named "Contextual species contributions to the robustness of coexistence" section).:

> **LINE 289** We can combine the species-level measures r_i^* and SV_i by viewing them as the species coordinates on a two dimensional map, in other words, plotting them against each other. Intuitively, the two should be strongly correlated: species that perceive a hostile biotic environment should also be the most vulnerable, and vice versa. But this need not always be so simple. If a species is itself very hostile towards an otherwise relatively favourable community, the coexistence of all species would require this hostile species to be vulnerable, as coexistence would only be possible in abiotic environments unfavourable to it. By contrast, a species to which the community is relatively hostile could nonetheless be relatively robust if its persistence at high enough densities is required for the persistence of other species. Here those various qualitative roles clearly depend on the biotic context. The same species could change roles depending of which community it is part of. This should lead to the definition of "two" particular roles: on the one hand, vulnerable and repressed, and on the other robust and facilitated. The results obtained by applying our measures to empirical data (Fig. 4) show that it doesn't always have to be this simple, and that it is possible to define two other non-trivial qualitative "roles".

Comment 4

Another point which I think is very interesting but in my opinion not well explained is the edge effects in high dimensions. That is the jump from larger distance to the edge to small distances. The authors put much importance in this result, but it is not well hilighted for the general reader, so it is hard to grasp which is the main finding. My take on this, and it is just a suggestion, is that the cumulative shape shows which will be the optimum number of species for random matrices at which there is a balance between gaining and losing species. At very high number of species as the author state a very small perturbation will make the community to loss species, while there will be an optimal point (by calculating the tagent) where species can maximize both diversity and the robustness against environmental perturbations.

Response: This is a very interesting comment. We agree with you on the possibility of studying the role of species richness in more detail via the behaviour of the $p(z)$ function. So we've added these thoughts to the discussion :

LINE 385 The ratio S/D can be used to understand how many species can be grouped together while maintaining a high percentage of robust states. More precisely, if we want to guarantee that a proportion p of coexistence states is robust to perturbations of intensity ϵ, then maximizing diversity amounts to solving

$$\max\{S \mid \frac{S}{D} \leq \sqrt{\pi/2}\frac{p}{\epsilon}\}$$

whose solution will take the form of $S = \sqrt{\pi/2}\frac{p}{\epsilon} \times D$, so proportional to D.

We believe that a more detailed study of the relationship between $p(z)$ and S would require further work.

Comment 5

Finally, the discussion section presents well the results, and provide nice discussion of the main findings and how this relates to prior work. I like the section when it is discussed that the environental perturbation was modelled as a perturbation that occur in all possible directions, and future work should focus on other types of environmental perturbation, including directional perturbation of changes of intrisinc growth rates in a given direction. A finally surprising fact is that many of the things that I have said before are acknowledged in the discussion section (lines 331-337), which seems odd to have this agreement at the end or not at the beginning. It seems that both works at some point where done simulateously without knowing from each other. I think there is a lot of room for more work on this geometric analysis of the feasibility domain, and the authors of this work are excellent researchers and they have ample material to highlight the novelty of their work. In sum, the classification of individual species according to their competitive roles, the implications fo these role for specific applied subjects, and the balance between diversity and tolerance to perturbations are key examples.

Response: I'm glad you're enjoying the discussion section. Having now changed the way in which this paper positions itself in relation to other recent research (through the modification we describe above in the Comment 2), we hope this won't be a problem anymore. Note that the line 331-337 of the previous manuscript you mentioned have been removed and partially included at the beginning of the manuscript.

Comment 6

With that said, it is very exicing to read and see that geometric analyses of species interactions are gaining momentum to better understand the mechanisms of species coexistence and make predictions of future dynamics in the community, that different labs reach same conclusions independently and that there is room for more significant improvements.

Response: We are particularly pleased to see that our work on this notion is of interest to a number of researchers in different laboratories. Indeed, as you said earlier, Allen-Perkins' work and ours were probably carried out in parallel, at the same time. Although this imposes an extra rigour in taking into account the respective work so as not to overshadow anyone, we're delighted to see that our approaches and results converge.

Authors’ Response to Reviewer 2

General Comments. Based on previous theoretical and empirical studies on feasibility in ecology, Desallais et. al made an insightful step forward that tries to unravel the relationship between the robustness of a community and the geometrical properties of its feasibility domain. Specifically, they ask the question: how should we interpret perturbation in the scope of feasibility domains, what is the distribution of extinction-inducing perturbation intensity, and what is the relevant geometry properties governing such distribution?

Combing intuition in planar regular triangles and the analysis from previous closed-form results in multi-species feasibility domains, Desallais et. al were able to argue that for the distribution p(z), where z is the extinction-inducing perturbation intensity, two geometrical distances D* and D emerge as the relevant controlling parameter across all shapes and dimensions of feasibility domains. D*, the minimal extinction-inducing perturbation intensity, defines the characteristic scale of p(z). While D, another effective distance in unit of length, controls how fast p(z) grows from z=0 (boundary of feasibility domain).

Response: Thank you for your comment! We’re glad that, despite the criticisms made of the first version of the manuscript, you were able to fully understand the article as we intended. We’d like to thank you for all your criticisms, which particularly helped to clarify the manuscript.

Comment 1

The authors have stated their methodology carefully and overarching framework clearly. However, it is worth pointing out one major concern about potential lack of assumptions, since the feasibility domain is not equivalent to the domain where all species could coexist/persist. That is, it requires certain assumptions of stability to make them equivalent.

For some interaction matrix A, sampling some r inside the feasibility domain with some initial abundance N0, the corresponding system (r, A, N0) will have a feasible fixed point N, but such N may not be (locally) stable, and thus the loss of coexistence can occur in this system. Similarly, different initial abundances N0's may start the system in different basins of attraction, leading to different extinction outcomes. Chances are that the boarder of feasibility domains might not be the actual boarder of "coexistence domain". In this line, I haven't seen enough details that could validate the basics of how the authors would define robustness, also not in the empirical data analysis part. One possible workaround is to impose global stability for interaction matrices, cf. Deng et. al PloS Compt. Biol. 2022. Meanwhile, assumptions on stability may have already restricted the robustness properties, or exclude the system inferred in empirical analysis, which the authors may need to leverage.

Response: You're right about the necessary conditions to make the feasibility domain equivalent to the domain where species actually coexist, in locally stable equilibria. To achieve this, we impose global stability by the D-stability of the A interaction matrices within our models, both for mathematical analyses and for the simulations carried out. To clarify this mathematical choice, we have taken care to specify it in the "The feasibility domain" section :

LINE 104 A growth rate vector $r = (r_i)$ is 'feasible' if the fixed point $N^*(r) = A^{-1}r$ of the above model is strictly positive, meaning that $N^*(r)_i > 0$ for all i. To guarantee the coexistence when the feasible equilibrium point is reached, we impose global stability of the system [1] by considering only D-stable interaction matrices.

[1] J. Deng, W. Taylor, and S. Saavedra, "Understanding the impact of third-party species on pairwise coexistence," PLOS Computational Biology, vol. 18, no. 10, pp. 1–21, Oct. 2022. DOI: `10.1371/journal.pcbi.1010630`.

But also at the beginning of the new mathematical appendix presenting all our analytical demonstrations (See "Data, script, code, and supplementary information availability" section to find it).

Comment 2

The authors stated their research significance as "...move beyond a purely geometric description of the feasibility domain, and explicitly characterize the robustness..." in the introduction (line 45-46), and "unlike previous measures of asymmetry of the feasibility domain..., we included the notion of disturbance in the mathematical definition of z..." in the discussion (line 318). While I could totally see important contribution to this subject in this work, personally I would not perceive its significance this way. For previous work and this work are built on the similar feasibility-domain-geometry ground, and specifically, perturbation (parameter change) being mearsured geometrically as exceeding the feasibility domain seems to be a widely accepted conception in these works, as the authors already cited throughout their manuscript.

Response: We have decided to rework part of the abstract, introduction and discussion section in order to better highlight the innovative elements of our study (the clarification of the $p(z)$ function and the deepening of the study at species level). The main aim of these modifications was to highlight our work without clashing with recent papers that have already incorporated the notion of disturbance into their demonstrations. See response to Comment 1 and Comment 2 of the reviewer 1 that contains the main changes made to the manuscript in this sense.

Comment 3

However, I do think this paper made a substantial step forward, in terms of providing an analytical approximation for the distribution of z, i.e, p(z). Such a distribution law, as mentioned by the authors, enables comparison across different interaction matrices and is thus universal knowledge for all systems. Moreover, it reveals the relevant geometrical parameters shaping this distribution. And to this end, we could be able to evaluate robustness across different systems in a consistent way, for example, a smaller z could be actually more robust than a larger z, solely because the first one lies in an "elongated" feasibility domain and the second one lies in a equilateral triangle feasibility domain; another example is the Figure 5 where one could immediately see how different communities operated at different status in terms of robustness.

Response: The characterization of $p(z)$ is indeed one of the results we'd like to highlight as one of the key-points, and we thank you for your positive criticism of the gain in understanding our work enables. As such, we wanted to put this result more forward. We have therefore modified the abstract and introduction for this purpose (See response to Comment 1 of the editor, and still Comment 1 and Comment 2 of the reviewer 1.).

But more generally, we have chosen to condense the results presented in the article in order to bring out certain essential equations more clearly (such as the expression of $p(z)$), which should necessarily give it importance. To do this, we have completely reworked the "Distribution of distances to the edge of feasibility" section. In the first version of the manuscript, it consisted of a number of subsections detailing how to arrive at $p(z)$. We have moved this to a mathematical appendix and kept only a single condensed section. In that sense, we also changed the title to be more precise on what this study is about.

Comment 4

In line 94, why does the cumulative mass function of the multivariate normal distribution $\Phi A^T A(0)$ have an argument 0? I find it a bit hard to follow the notation here.

Response: $P(D_f)$ corresponds to the probability of obtaining a feasible equilibrium, i.e. of positive abundance (cf. line 89 in the first manuscript). This "0" argument therefore corresponds to $N^* > 0$. To avoid these problems of understanding the article due to highly technical or unclear mathematical notation, we both have rework this section and provided a new and more detailed mathematical appendix :

LINE 120 $\mathbf{P}(D_f)$ is thus the cumulative distribution of a multivariate normal distribution noted $\Phi_{A^\top A}(0)$, centered on 0 and with covariance matrix $C = (A^\top A)^{-1}$. The probability $P(r \in D_f)$ can therefore be computed as the cumulative distribution, evaluated at 0, of a normal distribution whose covariance matrix is determined by the interaction matrix A (this covariance matrix is $(A^\top A)^{-1}$). In the absence of interactions $P(r \in D_f) = 2^{-S}$. To focus on the effect of interactions it is thus convenient to define a ratio of probabilities :

$$\Omega\underline{(A^\top A)} = 2^S \cdot \underline{\Phi_{A^\top A}(0)}\mathbf{P}(r \in D_f) \tag{1}$$

Comment 5

In line 101, the logic seems to be subtle if asking "minimum disturbance...without leaving the domain": which should be zero. I guess the authors are asking about maximum torlerable disturbance? Or the "minimum" is with respect to choices of borders of the feasibility domain? Please clarify.

Response: Thanks for this remark. It's important to be clear on this passage to avoid any confusion on this key point. Line 101, it's indeed a typo on our part and it should be "the maximum". As a result of the reworking of the various technical sections of the article, this passage has been removed from the new manuscript. But a similar passage is now in the new manuscript :

LINE 183 For any point r in a feasibility domain (so a feasible growth rate vector), we can measure its distance from the edge of the domain as the minimal perturbation intensity capable of leading at least one species to extinction.

Overall, it's a question of finding the smallest disturbance that will at least extinguish a species, given a given feasibility domain. Formulated in this way, this corresponds to the "maximum tolerable disturbance" and not to the smallest tolerable disturbance, which would effectively and trivially be zero or infinitesimal.

Comment 6

In line 133, what are the implications when considering "long-time scale"? It would be nice if the authors could further explain here. In line 146, what does the maximal sensitivieis $v = (||v^i||)$ mean here? Again, a bit explanation of the notation might help here. (is the parentheses constructing a vector here?)

Response: We consider a long time scale to assume that the ecological dynamics of the GLV model are at equilibrium. This is a necessary step in the study of the feasibility domain, notably to consider the equation $N_i^* = V \cdot r_i$ which calculates equilibrium abundances. To clarify this passage, we have added an indication:

LINE 179

Following Allen-Perkins et al. (2023), Cenci, Montero-Castaño, et al. (2018) and De Laender et al. (2023), we consider perturbations as changes in environmental conditions that occurs on a long-time scale (so that new equilibrium are reached)

Concerning maximum sensitivity, the parentheses are indeed constructing a vector here (note it was the same for $\delta r = (\delta r_i)$ in the first manuscript). As a result of the simplification of the technical and mathematical elements of the article, the passage including this notation has been removed from the new manuscript. Instead, we explicitly tell that v and w (added to the new version) are vectors and we describe what $||\cdot||$ means (euclidian norm). These notations remain in the new mathematical appendix but are better described.

Comment 7

In line 167, the derivation that is "simple to carray out" is actually nontrivial for me. I guess this is partially related to point 1 on the argument of the Φ function. I wonder if the authors could provide the details perhaps in appendix or other places? Meanwhile, perhaps it's also worth clarifying how is Ω/i defined. Does $A^T A/i$ means the product of A^T/i and A/i, which is A removed the i'th column or row, and has dimension of $S-1$, or is i'th column or row of A replaced with zeros, but still dimension of S?

Response: To answer the precise question, $A^T A/i$ means the product of A^T and A, from which is removed the i'th column and row and has dimension of $S-1$. In the new manuscript, we explain this :

LINE 229 with $X_{/i}$ notation meaning for any matrix X, the corresponding matrix without the $i-th$ row and column, and $\Omega_{/i}$ is essentially the relative volume of the feasibility domain for the community without species i (but see the appendix for precise expression and derivation)

This is part of a more general request from reviewers for more mathematical simplicity within the article and for a detailed appendix presenting all the calculations required for the article. We particularly agree with these criticisms. We have therefore made a number of changes within the body of the text, with a whole new "Distribution of distance to the edge of feasibility" that just present the $p(z)$ function and its core elements. (See Comment 1 of the editor).

In return, we have written a purely mathematical appendix presenting all the calculations made. We hope that this will make the text easier to read for a wide audience of ecologists, while still being comprehensible in detail and precision for anyone wishing to delve deeper into the mathematical aspects of our work.

Comment 8

In line 199-215, I think in general the SCVi measurement needs more justification. I could be comfortable calling these individual i's terms (summed up to D Equation 10) as the species i's contribution to the community robustness. However, what is the justification for interpreting SCVi as the individual vulnerability? The authors said "Indeed, they determine the distribution of coexistence states for which that each species is within a certain pertubation distance from extinction". But I found this statement to be equivalent to "they determin the p(z) function", which is the property of community, not of one species.

Response: Thanks for your comment! As this is one of the key measures in our article, it is important that it is particularly clear. The SV_i elements are indeed components of D which is a community property of robustness of coexistence. But note the "1/" at the beginning of the equation defining D (equation 10 in the previous manuscript and equation 9 in the new version). The higher these these individual i's terms, the smaller D and the higher the initial slope of the distribution function $p(z)$ (See Fig. 3 of the new manuscript). That's why we interpret them as individual vulnerability. To be clearer, we've changed the manuscript as follow :

> **LINE 278** We now turn our attention to D, which describes the shape of feasibility domain near its edges. It consists of On the other hand, the distance D describes the edges of the feasibility domain. It reads as the inverse of an average of S elements, one for each species: i. Those elements relate to the individual vulnerability of each species.
>
> $$\mathrm{SV}_i = w_i \sqrt{\frac{|A^\top A|}{|(A^\top A)_{/i}|}\frac{\Omega_{/i}}{\Omega}} \tag{2}$$
>
> Indeed, they Those terms determine the distribution of coexistence states for which that species is within a certain perturbation distance from extinction. Hence, they relate to the individual vulnerability of each species.

Authors' Response to Reviewer 3

General Comments. This manuscript studies the robustness of the equilibria of a GLV model to perturbations on the growth rate of species. That is, given a feasible (but not necessarily stable) equilibria where all the species abundances are positive, they ask how big can a perturbation on the growth rates be until the first species abundance becomes negative. They embed the perturbation z in a function p(z) that is able to account for the geometry of the feasibility domain. In this way they are able to incorporate in their analysis both the sheer volume of such space, and its shape; an equally important factor in determining effective feasibility regions. The way they incorporate this shape information is by chipping away from the original feasibility region any volume that lies outside the the largest inscribed hypersphere with center given by the growth rate vector. That is, essentially the geometry is erased, collapsed to a hypersphere. Armed with this framework, they derive approximations to the distribution of hypersphere radii for general feasibility domains, and validate them by computing these radii for random parametrizations. Additionally, they also consider this analysis at the species level, computing a quantity called vulnerability to perturbation given coexistence, and another quantity called "how hostile the community is to species i". They then look at this quantities using parameter values derived in Barbier et al (2021) for grassland experiment data. This allows them to map experimental points onto their theory, providing an example of what type of insights this theory could give. The results make sense. A sensible, intuitive correlation is found: overall, species that perceive a hostile biotic environment are also the most vulnerable, pointing at the fact that this whole framework makes sense. The manuscript seems to be technically sound, even though I was not able to check the math because I couldn't find the supplementary material. The results promise to be potentially useful. However, I have two major concerns that are preventing me from recommend this manuscript for publication.

Response: Thank you for this detailed general comment. We're delighted to see that the core ideas of our article, in its first version, have been fully understood. This makes the reviews addressed to the article very audible and we hope we have dealt with them in the right way.

Comment 1

Most importantly, there is no real experimental validation. This is a problem because the biggest promise of the introduction (lines 41-46) is precisely to characterise this robustness to perturbations: "These results are promising, but do not yet test quantitative predictions regarding the robustness of species coexistence to actual perturbations. [...

. Here to make feasibility theory more directly interpretable, we will include ecological perturbations in its formulation. The goal would then be to move beyond a purely geometric description of the feasibility domain, and explicitly characterise the robustness of coexistence to environmental perturbations." The experimental part is merely a proof of concept which conclusion is: this framework is sound, and could be used to predict robustness of communities and species to environmental perturbations. As such, not using it to predict such robustness to perturbations in real systems (or simulations with biologically informed parameters) leaves the manuscript incomplete. I believe an experiment of the type: Calculate p(z) for some parameterisation, then run simulations where one gradually perturbs r and checks if such p(z) can predict when a species will go extinct. I am aware that real experimental validation is hard, because it would first require assuming that GLV models are sufficient to explain ecological dynamics, which is not always true. But checking it in simulations, with realistically informed GLV parameters is a feasible task very natural to do next.

Response: Thank you for this comment, which we can only agree with. It would indeed have been very welcome to have suitable data to test the effectiveness of our predictors. As we didn't have them, we followed your advice to use a generalized L-V model to test the effectiveness of D_* and SV_i in measuring the persistence of communities and/or species over time in a system subjected to disturbance. The whole approach is presented in a new appendix section, "Application to persistence of species in simulated ecological systems", available online. The results are shown in a new figure presented in this appendix (Supplementary figure B1).

We have chosen to include this verification in an appendix rather than at the heart of the article because (1) it is not as novel as the formulation of $p(z)$ or the characterisation of biotic roles in coexistence, the two results we wish to highlight and (2) because we do not consider such tests using L-V system with random parameters and perturbations as a strong proof. So as not to lengthen the already long article, we have chosen to place this alongside the other appendices. Still, these new elements are mentioned in the introduction and the article:

> **LINE 80** We apply our methods to simulated ecological communities, either drawing parameters at random (See appendix B) or inferring them from experimental plant community experiments. The results (in line with Allen-Perkins et al. (2023)) confirm the link between the coexistence measures we derive from our work and the actual persistence of species through time in a changing environment.

> **LINE 254** This approximate relationship, which must be understood as an equivalence of orders of magnitudes and not of precise values, taken together with equation 7 connects the size of the domain to its shape, and to the probability that a given perturbation can push species to extinction. In the appendices we expand on the latter point by considering randomly changing environments trough time. We show in simulations that the equivalent metrics of equation 10 do predict the duration of stable coexistence periods (See supplementary figure B1 in the "Persistence of species in simulated ecological system" appendix).

Comment 2

The parameters used to illustrate the theory are not derived in the paper, but taken from another paper; Barbier et al (2021). This means that in order for me to support the results presented here, I need to also endorse the results of the previous paper, which I haven't read/reviewed. Science should be reproducible, so I would be more comfortable if the way those parameters were obtained in the previous paper was reproduced and explained in detail here.

Response: We agree with you on the importance of research reproducibility. However, it is difficult for us to re-explain the complete methodology of Barbier et al (2021) which led to the quantification of pairwise interaction forces and carrying capacities on the basis of their experiments, as this is not part of our research work. This paper has been peer-reviewed and published by a trusted publisher and their full methodology is available by reading their paper and supplementary materials.

However, for the sake of transparency, all the data sent to us by the researchers and taken from their article is made available online and clearly indicated in the paper:

> **LINE 316** In this parametrization, A_{ij} has no dimensions and satisfies $A_{ii} \equiv 1$. On the basis of pairwise interaction values, we then reconstruct interaction matrices consisting of 4 species, which have been experimentally realized. All the pairwise interactions values and carrying capacity values derived from their work and used in this study are availiable on Zotero (See Data, script, code, and supplementary information availability section below).

Comment 3

Line 9. But you are reducing it to a single number: p(z), right?

Response: $p(z)$ is the function linking the probability of loss of coexistence knowing a disturbance of intensity z and this intensity z. In this sense, we believe it is much more than a simple number. It's true that the study of this function allows us to obtain various metrics such as the characteristic distances D and D_*, or SV_i at the species level. These metrics are highlighted in the article for their practical uses, notably for experimental and applied purposes. However, the study of the $p(z)$ function provides a complete understanding of the relationship between intensity of perturbation, size and shape of the feasibility domain and robustness of coexistence, and allows us to go further than simple number. Finally, we believe (see response to Comment 4 of reviewer 1) that the study of $p(z)$ would allow us to go even further in the study of the feasibility and persistence of species.

Comment 4

Lines 12-13. What is the "expected" robustness? Is there a null model?

Response: As $p(z)$ is a distribution function, null model would be $p(z) = 0.5$, meaning that for a certain z value, there are as many more robust communities as less robust communities enabled by this set of biotic interactions. If a community lies close to $p(z_r) = 1$ on the $p(z)$ curve (Fig. 5) defined by its interaction network, this means that this realized community is much more robust than a majority of other possible communities with regard to this interaction network. This is why we speak of more robust communities "than expected".

Comment 5

Line 15. Stability is one of the keywords of the paper. However, it only appears in the abstract and in the discussion. There is no stability analysis in the methods and the results. If we consider that in GLV models feasibility does not imply stability, I think haveing stability be part of the keywords of the paper does not faithfully reflect its content, nor does having the sentence "our work helps synthesise coexistence and ecological stability theories" in the abstract.

Response: Thank you for your comment. It is true that a large proportion of studies in theoretical ecology dealing with stability refer to stability of linear systems (stable equilibrium point). In this sense, there is effectively no analysis of the stability of systems and it therefore seems wiser to remove this term from the keywords. We did so in the new manuscript.

To explain the initial approach: stability can be understood as a more general facet of the study of the response of systems to disturbance. It is in this sense that the phrase "our work helps synthesise coexistence and ecological stability theories" was written. To avoid any confusion, we have removed most of the references to "stability" in the article, except those specifically concerning the D-stability of interaction matrices (newly added at the request of reviewer 2, see its Comment 1) and concerning certain parts of the discussion section where stability precisely refer to the broader definition we've mention just above.

Comment 6

Line 119: Confusing notation – A is both area of triangle and matrix of interactions.

Response: Thanks for the correction! Indeed, the two similar notations are confusing. To remedy this, we've decided to designate the area of a triangle by $\mathcal{A}$, with a different font, to keep A for the interaction matrix. This new notation has been added everywhere in the "Distribution of distances from the edge in standard triangles" section of the new manuscript.

Comment 7

Figure 3: Left panel: p(z) is calculated for pairs of A, and r, but I think it would be more useful to also see simulations where for a fixed interaction matrix A, we see the distribution for many samplings of r. Left panel: How does this distribution changes with the number of species S is not shown, and it would be good to see. Right panel: confounding factor. How much of the gain in accuracy is due to the fact that p(z) decreases exponentially with S?

Response: In an effort to simplify certain technical and mathematical details, certain equations have been removed from the article in favour of a complete and detailed mathematical appendix (availiable online alongside the others). As a result, we felt that the panel on the right of Figure 3 was less relevant and has been removed. The remaining figure (corresponding to the left panel of the old manuscript) was generated by selecting a given interaction matrix A and sampling z values in order to obtain the $p(z)$ distribution empirically. The aim is precisely to illustrate by example the relevance of the $p(z)$ analytical approximation. It fits with the description you propose! To be clearer, we have also added the initial slope of $p(z)$ on the figure and a scheme that explain graphically D_* and r^*. In doing so, we want the Fig. 3 to be seen as the direct continuation of the Fig. 2.

Comment 8

Lines 209-213: Extremely confusing description. I suggest unpacking the non-trivial outcomes of the analysis further, and even add examples to help the reader follow.

Response: Thanks for this remark. As this part has been given even greater emphasis in the new version of the manuscript (with the overall simplification), it is all the more important that its results are clear and intelligible. We have taken your advice and kept the analysis of non-trivial cases for the following section and the description of the Fig. 4. The related paragraph now reads as follows:

LINE 289

To understand the role of each species, we can correlate redtheir relative SCV_i values with how hostile the community is to the species, as measured by relative r_i^c/D_* values. We can combine the species-level measures r_i^* and SV_i by viewing them as the species coordinates on a two dimensional map, in other words, plotting them against each other (see Fig. 4). Intuitively, the two should be strongly correlated: species that perceive a hostile biotic environment should also be the most vulnerable, and vice versa. But this need not always be so simple. If a species is itself very hostile towards an otherwise relatively favourable community, the coexistence of all species would require this hostile species to be vulnerable, as coexistence would only be possible in abiotic environments unfavourable to it. By contrast, a species to which the community is relatively hostile could nonetheless be relatively robust if its persistence at high enough densities is required for the persistence of other species. Here those various qualitative roles clearly depend on the biotic context. The same species could change roles depending of which community it is part of. This should lead to the definition of "two" particular roles: on the one hand, vulnerable and repressed, and on the other robust and facilitated. The results obtained by applying our measures to empirical data (Fig. 4) show that it doesn't always have to be this simple, and that it is possible to define two other non-trivial qualitative "roles".

The non-trivial cases are more detailed in the "Application to data from a grassland experiment" section where examples are precisely given through the result presented Fig. 4.

Comment 9

Lines 329-338: Need to be more clear how the results here extend the work of Allen-Perkins et al. As of now it doesn't seem like the current manuscript adds much. Again, I think testing the prediction power of this theory, at least in realistically informed GLV models, would highly increase the strength of the manuscript.

Response: To better distancing ourselves from the work of Allen-Perkins et al. 2023 (and other recent work on feasibility), we have reworked the way we introduce our research and how it fits into this existing recent literature (See specifically our response to Comment 1 and Comment 2 of reviewer 1 for more details).

As well as the new abstract and introduction, the reorganisation of the different sections of the article was also intended to put more emphasis on two of the innovative interests of our paper: the complete characterization of $p(z)$ and the deepening of the study at the species level with their biotic role in coexistence due to interactions. Consequently, lines 329-338 of the previous manuscript have been removed.

Comment 10

Does the title clearly reflect the content of the article? [

Yes,[X

No (please explain), [ ] I don't know The title is vague because it does not state/hint to the main result of the paper.

Response: To better reflect the results presented in the article, we have changed the title to "The distribution of distances to the edge of species coexistence".

Comment 11

Are the methods and analyses sufficiently detailed to allow replication by other researchers? [ ] Yes, [X] No (please explain), [ ]
I don't know I could not find the supplementary material with the links provided, so there are calculations and proofs that cannot be checked, or I have a hard time following (eq between lines 119 and 120, eqs 7, 8, 9, eqs between lines 168 and 169, eqs 10 and 11). The part of the paper where data is used makes use of parameters that are inferred in a different paper. This makes the whole data section hinge on results that are not reproducible with the information given in the paper.

Response: With this second version of the manuscript, we've included a comprehensive mathematical appendix which we hope will clarify any points that are difficult to follow. It is available online alongside the other appendices : https://doi.org/10.5281/zenodo.12744286
Concerning the data used, see response to Comment 2.

Comment 12

Are the methods and statistical analyses appropriate and well described? [] Yes, [X] No (please explain), [ ] I don't know See above answer. Nothe that this might be my problem, since I couldn't find the Supplementary Material

Response: The technical and mathematical part of the first manuscript was particularly detailed (but not detailed enough), so we didn't provide a mathematical appendix at first. Since we have decided to greatly simplify the manuscript, we have included an even more detailed mathematical appendix (See previous comment).

Comment 13

Are the conclusions adequately supported by the results (without overstating the implications of the findings)? [] Yes, [X] No (please explain), [ ] I don't know Refer to my first major concern.

Response: We understand. As we've already answered your first comment, we've taken this into account and applied our metrics to simulated communities (gLV model), as you suggested. We hope this will give our work a little more credibility.

The distribution of distances to the edge of species coexistence

Mario Desallais[1], Michel Loreau[1,2], and Jean-François Arnoldi[1]

https://doi.org/10.1101/2024.01.21.575550

Abstract

In Lotka-Volterra community models, given a set of biotic interactions, recent approaches have analysed the probability of finding a set of species intrinsic growth rates (representing intraspecific demographic features) that will allow coexistence. Several metrics have been used to quantify the fragility of coexistence in the face of variations in those intrinsic growth rates (representing environmental perturbations), thus probing a notion of 'distance' to the edge of coexistence of the community. Here, for any set of interacting species, we derive an analytical expression for the whole distribution of distances to the edge of their coexistence. Remarkably, this distribution is entirely driven by (at most) two characteristic distances that can be directly computed from the matrix of species interactions. We illustrate on data from experimental plant communities that our results offer new ways to study the contextual role of species in maintaining coexistence, and allow us to quantify the extent to which intraspecific features and biotic interactions combine favorably (making coexistence more robust than expected), or unfavourably (making coexistence less robust than expected). Our work synthesizes different study of coexistence and proposes new, easily calculable metrics to enrich research on community persistence in the face of environmental disturbances.

Keywords: Generalized Lotka-Volterra models; Community ecology; Perturbations; Feasibility domain; Persistence

[1]Station d'écologie théorique et expérimentale (SETE), Centre National de la Recherche Scientifique – Moulis, France, [2]Institute of Ecology, College of Urban and Environmental Sciences, Peking University – Beijing, China

Correspondence
mario.desallais@sete.cnrs.fr

Introduction

Understanding why and how species coexist is a central question in community ecology (Armstrong and McGehee, 1976; Chesson, 2000; Hastings, 1980; Hutchinson, 1961). Many studies have focused on what makes coexistence possible, and in particular on the role of the network of interactions between species (Abrams, 1984; Abrams et al., 2003; Brose et al., 2006; Otto et al., 2007; Williams, 2008). In the context of Lotka-Volterra models (the simplest mathematical representations of the population dynamics of interacting species), to quantify the role played by biotic interactions in species coexistence, a recent and growing body of theoretical work proposes to study the volume of a community's so called 'feasibility domain' (Rohr et al., 2014, 2016; Saavedra et al., 2017; Song et al., 2018b). Given the set of biotic interactions between species, this feasibility domain is defined as the range of species intrinsic features (thought to reflect abiotic conditions that do not depend on the presence of the other species considered, such as intrinsic growth rates or carrying capacities) that allow species to coexist (Fig. 1). The idea here is that the larger this domain, the more likely a community is to withstand environmental disturbances while maintaining coexistence (Bartomeus et al., 2021; Song et al., 2018a).

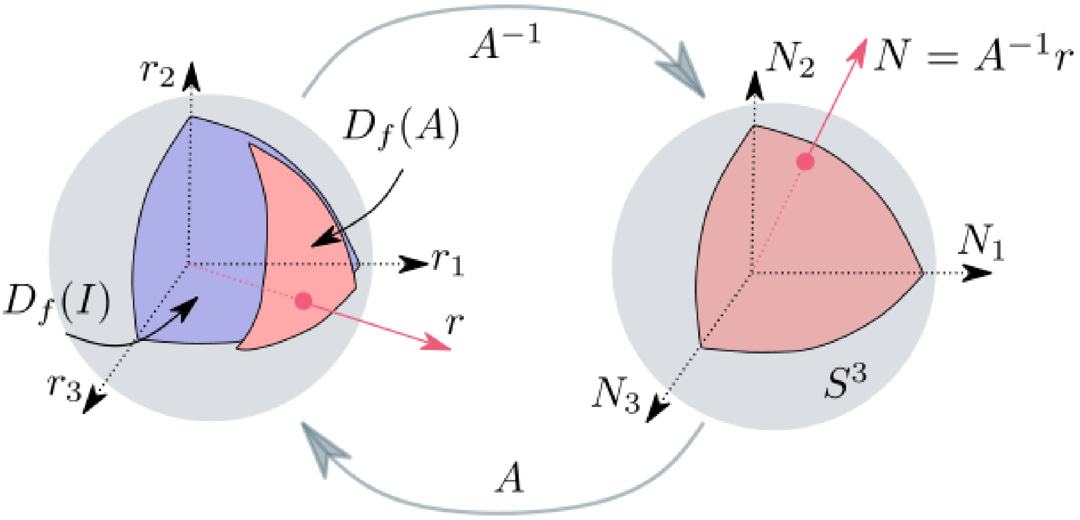
However, the fact that a large set of conditions allows coexistence does not necessarily mean that coexistence is robust to environmental change. A thin elongated feasibility domain could have a large volume, yet only contain fragile coexistence states, vulnerable to small changes in abiotic conditions. This observation reflects the tenuous distinction between two seemingly equivalent questions: "how likely will species coexist?", whose answer, in L-V models, corresponds to the size of the feasibility domain, and "If species do coexist, how fragile will this coexistence be?". This difference between raw and conditional probabilities of coexistence has led to the emergence of shape metrics of feasibility domains (Allen-Perkins et al., 2023; Grilli et al., 2017; Saavedra et al., 2017).

In line with these recent approaches, the aim of our study is to expand on the study of feasibility by proposing an explicit mathematical relationship between the robustness of coexistence in the face of environmental disturbances, and the shape and size of a feasibility domain. To do so, we model ecological perturbations as long term changes of species intrinsic features (such as their growth rates or carrying capacities) and define, for any realized coexistence state, a notion of distance to the edge of coexistence. This distance is the minimal environmental perturbation intensity z able to lead at least one species to extinction. Our goal is to determine, amongst all coexistence states, the proportion $p(z)$ that lie within distance z from the edge of feasibility. For a given feasibility domain, this function $z \mapsto p(z)$ describes the distribution of distances to its edges, thus characterizing both the size and shape of the domain. If the function $p(z)$ rapidly reaches 1 as z grows this means that coexistence is typically fragile. The (cumulative) function $p(z)$ quantifies the interrelation between species growth rates and their interactions. For instance, if in a given state, $p(z)$ is close to 1, this means that in this environment, the set of species intrinsic growth rates and the set of their biotic interactions combine favourably. Our mathematical analysis will reveal the essential features of the function $p(z)$ that can be directly computed from the matrix of biotic interactions.

As we hinted above, our the description of the distribution of distances to the edge of coexistence, is in line with recent work by Allen-Perkins et al. (2023). Using a similar logic to study the asymmetry of the feasibility domain (but different analytical calculations) these authors introduced different metrics related to the robustness of coexistence of the community. Remarkably, they used one of these metrics, the so-called "probability of exclusion", to characterize species vulnerability in grasslands, showing that theoretical predictions based on the shape of the feasibility domain are consistent with observed population dynamics. In a similar vein we show here how to use features of the function $p(z)$ to study the relative vulnerability of species. The idea is to address the biotic role played by each species in the robustness of coexistence, in the context of the community to which it belongs.

We apply our methods to simulated ecological communities, either drawing parameters at random (See appendix B) or inferring them from experimental plant community experiments (Van Ruijven and Berendse, 2009). The results (in line with Allen-Perkins et al. (2023)) confirm the link between the coexistence measures we derive from our work and the actual persistence of species through time in a changing environment. Applied to experimental plant community data, our analysis reveals the role played by the various plant species in maintaining coexistence, which we relate to the degree of facilitation or competition experienced by each. We also quantify the adequacy, in terms of coexistence, between biotic and abiotic conditions in those plant communities. Our work constitutes a proof of concept, demonstrating a theoretical method for future experiments aimed at characterizing a particular type of environment and how well it matches a particular assemblage of species in terms of maintaining coexistence.

Figure 1 – The feasibility domain $D_f(A)$ (in light red on the left) is defined as the subset of growth rate directions that, given a pair-wise interaction matrix A, allows coexistence between all species. It is the intersection of the sphere with the image in r-space (via the matrix A) of the positive quadrant in N-space (shown on the right). The shape and volume of the feasibility domain corresponds to the shape and volume of the light red surface on the left. The probability of feasibility $\mathbb{P}(r \in D_f)$ is the ratio between the volume of D_f and the volume of the unit sphere.



The feasibility domain

Consider a community of S species. Let N_i define the abundance of species i and r_i its intrinsic growth rate (which could be negative if the species cannot establish on its own), which encodes the effect of the environment on the ability of the species to grow if it were alone (Coulson et al., 2017; Levins, 1968; Meszéna et al., 2006; Roughgarden, 1975). The central object of study of feasibility is the matrix $A = (A_{ij})$ of pairwise biotic interactions between all S species in the community. A_{ij} encodes how a change in the abundance of species j, impacts the growth of species i. This can represent competition or facilitation depending on the sign of A_{ij}. The diagonal terms A_{ii} represent intraspecific competition, and will be assumed non-zero in our analysis. The generalized Lotka-Volterra (L-V) model (Volterra, 1926) prescribes the population dynamics of all species as:

$$(1) \qquad \frac{dN_i}{dt} = N_i \cdot \left(r_i - \sum_{j=1}^{S} A_{ij} N_j \right) \quad \text{for } i = 1, \dots, S$$

A growth rate vector $r = (r_i)$ is 'feasible' if the fixed point $N^*(r) = A^{-1}r$ of the above model is strictly positive, meaning that $N^*(r)_i > 0$ for all i. To guarantee the coexistence when the feasible equilibrium point is reached, we impose global stability of the system (Deng et al., 2022) by considering only D-stable interaction matrices (Grilli et al., 2017). To define the feasibility domain one has to assume that variations in growth rates are the result of a variation in abiotic conditions impacting the ability of species to grow on their own, but not their interactions (but see discussion). This abstraction leads to a definition of the feasibility domain associated with the interaction matrix A (Rohr et al., 2014): the set of growth rate vectors $D_f(A)$ such that the equilibrium abundances are non-zero. However, in the L-V model, multiplying all growth rates by a constant does not change coexistence. Thus the feasibility domain has to be defined as a set of *directions*, isomorphic to a solid angle in the r-vector space (Ribando, 2006; Saavedra et al., 2017; Song et al., 2018b), so a convex subset of the sphere (Fig. 1):

$$D_f(A) = \{r/||r|| \mid \text{such that } N^* = A^{-1}r \text{ is strictly positive}\} \tag{2}$$

We can also think of the relative volume of the domain as the probability $\mathbb{P}(r \in D_f)$ of randomly drawing growth rates r which lead to positive abundances (Grilli et al., 2017). The random sampling must be though of as uniform in the space of growth rate *directions*. Importantly, drawing each species' growth rate r_i independently from a standard Gaussian distribution yields such a uniform sampling of growth rate directions. This remark, followed by the linear change of variables $A^{-1} : r \mapsto N$ then leads to the following formula:

$$\mathbb{P}(r \in D_f) = \frac{|A|}{\sqrt{2\pi}^S} \int_0^\infty e^{-\frac{||AN||^2}{2}} d^S N \tag{3}$$

The probability $\mathbb{P}(r \in D_f)$ can therefore be computed as the cumulative distribution, evaluated at 0, of a normal distribution whose covariance matrix is determined by the interaction matrix A (this covariance matrix is $(A^\top A)^{-1}$). In the absence of interactions $\mathbb{P}(r \in D_f) = 2^{-S}$. To focus on the effect of interactions it is thus convenient to define a ratio of probabilities (Saavedra et al., 2017):

$$\Omega = 2^S \cdot \mathbb{P}(r \in D_f) \tag{4}$$

Ω corresponds to the effect of species interactions on the probability of coexistence and is equal to 1 in the non-interaction case.

These are well known results, and since their first introduction to ecology by (Rohr et al., 2014), have been applied to study the coexistence of many ecological systems. Yet the volume of the feasibility domain (also called structural stability) does not, a priori, tell us anything about the shape of the domain, nor how to relate its value to the probability that a given perturbation will push some species to extinction. Our goal in the next section is to provide such a connection.

Distribution of distances from the edge of a triangle

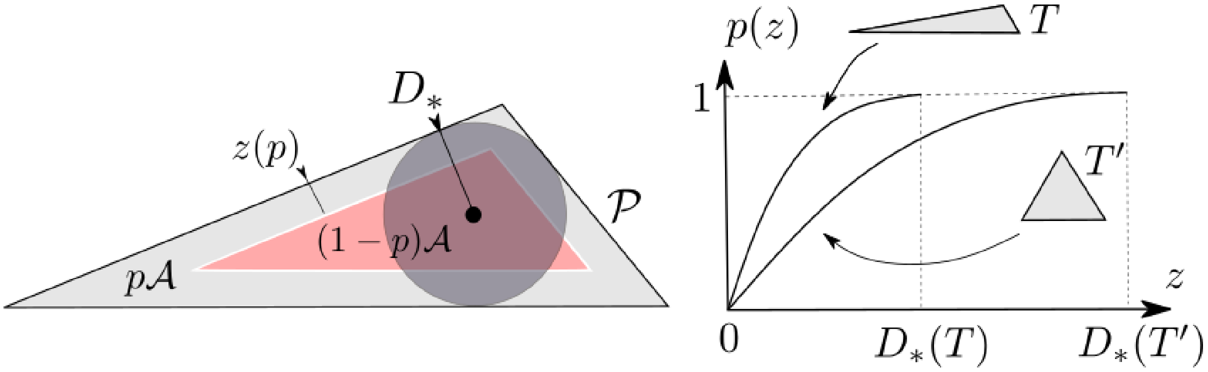
If the community is made of three species ($S = 3$) , the feasibility domain corresponds to a solid angle, a triangle drawn on a sphere (see Fig. 1). We thus start with a simplified analysis of regular triangles. This analogy allows us to gradually introduce the logic behind our geometrical approach (see Fig. 2). In this detour into simple trigonometry, which may seem removed from the initial ecological question, we will create a shape metric capable of encapsulating all the subtleties of shape differences between triangles (see Fig. 2).

The probability of a point to be at a distance greater than z from one of the triangle's edges corresponds to the relative area of the inscribed triangle whose own edges are exactly at a distance z from the boundaries of the original one (see the left panel of Fig. 2). Knowing $\mathcal{A}$, the area of the original triangle, and $\mathcal{A}'$, the area of the inscribed triangle, the proportion $p(z)$ of points that lie within a distance z from an edge is thus $p(z) = \dfrac{\mathcal{A} - \mathcal{A}'}{\mathcal{A}}$. It is an entertaining exercise to show that

$$p(z) = 1 - \left(1 - \frac{z}{D_*}\right)^2$$

showing that $p(z)$ is fully parameterized by a single number D_*, which is the radius of the largest disc contained in the triangle (indeed $p(D_*) = 1$). One can show that $D_* = 2\mathcal{A}/\mathcal{P}$ where $\mathcal{P}$ is the perimeter of the original triangle. For a fixed area $\mathcal{A}$, D_* is maximal for equilateral triangles (right panel of Fig. 2). This single distance measure D_* therefore allows us to quantitatively express differences in shape and size between triangles, and quantify, via the function $p(z)$ which encodes the whole distribution of distances to the triangle's edges.

Figure 2 – Left: Triangles are parameterized by their area $\mathcal{A}$ and perimeter $\mathcal{P}$. We are interested in the fraction $p(z)$ of points that lie within a distance z from an edge. We can show that $p(z)$ is fully parameterized by $D_* = 2\mathcal{A}/\mathcal{P}$, the radius of the inscribed disc, whose center is equidistant to all edges of the triangle. Right: at a fixed area $\mathcal{A}$, D_* grows as triangles become equilateral.



In the next section we will generalize this geometrical ideas to feasibility domains, that are not simple triangles, and can be of any dimension (i.e. any number of species). The aim is now to derive a similar function $p(z)$ applicable to ecological systems (L-V models).

Distribution of distances from the edge of a feasibility domain

Following Allen-Perkins et al. (2023), Cenci et al. (2018a), and De Laender et al. (2023), we consider perturbations as changes in environmental conditions that occurs on a long-time scale (so that a new equilibrium can be reached). Mathematically, we model a perturbation as a vector of variation δr of species intrinsic growth rates (i.e. whose components are the species-level variations δr_i). Using the euclidean norm of vectors $||\cdot||$ we then measure the relative intensity of this perturbation as (we will see why below)

$$\text{intensity} = \sqrt{S}\frac{||\delta r||}{||r||} \tag{5}$$

For any point r in a feasibility domain (so a feasible growth rate vector), we can measure its distance from the edge of the domain as the minimal perturbation intensity capable of leading at least one species to extinction. In the appendix we show that this distance can be directly computed as

$$d = \min\{\text{intensity};\ \text{such that } N_i(r+\delta r) = 0 \text{ for some } i\} = \min_i \frac{\sqrt{S}}{||r||}\frac{N_i(r)}{w_i} \tag{6}$$

in the last term, for any species i, w_i is the euclidean norm of the corresponding row of the inverse interaction matrix, which encodes that species sensitivity to environmental perturbations, with w_i measuring its maximal sensitivity (thus $w_i^2 = \sum_j (A^{-1})_{ij}^2$). Our main result, illustrated in Fig. 3, is a simple formula for the distribution of such distances, in the form of a cumulative function $p(z) = \mathbb{P}(d \leq z)$, which mimics the one given in the previous section for standard triangles, and is entirely parameterized by two characteristic distances and species richness S:

$$p(z) = \mathbb{P}(d \leq z) \approx 1 - \left(1 - \frac{z}{D_*}\right)^{S\sqrt{\frac{2}{\pi}}\frac{D_*}{D}} \tag{7}$$

As for standard triangles, D_* represents the largest distance within the domain, associated with its incenter r^*, also the most robust state of coexistence given the set of biotic interactions. Remarkably, we can deduce a simple formula for both D_* and r^*. Indeed, in the appendix we show that, if w is the vector of maximal species sensitivities (i.e. whose components are the species-level values w_i), then

$$r^* = Aw \text{ and } D_* = \frac{\sqrt{S}}{||Aw||} \tag{8}$$

One can check that in the absence of interactions, and thus when A is diagonal, we have $D_* = 1$ (this is a consequence of our choice of normalisation of perturbation intensity). The formula for the distribution of distances differs from the one for triangles in that the maximal distance is not the only relevant distance, the one driving the behaviour at small z values, so near the edge of the domain, is in fact

$$D = 1/\frac{1}{S}\sum_{i=1}^{S} w_i \sqrt{\frac{|A^\top A|}{|(A^\top A)_{/i}|}}\frac{\Omega_{/i}}{\Omega} \tag{9}$$

with $X_{/i}$ notation meaning for any matrix X, the corresponding matrix without the $i-th$ row and column, and $\Omega_{/i}$ is essentially the relative volume of the feasibility domain for the community without species i (but see the appendix for a more precise expression and derivation). The initial slope of $p(z)$ is given by $S\sqrt{\frac{2}{\pi}}/D$ and determines the behavior of $p(z)$ at small z values, so near the edge of the domain. We see that this slope explicitly grows with species richness. The latter behavior occurs because when there are many species present, it is ever more likely that one of them is close to local extinction. This diversity effect will tend to take a dominant part in shaping the function $p(z)$. Geometrically speaking, this effect comes from the fact that in high dimensions, even very thin neighbourhoods of the edge of a closed object will cover a dominant fraction of the overall volume of that object. The expression for D and D_* clearly differ. Nonetheless those two distances are closely related and take very similar values, with $D \approx D_*$ for the vast majority of random interactions matrices that we generated, and even more so when considering empirically inferred matrices (See supplementary figure A2). Finally, we can connect the characteristic distances D_* and D with the relative volume of the domain Ω (from the first section). In supplementary figure A1 and in the mathematical appendix we explain why we may expect that, roughly speaking

$$D_* \approx D \approx \Omega^{\frac{2}{S}} \tag{10}$$

This approximate relationship, which must be understood as an equivalence of orders of magnitudes and not of precise values, taken together with equation 7 connects the size of the domain to its shape, and to the probability that a given perturbation can push species to extinction. In the appendices we expand on the latter point by considering randomly changing environments trough time. We show in simulations that the equivalent metrics of equation 10 do predict the duration of stable coexistence periods (See supplementary figure B1 in the "Persistence of species in simulated ecological system" appendix).

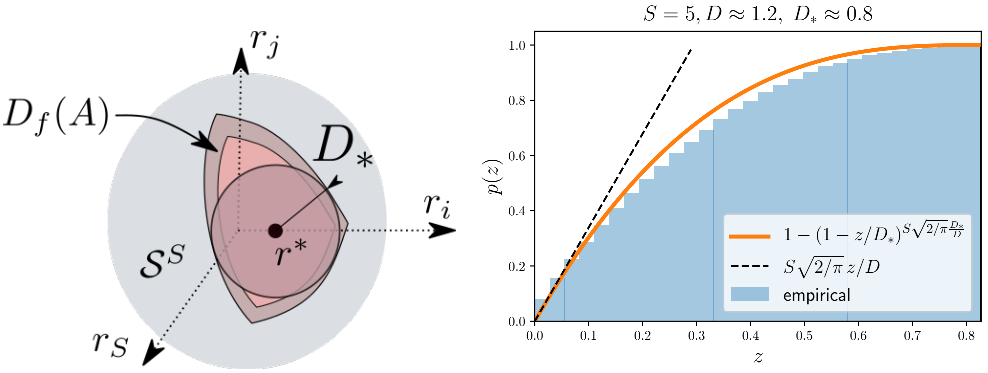


Figure 3 – Empirical distribution of distances $p(z)$ for a random interaction matrix of $S = 5$ species, compared with its analytical approximation (equation 7 in orange line). To obtain the empirical distribution, we randomly sampled distances to the edge z and calculated, for each distance z, the fraction $p(z)$ of coexistence states closest to the edge.

Contextual species vulnerability

The above analyses of the distribution of distances to the edge of feasibility enable us to characterize the robustness of coexistence of an ecological community. We now take the analysis further to show that the characteristic distance D and the incenter r_* (that determine the distribution of distance to the edge of coexistence) can be used to study the contributions and contextual roles of species in maintaining coexistence. To understand why, we can start with the incenter components

$$r_i^* = \sum_j A_{ij} w_j \tag{11}$$

and see that it can be interpreted a measure of the strength of competition exerted by the community on species i, the sum of interactions felt by that species, but where each per-capita interaction term A_{ij} is weighted by the partner's maximal sensitivity to perturbations (the terms w_j). Here a weak interaction with a highly sensitive species (a large w_j) can contribute more than a weak interaction with a highly stable population (a small w_j). If $r_i^* = 1$, the community has a neutral effect, equal to that of the species on its own. If it is less than 1, the community facilitates that species (see supplementary figure C1). On the other hand, the distance D describes the edges of the feasibility domain. It reads as the inverse of an average of S elements, one for each species:

$$\mathrm{SV}_i = w_i \sqrt{\frac{|A^\top A|}{|(A^\top A)_{/i}|} \frac{\Omega_{/i}}{\Omega}} \tag{12}$$

Those terms determine the distribution of coexistence states for which that species is within a certain perturbation distance from extinction. Hence, they relate to the individual vulnerability of each species.

We can combine the species-level measures r_i^* and SV_i by viewing them as the species coordinates on a two dimensional map, in other words, plotting them against each other (see Fig. 4). Intuitively, the two should be correlated: species that perceive a hostile biotic environment should also be the most vulnerable, and vice versa. This should lead to the definition of "two" particular roles: on the one hand, vulnerable and repressed, and on the other robust and facilitated. The results obtained by applying our measures to empirical data (Fig. 4) show that it doesn't always have to be this simple, and that it is possible to define two other non-trivial qualitative "roles".

Application to data from a grassland experiment

Our approach characterizes the robustness of coexistence at two levels: at the scale of the community as a whole, but also at the species scale. Here, we illustrate the insights that this approach can generate for real ecological communities. We revisit data from Van Ruijven and Berendse (2009) and its subsequent analysis by Barbier et al. (2021), compiled from long-term studies of plant communities in the experimental gardens of Wageningen University, Netherlands. Here we directly use the results of Barbier et al. (2021), who estimated the interaction strengths between 8 plant species, as well as their carrying capacities. Interactions refer here to a Lotka-Volterra parametrization that differs from the one that implicitly follows from equation 1. Indeed, monocultures where used to infer species' carrying capacities K_i, and it is those that we consider as proxies for the abiotic conditions (and not intrinsic growth rates r_i). The relevant interaction matrix, inferred using duo-culture experiments, follows from re-writing the L-V equations as

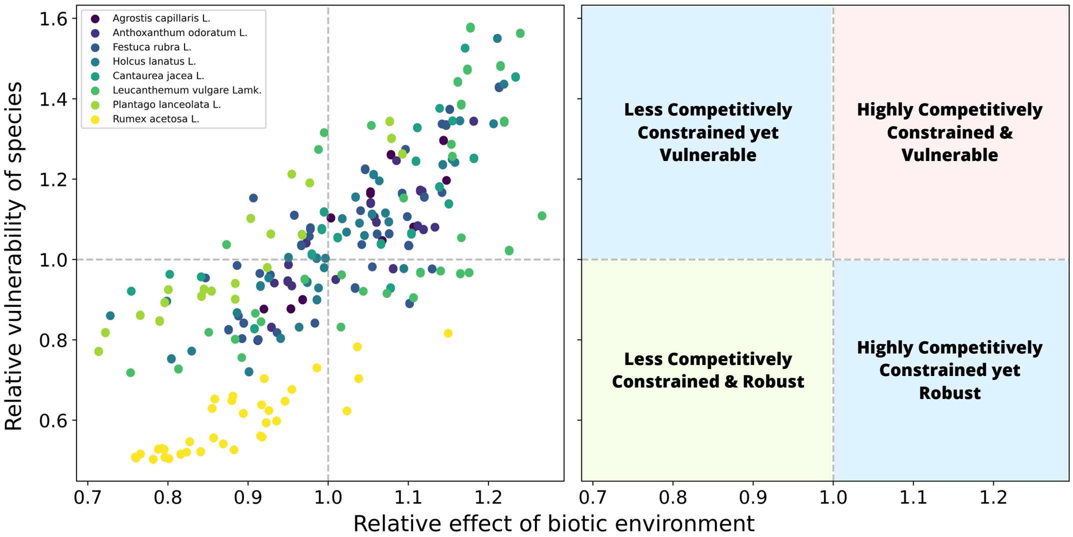
$$\frac{dN_i}{dt} = \frac{r_i N_i}{K_i}\left(K_i - \sum_j A_{ij} N_j\right) \tag{13}$$

In this parametrization, A_{ij} has no dimensions and satisfies $A_{ii} \equiv 1$. On the basis of pairwise interaction values, we then reconstruct interaction matrices consisting of 4 species, which have been experimentally realized (Van Ruijven and Berendse, 2009). All the pairwise interactions values and carrying capacity values derived from their work are available on Zotero (See Data, script, code, and supplementary information availability section below).

To show the role of the same species in different communities, we calculated SV_i and r_i^* for each species within all four-species communities (See left panel of Fig. 4). We normalized these values by the mean value within each community to obtain relative species vulnerability and relative biotic effects on species, as the same species can hold different roles for the robustness of coexistence, depending on the biotic environment. Furthermore, while we unsurprisingly find the same trend as in the supplementary figure C1 (The majority of points being located in the red and green areas and being either "Highly Competitively Constrained and Vulnerable" or "Less Competitively Constrained and Robust"), we can observe non-trivial cases (blue areas of the figure). In these cases, the biotic interactions affecting the species in question are not sufficient to explain its vulnerability.

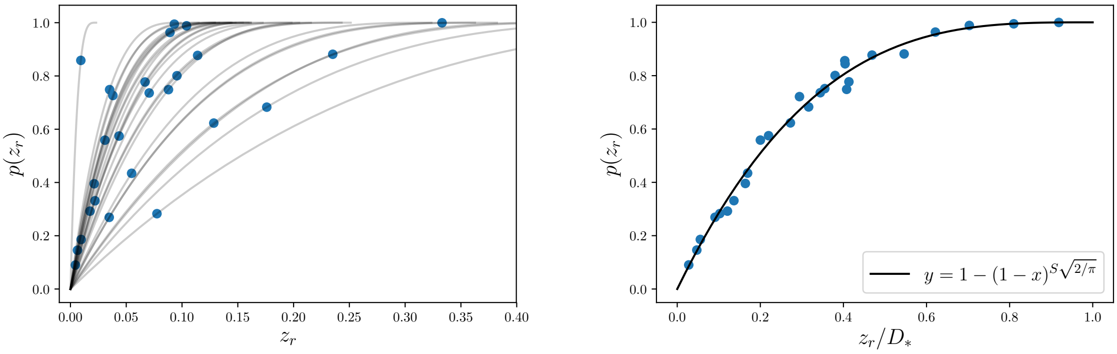
The "Less Competitively Constrained yet Vulnerable" points correspond to a case where the strong vulnerability comes from its competitive forces applied to (and not received by) other species. Indeed, to achieve coexistence, it must necessarily be of low abundance and therefore vulnerable, so that other species do not suffer too greatly from its presence. The "High Competitively Constrained yet Robust" points correspond to the case where species are useful for the coexistence of others and therefore have a high abundance (and a low vulnerability of coexistence) despite higher competitive forces experienced. These non-trivial cases explain why some points in supplementary figure C1 deviate from the expected relation.

Figure 4 – Analysis of the robustness of coexistence at the species scales. Each point on the left graph represents all the individual positions of the 8 species of the dataset within the 35 possible 4-species communities where they are present. On the x-axis, the relative effect of interactions (biotic environment) is indicated (r_i^* divided by the mean value for all species in the community). On the y-axis, the vulnerability of each species is indicated (SV_i, divided by the average on all species of the community). This allows us to define 4 notable cases, represented on the graph on the right by the different colors.



Interestingly, the points cluster relatively well by species. This suggests that within the different 4-species communities formed by the 8 selected species, the species tend to maintain a relatively identical biotic role. Note that the abiotic environment in which these species have grown is supposedly the same. This makes ecological sense, as the biotic roles of each species depend on their phenotypic traits, and are therefore fixed by the biology of each species. For example, *Rumex Acetosa L.* is predominantly found in the green zone in Fig. 4, suggesting good persistence through low competitive forces. This fits rather well with its characterization as a weed species, present in a wide range of environments and able to coexist and persist within many ecosystems (Korpelainen and Pietiläinen, 2020).

Figure 5 – Using the empirically inferred interaction matrix between 8 plant species (Barbier et al., 2021) and their carrying capacities (taking median values for simplicity), we assembled all theoretically feasible 4-species communities (27 out of the 70 different combinations turn out to be feasible). Left: The interaction matrix for each community defines a curve, and the realized community gives the point on the curve. Large values of z_r (x-axis) implies high robustness (i.e. large distance from the edge of feasibility), whereas large values of $p(z_r)$ means that most communities with similar interactions are less robust. The higher this value, the better the match between the realized intrinsic parameters and biotic interactions. Right: how well suited interactions and carrying capacities go together is more clearly visualised by rescaling realized distances by the maximal distance D_*. Indeed all curves collapse on a single one and we see that the communities span the whole range of $p(z)$, meaning that some are as robust as they could be, while others are much more vulnerable than what could have been expected. The analytical graph is $p(z) = 1 - (1 - \frac{z}{D_*})^{S\sqrt{\frac{2}{\pi}}}$ (here $S = 4$), its accuracy to predict the actual $p(z)$ values is due to the fact that $D_* \approx D$ (see supplementary figure A2)



Since the abiotic environment was assumed to be the same across the experiment, we can now determine how well or ill suited it was to particular species combinations (in terms of favouring robust coexistence). Indeed, using the carrying capacities determined during the experiments, we can determine z_r, the minimal distance to the edge of the realized community, and $p(z_r)$, the proportion of points within this distance. This allows us to place all the communities on the z/D_* and $p(z)$ curve (see Fig. 5). If $p(z_r) \approx 1$, it means that in this environment, the realized community had the most advantageous combination of biotic interactions (interaction matrix A) and intrinsic species parameters (carrying capacity K), in terms of robustness of coexistence. If $p(z_r)$, it means that this environment has led to a kind of mismatch between species interactions and species growth rates, making coexistence far less robust than what it could have been, given the set of species and their interactions.

Discussion

For a given community on interacting species, the function $z \mapsto p$ maps a value of environmental perturbation intensity z to the fraction p of coexistence states from which coexistence can be lost following such perturbations. We showed here that $p(z)$ is a rich object to study the robustness of species coexistence, and how biotic interactions affect it, while not reducing robustness to a single number.

In Lotka-Volterra models, $p(z)$ precisely characterizes the shape of the feasibility domain, which is the set of growth rate vectors that allow stable coexistence between all species. Indeed $p(z) = \mathbb{P}(d \leq z)$ determines the distribution of distances to the edge of coexistence (see Fig. 3), where for a given coexistence state, the distance to an edge corresponds to the state's "robustness" or "full resistance", as defined by Lepori et al. (2024) and Medeiros et al. (2021).

We showed that the function $p(z)$ is fully parameterized by species richness S and two characteristic distances D and D_*, both equal to 1 in the absence of interactions. More precisely

$$p(z) = 1 - (1 - \frac{z}{D_*})^{S\sqrt{\frac{2}{\pi}}\frac{D_*}{D}}$$

D_* is the maximal distance within the feasibility domain and thus represents the robustness of the most robust state, r^*, such that $p(z = D_*) = 1$. We derived remarkably simple formulas for D_* and r^* (See Eq. 9 and 11 and mathematical appendix), based on the interaction matrix and its inverse. Unpacking the expression for r_* allowed us to give a species-level characterization that can be interpreted as measuring the effective amount of competition that any given species feels, where its interactions are weighted by the sensitivities of its interacting partners.

The other important distance, D, once divided by S, determines the behaviour of $p(z)$ at small perturbation intensity values, in the sense that $p(z) \approx S\sqrt{\frac{2}{\pi}}\frac{z}{D}$ describing the edges of feasibility, which take up most of its volume if S is large[1]. The ratio S/D can be used to understand how many species can be grouped together while maintaining a high percentage of robust states. More precisely, if we want to guarantee that a proportion p of coexistence states is robust to perturbations of intensity ϵ, then maximizing diversity amounts to solving

$$\max\{S \mid \frac{S}{D} \leq \sqrt{\pi/2}\frac{p}{\epsilon}\}$$

whose solution will take the form of $S = \sqrt{\pi/2}\frac{p}{\epsilon} \times D$, so proportional to D.

The expression for D is less simple than the one for D_*, but can also be used to give complementary species-level characterization of coexistence. In line with Allen-Perkins et al. (2023), we can decompose D to measure the robustness of each species persistence conditioned on overall coexistence. This interpretation, together with the one relating r_* to effective competition pressure, can be used to reveal the contextual roles of species in maintaining coexistence. The biotic context created within a coexisting community can be favorable or unfavorable to individual species through the balance of interactions they receive and emit and how hostile they are to others (See different panels of Fig. 4). It is interesting to note that the species present in the dataset used in the study seem to retain relatively the same role regardless of community composition. It would be interesting to extend this analysis to larger datasets to study the consistency of species roles in maintaining robust coexistence. If we consider the contribution to the community-scale robustness of coexistence as a function rendered by a species within the community, it is likely that certain species correspond to "key species" (Power et al., 1996; Whittaker and Cottee-Jones, 2012).

Broadly speaking, our theory highlights a negative effect, amplified by species richness, of the intensity of the interaction forces and the sensitivity of the species on the robustness of coexistence. Figure 4 and supplementary figure C1 also show the relationship between strong inter-specific competition faced by species and their vulnerability of coexistence. These results are consistent with the existing literature on the effect of interactions on community coexistence or stability under environmental perturbations (Barabás et al., 2016; Chesson, 2000; Hale et al., 2020; Mccann et al., 1998; Vallina and Le Quéré, 2011). The fact that features of the inverse interaction matrix are present in both D and D_* highlights the importance of network structure, as the inverse matrix encodes net effects between species, via all indirect interaction pathways. For the same overall mean interaction strength, net effects can be very different depending on the way the matrix A is organized. This is consistent with previous research on the effect of network structure on coexistence (especially in cases with more than two species) as on other stability notions (Barabás et al., 2016; Cenci et al., 2018b; Lurgi et al., 2016; Serván et al., 2018). This leads to an important ecological conclusion: vulnerability to extinction depends on how a species is affected by others through direct interactions, combined with the sensitivities of

---

[1] This last remark is only a geometrical way of saying that for many interacting species, in the absence of prior knowledge of abiotic conditions, there is a high chance that at least one of those species is close to extinction.

those species (how they amplify environmental change). Here sensitivity is a potentially collective notion that arises from indirect interactions between species, and is thus sensitive to the interaction structure.

As in previous studies of asymmetry of the feasibility domain, our theory strongly depends on the way environmental disturbances are modeled (Allen-Perkins et al., 2023; Cenci et al., 2018a; De Laender et al., 2023; Lepori et al., 2024). This highlights the importance of taking into account the type of disturbance when studying the stability of a community (Arnoldi et al., 2018; Arnoldi et al., 2019; Bender et al., 1984) and suggests that different results could be obtained by considering other types of disturbance (ie. that vary through time, and/or scale with species standing biomass). Deepening our theory to account for more general types of disturbance could be an interesting direction.

Coexistence is defined as the maintenance of positive abundance of all species in a community. No attention is paid to total biomass, ecosystem functions, turnover, or processes at the meta-community level. Our results should therefore not be interpreted as evidence of a negative effect of biodiversity on stability in the sense of maintaining biomass or ecological function over time (Loreau and Mazancourt, 2013), nor on the resistance or resilience of the the community (Arnoldi et al., 2016; Kéfi et al., 2019). It simply highlights the difficulty for complex interaction networks to generate communities that can tolerate environmental disturbances without losing any species. This vision of a fixed community and coexistence seen as the absolute persistence of all species over time is, however, clearly limited and open to criticism. It would be interesting to develop approaches that include turnover or variations in species interactions over time.

Another caveat is the supposed independence between biotic and abiotic parameters. This unrealistic assumption means that a change in abiotic environmental conditions (disruption of growth rates or carrying capacity) should not change biotic interactions. This assumption is necessary to define the feasibility domain (Saavedra et al., 2017). However, the empirical applications we present (determination of the biotic role of different species within several communities; quantification of the adequation between a given abiotic environment and a certain biotic assemblage) illustrate how to overcome this issue. Indeed, in the experimental data, the abiotic environment is the same for each community studied and is not subject to change.

Overall, this study provides an understanding of the link between the conditions under which communities coexist and the robustness of this coexistence. On the one hand, the analytical results provide a clear explanation of the relationships between the various mathematical elements involved in feasibility domain analysis. On the other hand, they enable us to link the interpretations made specifically through the analysis of the notion of feasibility domain to more general notions of community ecology. In doing so, we have linked different measures of stability and placed the robustness of coexistence within the multidimensional concept of ecological stability (Donohue et al., 2016; Radchuk et al., 2019).

Acknowledgements

We thank Matthieu Barbier for the data provided and useful discussions. We also thank Rudolf Rohr and Nicolas Loeuille for inspiring discussions throughout the article writing process. We finally thanks reviewers and editor for their very constructive feedback, valuable suggestions and timely handling of our manuscript.

Fundings

This work was supported by the TULIP Laboratory of Excellence (ANR-10-LABX-41) and the BIOSTASES Advanced Grant, funded by the European Research Council (ERC) under the European Union's Horizon 2020 research and innovation programme (666971).

Conflict of interest disclosure

The authors declare that they comply with the PCI rule of having no financial conflicts of interest in relation to the content of the article.

Data, script, code, and supplementary information availability

Script and data useful for "Application to data from a grassland experiment" section are available online: 10.5281/zenodo.10534234; For more information on the dataset, please contact Barbier et al. (2021) ; Desallais et al., 2024b

Supplementary information, including appendices A, B and C and a mathematical appendix, is available online : (`https://zenodo.org/doi/10.5281/zenodo.12744286`; Desallais et al., 2024a

References

Abrams (1984). *Variability in resource consumption rates and the coexistence of competing species*. *Theoretical Population Biology* **25**, 106–124. `https://doi.org/https://doi.org/10.1016/0040-5809(84)90008-X`. URL: `https://www.sciencedirect.com/science/article/pii/004058098490008X`.

Abrams, Brassil, Holt (2003). *Dynamics and responses to mortality rates of competing predators undergoing predator–prey cycles*. *Theoretical Population Biology* **64**, 163–176. `https://doi.org/10.1016/S0040-5809(03)00067-4`. URL: `https://linkinghub.elsevier.com/retrieve/pii/S0040580903000674` (visited on 01/09/2023).

Allen-Perkins A, García-Callejas D, Bartomeus I, Godoy O (2023). *Structural asymmetry in biotic interactions as a tool to understand and predict ecological persistence*. *Ecology Letters* **n/a**. `https://doi.org/https://doi.org/10.1111/ele.14291`.

Armstrong RA, McGehee R (1976). *Coexistence of species competing for shared resources*. *Theoretical Population Biology* **9**, 317–328. `https://doi.org/10.1016/0040-5809(76)90051-4`. URL: `https://linkinghub.elsevier.com/retrieve/pii/0040580976900514` (visited on 02/08/2022).

Arnoldi JF, Loreau M, Haegeman B (2016). *Resilience, reactivity and variability: A mathematical comparison of ecological stability measures*. *Journal of Theoretical Biology* **389**, 47–59. `https://doi.org/10.1016/j.jtbi.2015.10.012`. URL: `https://linkinghub.elsevier.com/retrieve/pii/S0022519315005056` (visited on 02/08/2022).

Arnoldi JF, Bideault A, Loreau M, Haegeman B (2018). *How ecosystems recover from pulse perturbations: A theory of short- to long-term responses*. *Journal of theoretical biology* **436**, 79–92. `https://doi.org/10.1016/j.jtbi.2017.10.003`. URL: `https://www.ncbi.nlm.nih.gov/pmc/articles/PMC5675055/` (visited on 01/08/2023).

Arnoldi JF, Loreau M, Haegeman B (2019). *The inherent multidimensionality of temporal variability: how common and rare species shape stability patterns*. *Ecology letters* **22**, 1557–1567. `https://doi.org/10.1111/ele.13345`. URL: `https://www.ncbi.nlm.nih.gov/pmc/articles/PMC6756922/` (visited on 07/18/2022).

Barabás G, J. Michalska-Smith M, Allesina S (2016). *The Effect of Intra- and Interspecific Competition on Coexistence in Multispecies Communities*. *The American Naturalist* **188**. Publisher: The University of Chicago Press, E1–E12. `https://doi.org/10.1086/686901`. URL: `https://www.journals.uchicago.edu/doi/full/10.1086/686901` (visited on 07/25/2022).

Barbier M, Mazancourt C, Loreau M, Bunin G (2021). *Fingerprints of High-Dimensional Coexistence in Complex Ecosystems*. *Physical Review X* **11**. Publisher: American Physical Society, 011009. `https://doi.org/10.1103/PhysRevX.11.011009`. URL: `https://link.aps.org/doi/10.1103/PhysRevX.11.011009` (visited on 07/18/2022).

Bartomeus I, Saavedra S, Rohr RP, Godoy O (2021). *Experimental evidence of the importance of multitrophic structure for species persistence*. *Proceedings of the National Academy of Sciences* **118**, e2023872118. `https://doi.org/10.1073/pnas.2023872118`. URL: `http://www.pnas.org/lookup/doi/10.1073/pnas.2023872118` (visited on 02/08/2022).

Bender EA, Case TJ, Gilpin ME (1984). *Perturbation Experiments in Community Ecology: Theory and Practice*. *Ecology* **65**, 1–13. `https://doi.org/10.2307/1939452`. URL: `http://doi.wiley.com/10.2307/1939452` (visited on 02/08/2022).

Brose U, Williams RJ, Martinez ND (2006). *Allometric scaling enhances stability in complex food webs*. *Ecology Letters* **9**, 1228–1236. `https://doi.org/10.1111/j.1461-0248.2006.00978.x`.

Cenci S, Montero-Castaño A, Saavedra S (2018a). *Estimating the effect of the reorganization of interactions on the adaptability of species to changing environments*. *Journal of Theoretical Biology* **437**, 115–125. `https://doi.org/https://doi.org/10.1016/j.jtbi.2017.10.016`. URL: `https://www.sciencedirect.com/science/article/pii/S0022519317304794`.

Cenci S, Song C, Saavedra S (2018b). *Rethinking the importance of the structure of ecological networks under an environment-dependent framework*. *Ecology and Evolution* **8**, 6852–6859. `https://doi.org/10.1002/ece3.4252`. URL: `https://onlinelibrary.wiley.com/doi/abs/10.1002/ece3.4252` (visited on 07/25/2022).

Chesson P (2000). *Mechanisms of Maintenance of Species Diversity*. *Annual Review of Ecology and Systematics* **31**, 343–366. `https://doi.org/10.1146/annurev.ecolsys.31.1.343`. URL: `https://www.annualreviews.org/doi/10.1146/annurev.ecolsys.31.1.343` (visited on 02/08/2022).

Coulson T, Kendall BE, Barthold J, Plard F, Schindler S, Ozgul A, Gaillard JM (2017). *Modeling Adaptive and Nonadaptive Responses of Populations to Environmental Change*. *The American Naturalist* **190**. PMID: 28829647, 313–336. `https://doi.org/10.1086/692542`. eprint: `https://doi.org/10.1086/692542`. URL: `https://doi.org/10.1086/692542`.

De Laender F, Carpentier C, Carletti T, Song C, Rumschlag SL, Mahon MB, Simonin M, Meszéna G, Barabás G (2023). *Mean species responses predict effects of environmental change on coexistence*. *Ecology Letters* **26**, 1535–1547. `https://doi.org/https://doi.org/10.1111/ele.14278`. eprint: `https://onlinelibrary.wiley.com/doi/pdf/10.1111/ele.14278`. URL: `https://onlinelibrary.wiley.com/doi/abs/10.1111/ele.14278`.

Deng J, Taylor W, Saavedra S (2022). *Understanding the impact of third-party species on pairwise coexistence*. *PLOS Computational Biology* **18**, 1–21. `https://doi.org/10.1371/journal.pcbi.1010630`. URL: `https://doi.org/10.1371/journal.pcbi.1010630`.

Desallais M, Loreau M, Arnoldi JF (2024a). *Appendices of the "The distribution of distances to the edge of species coexistence" scientific article*. `https://doi.org/10.5281/zenodo.12744286`. URL: `https://doi.org/10.5281/zenodo.12744287`.

Desallais M, Loreau M, Jean-François A (2024b). *Functions and scripts used on the "The distribution of distances to the edge of species coexistence" scientific article*. `https://doi.org/10.5281/zenodo.10534234`. URL: `https://doi.org/10.5281/zenodo.10905742`.

Donohue I, Hillebrand H, Montoya JM, Petchey OL, Pimm SL, Fowler MS, Healy K, Jackson AL, Lurgi M, McClean D, O'Connor NE, O'Gorman EJ, Yang Q (2016). *Navigating the complexity of ecological stability*. *Ecology Letters* **19**. Ed. by Frederick Adler, 1172–1185. `https://doi.org/10.1111/ele.12648`. URL: `https://onlinelibrary.wiley.com/doi/10.1111/ele.12648` (visited on 02/08/2022).

Grilli J, Adorisio M, Suweis S, Barabás G, Banavar JR, Allesina S, Maritan A (2017). *Feasibility and coexistence of large ecological communities*. *Nature Communications* **8**, 14389. `https://doi.org/10.1038/ncomms14389`. URL: `http://www.nature.com/articles/ncomms14389` (visited on 02/08/2022).

Hale KRS, Valdovinos FS, Martinez ND (2020). *Mutualism increases diversity, stability, and function of multiplex networks that integrate pollinators into food webs*. *Nature Communications* **11**, 2182. `https://doi.org/10.1038/s41467-020-15688-w`. URL: `https://doi.org/10.1038/s41467-020-15688-w`.

Hastings A (1980). *Disturbance, coexistence, history, and competition for space*. *Theoretical Population Biology* **18**, 363–373. `https://doi.org/10.1016/0040-5809(80)90059-3`. URL: `https://linkinghub.elsevier.com/retrieve/pii/0040580980900593` (visited on 02/08/2022).

Hutchinson GE (1961). *The Paradox of the Plankton*. *THE AMERICAN NATURALIST*, 9.

Kéfi S, Domínguez-García V, Donohue I, Fontaine C, Thébault E, Dakos V (2019). *Advancing our understanding of ecological stability*. *Ecology Letters* **22**. Ed. by Tim Coulson, 1349–1356. `https://doi.org/10.1111/ele.13340`. URL: `https://onlinelibrary.wiley.com/doi/10.1111/ele.13340` (visited on 02/07/2022).

Korpelainen H, Pietiläinen M (2020). *Sorrel (Rumex acetosa L.): Not Only a Weed but a Promising Vegetable and Medicinal Plant*. *The Botanical Review* **86**, 234–246. `https://doi.org/10.1007/s12229-020-09225-z`. URL: `https://doi.org/10.1007/s12229-020-09225-z`.

Lepori VJ, Loeuille N, Rohr RP (2024). *Robustness versus productivity during evolutionary community assembly: short-term synergies and long-term trade-offs*. *Proceedings of the Royal Society B: Biological Sciences* **291**, 20232495. `https://doi.org/10.1098/rspb.2023.2495`. eprint: `https://royalsocietypublishing.org/doi/pdf/10.1098/rspb.2023.2495`. URL: `https://royalsocietypublishing.org/doi/abs/10.1098/rspb.2023.2495`.

Levins R (1968). *Evolution in Changing Environments: Some Theoretical Explorations*. Princeton University Press. URL: `https://books.google.fr/books?id=yOQ9DwAAQBAJ&lr=&source=gbs_navlinks_s`.

Loreau M, Mazancourt C (2013). *Biodiversity and ecosystem stability: a synthesis of underlying mechanisms*. *Ecology Letters* **16**. Ed. by Emmett Duffy, 106–115. `https://doi.org/10.1111/ele.12073`. URL: `https://onlinelibrary.wiley.com/doi/10.1111/ele.12073` (visited on 02/07/2022).

Lurgi M, Montoya D, Montoya JM (2016). *The effects of space and diversity of interaction types on the stability of complex ecological networks*. *Theoretical Ecology* **9**, 3–13. `https://doi.org/10.1007/s12080-015-0264-x`. URL: `https://doi.org/10.1007/s12080-015-0264-x` (visited on 08/12/2022).

Mccann K, Hastings A, Huxel G (1998). *Weak Trophic Interactions and the Balance of Nature*. *Nature* **395**, 794–798. `https://doi.org/10.1038/27427`.

Medeiros LP, Song C, Saavedra S (2021). *Merging dynamical and structural indicators to measure resilience in multispecies systems*. *Journal of Animal Ecology* **90**, 2027–2040. `https://doi.org/10.1111/1365-2656.13421`. URL: `https://onlinelibrary.wiley.com/doi/10.1111/1365-2656.13421` (visited on 02/08/2022).

Meszéna G, Gyllenberg M, Pásztor L, Metz JA (2006). *Competitive exclusion and limiting similarity: A unified theory*. *Theoretical Population Biology* **69**, 68–87. `https://doi.org/https://doi.org/10.1016/j.tpb.2005.07.001`. URL: `https://www.sciencedirect.com/science/article/pii/S004058090500095X`.

Otto SB, Rall BC, Brose U (2007). *Allometric degree distributions facilitate food-web stability*. *Nature* **450**, 1226–1229. `https://doi.org/10.1038/nature06359`.

Power ME, Tilman D, Estes JA, Menge BA, Bond WJ, Mills LS, Daily G, Castilla JC, Lubchenco J, Paine RT (1996). *Challenges in the Quest for Keystones: Identifying keystone species is difficult—but essential to understanding how loss of species will affect ecosystems*. *BioScience* **46**, 609–620. `https://doi.org/10.2307/1312990`. eprint: `https://academic.oup.com/bioscience/article-pdf/46/8/609/650270/46-8-609.pdf`. URL: `https://doi.org/10.2307/1312990`.

Radchuk V, Laender FD, Cabral JS, Boulangeat I, Crawford M, Bohn F, Raedt JD, Scherer C, Svenning JC, Thonicke K, Schurr FM, Grimm V, Kramer-Schadt S (2019). *The dimensionality of stability depends on disturbance type*. *Ecology Letters* **22**. Ed. by Ian Donohue, 674–684. `https://doi.org/10.1111/ele.13226`. URL: `https://onlinelibrary.wiley.com/doi/10.1111/ele.13226` (visited on 02/01/2022).

Ribando JM (2006). *Measuring Solid Angles Beyond Dimension Three*. *Discrete & Computational Geometry* **36**, 479–487. `https://doi.org/10.1007/s00454-006-1253-4`. URL: `http://link.springer.com/10.1007/s00454-006-1253-4` (visited on 02/08/2022).

Rohr RP, Saavedra S, Bascompte J (2014). *On the structural stability of mutualistic systems*. *Science* **345**, 1253497. `https://doi.org/10.1126/science.1253497`. URL: `https://www.science.org/doi/10.1126/science.1253497` (visited on 02/08/2022).

Rohr RP, Saavedra S, Peralta G, Frost CM, Bersier LF, Bascompte J, Tylianakis JM (2016). *Persist or Produce: A Community Trade-Off Tuned by Species Evenness*. *The American Naturalist* **188**. PMID: 27622875, 411–422. `https://doi.org/10.1086/688046`. eprint: `https://doi.org/10.1086/688046`. URL: `https://doi.org/10.1086/688046`.

Roughgarden J (1975). *A Simple Model for Population Dynamics in Stochastic Environments*. *The American Naturalist* **109**, 713–736. URL: `http://www.jstor.org/stable/2459866`.

Saavedra S, Rohr RP, Bascompte J, Godoy O, Kraft NJB, Levine JM (2017). *A structural approach for understanding multispecies coexistence*. *Ecological Monographs* **87**, 470–486. `https://doi.`

org/10.1002/ecm.1263. URL: https://onlinelibrary.wiley.com/doi/10.1002/ecm.1263 (visited on 02/07/2022).

Serván CA, Capitán JA, Grilli J, Morrison KE, Allesina S (2018). *Coexistence of many species in random ecosystems*. *Nature Ecology & Evolution* **2**. Number: 8 Publisher: Nature Publishing Group, 1237–1242. https://doi.org/10.1038/s41559-018-0603-6. URL: https://www.nature.com/articles/s41559-018-0603-6 (visited on 07/25/2022).

Song C, Altermatt F, Pearse I, Saavedra S (2018a). *Structural changes within trophic levels are constrained by within-family assembly rules at lower trophic levels*. *Ecology Letters* **21**. Ed. by Josè Marìa Gomez, 1221–1228. https://doi.org/10.1111/ele.13091. URL: https://onlinelibrary.wiley.com/doi/10.1111/ele.13091 (visited on 02/08/2022).

Song C, Rohr RP, Saavedra S (2018b). *A guideline to study the feasibility domain of multi-trophic and changing ecological communities*. *Journal of Theoretical Biology* **450**, 30–36. https://doi.org/10.1016/j.jtbi.2018.04.030. URL: https://linkinghub.elsevier.com/retrieve/pii/S0022519318302042 (visited on 02/08/2022).

Vallina SM, Le Quéré C (2011). *Stability of complex food webs: Resilience, resistance and the average interaction strength*. *Journal of Theoretical Biology* **272**, 160–173. https://doi.org/10.1016/j.jtbi.2010.11.043. URL: https://www.sciencedirect.com/science/article/pii/S0022519310006387 (visited on 07/25/2022).

Van Ruijven J, Berendse F (2009). *Long-term persistence of a positive plant diversity–productivity relationship in the absence of legumes*. *Oikos* **118**, 101–106. https://doi.org/10.1111/j.1600-0706.2008.17119.x. URL: https://onlinelibrary.wiley.com/doi/abs/10.1111/j.1600-0706.2008.17119.x (visited on 12/23/2023).

Volterra V (1926). *Fluctuations in the Abundance of a Species considered Mathematically1*. *Nature* **118**, 558–560. https://doi.org/10.1038/118558a0. URL: https://www.nature.com/articles/118558a0 (visited on 07/15/2022).

Whittaker R, Cottee-Jones HEW (2012). *The keystone species concept: a critical appraisal*. *Frontiers of Biogeography* **4**, 117–127. https://doi.org/10.21425/F54312533.

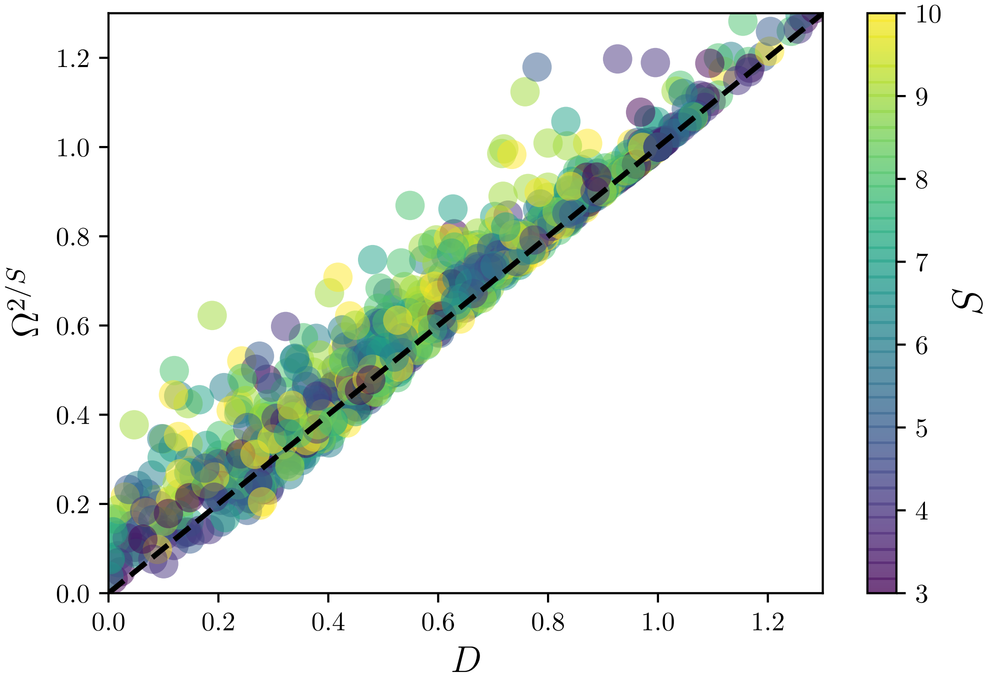
Williams RJ (2008). *Effects of network and dynamical model structure on species persistence in large model food webs*. *Theoretical Ecology* **1**, 141–151. https://doi.org/10.1007/s12080-008-0013-5. URL: http://link.springer.com/10.1007/s12080-008-0013-5 (visited on 07/19/2022).

Appendix A
The distribution of distances to the edge of species coexistence
Relation between Ω, D and D_*

Mario Desallais, Michel Loreau, and Jean-François Arnoldi

In this study, we highlight the importance of taking into account the shape of the feasibility domain and its size to characterize the robustness of coexistence induced by species interactions. Two different measures, therefore, emerge: Ω, a proxy for the probability of coexistence, and D (or D_*), a proxy for the robustness of coexistence. However, these are not independent. Fig. 1 shows the relationship between the two values.

Figure 1 – Relation between D and $\Omega^{2/S}$ for random matrices of size S x S. If D controls the distribution of distance to the edge of feasibility, Ω corresponds to a relative volume of the feasibility domain. The two notions are not equivalent, but still are closely related, which is seen here by comparing D to $\Omega^{2/S}$.

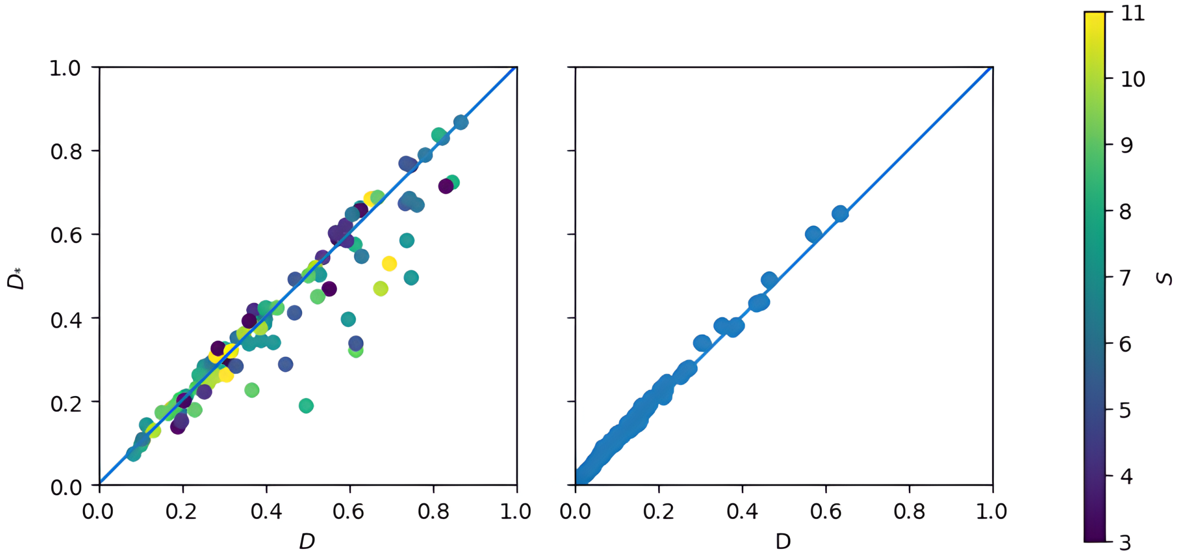


Interestingly, D and D_* are also very closely related. This is somewhat visible in their respective expressions, and confirmed numerically (see Fig. 2). This a useful thing to note because D_* is much simpler to compute, interpret, and manipulate than D, although it is the latter that is expected to driver of the major part of the function $p(\hat{z})$, at least when considering species-rich communities.

References

Barbier M, Mazancourt C, Loreau M, Bunin G (2021). *Fingerprints of High-Dimensional Coexistence in Complex Ecosystems*. *Physical Review X* **11**. Publisher: American Physical Society, 011009. `https://doi.org/10.1103/PhysRevX.11.011009`. URL: `https://link.aps.org/doi/10.1103/PhysRevX.11.011009` (visited on 07/18/2022).

Figure 2 – Correlation between robustness of coexistence D and characteristic distance D_*. Left panel shows this correlation for randomly generated matrices of variable size (S between 3 and 11). The right panel shows this correlation for matrices from real communities (S=4), based on the Barbier et al. (2021) dataset (see section "Application to data from grassland experiment" from the main article). The diagonal blue line corresponds to the x=y line in both cases. Note that D and D* are indeed closely related, for random matrices as for empirical matrices.



Appendix B
The distribution of distances to the edge of species coexistence
Persistence of species in simulated ecological systems

Mario Desallais, Michel Loreau, and Jean-François Arnoldi

In this appendix, we investigate the relationship between the metrics proposed in this study (See Eq. 10 and 11 of the main article) and actual species and community persistence. Although our main results consist in the distribution function of distances to the edge of species coexistence and the characterization of biotic roles at the species level in this coexistence, this verification remains essential to ensure the relevance of speaking about persistence, robustness or vulnerability of coexistence.

In the absence of an adequate dataset for this test, we employ numerical simulations using a generalized Lotka-Volterra model (See Eq. 1 of the main article) with random parameters. The simulation we used for this test follows this procedure : We generate n interaction matrices A of size $S \times S$, ensuring the are D-stable. We then sample a single set of growth rates (vector r). By construction, this defines n sets of species abundances (recall, $N^*(r) = A^{-1}r$) and therefore n communities. The coexistence and feasability of the latter depend on wheter the vector r is within the feasibilty domain defined by the interaction matrix A associated.

To simulate a changing environment, we randomly perturb the growth rate vector r, applying a small δr_i at each time step of the simulation. Consequently, at each time step, new species abundances N_i are defined. Coexistence occurs if r lies within the feasibility domain and all $N_i > 0$, otherwise it does not. When a community coexists for at least one time step, this defines a stable coexistence period. This period ends if, at any subsequent time step, any abundance falls to zero or below.

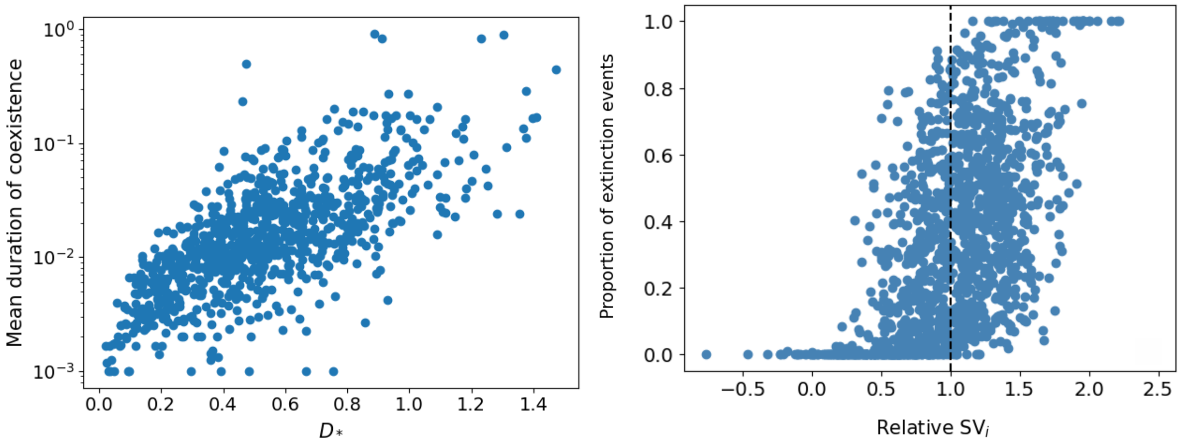
At the end of the simulation, we calculate the average duration of these "stable coexistence periods" for each community, over the total duration of the simulation. These values are represented on the Y axis of the left-hand panel in Figure 1 and we hypothesize that D_* is a good predictor of these values.

During the simulation, we also count each coexistence loss event for each community. Since each coexistence loss can be associated with a species (the one whose abundance falls to 0 or below), we calculate the proportion of extinction / coexistence loss events associated with each species in the community. This defines the Y axis of the right-hand panel in figure 1. We hypothesize that the value of SV_i for each species (relative to the mean in the community) predicts this proportion.

These two hypotheses are verified, the results of which are shown in Figure 1. A large value of D_* can allow a relatively long average duration of stable coexistence (although this does not guarantee it), while a small value of D_* forces a small average duration of coexistence. Regarding species extinctions, a low value of SV_i seems to guarantee a lower risk of extinction throughout the simulation, while a high value of SV_i are associated with a high proportion of extinction. This is particularly true for extreme values. The results obtained for SV_i values (relative to the community average) close to 1 are expected given the randomness of disturbances on δ_r and the fact that at SV_i close to 1, the species concerned is neither particularly favoured nor particularly damaged by other species.

The results obtained (1) are particularly sensitive to the parameterization of the simulation (especially, total time of the random walk, initial value of r_i, and value taken by the δr_i during the random walk), but the conclusions drawn above are consistent over the simulations. Nonetheless, these simulations should only be considered as a first check, and deserve to be further explored to provide more solid proof of the link between distances to the edges of the feasibility domain and persistence. In this sense, it would be very interesting to extend this study with a dataset from real experiment that quantified biotic interactions between species and the persistence of communities and species in the face of disturbance.

Figure 1 – Simulations-based testing of the link between community persistence (left panel) or species persistence (right panel) and measures of coexistence derived from the $p(z)$ function. Y-axis of left panel is displayed in log scale. Here, $n = 1000$, $S = 3$, non-diagonal elements of A are comprise between -1 and 1 and δr_i is drawn in a normal distribution with mean 0 and variance 1.



Appendix C

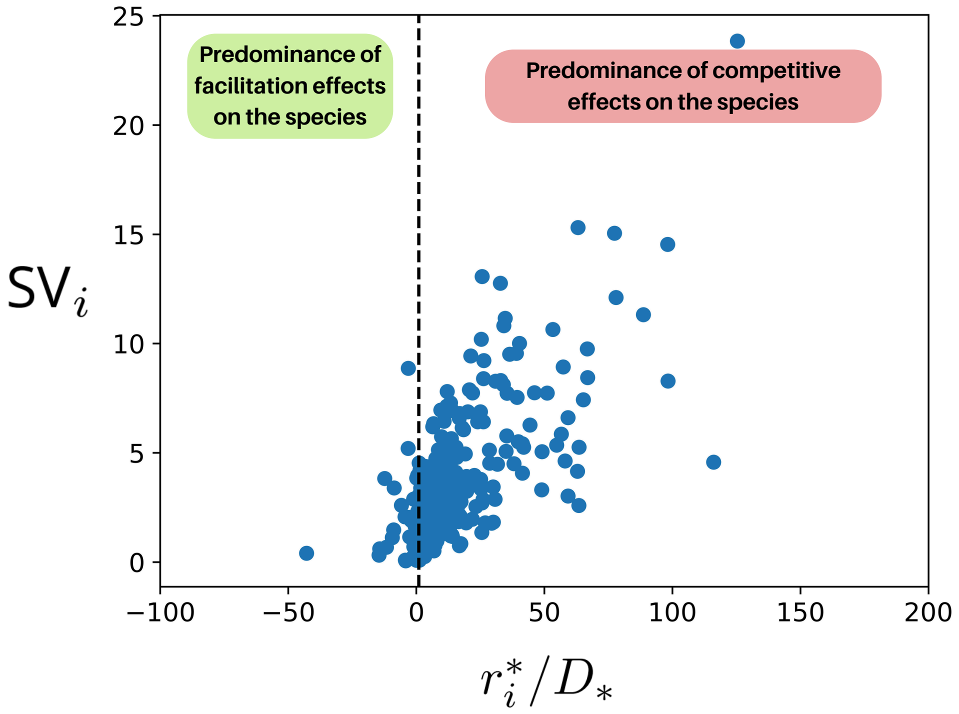
The distribution of distances to the edge of species coexistence

Role of absolute interaction strength in species vulnerability

Mario Desallais, Michel Loreau, and Jean-François Arnoldi

Although in the main article we used the SV_i and r_i^*/D_* measures relative to their average in the community, it is also possible to use and compare them in absolute terms. By doing this, a strong correlation between these two values can be observed (Fig. 1). This suggests that species that are constrained by others (through competition, highlighted in red in Fig. 1) are generally the ones that are mostly vulnerable. Conversely, species that tend to benefit from others (through facilitation, highlighted in green in Fig. 1) are those that are less vunerable.

Figure 1 – Correlation between the vulnerability of each species within a community (SV$_i$) and the effect of the biotic environment (interaction between species) on each (r_i^*/D_*). Each point represents one species within a community of 10 species (500 points in total). The vertical dot line corresponds to x=1, the qualitative threshold of the biotic effect on species. If this value is less than 1 (green box on the figure), this implies that the biotic environment is overall favorable (facilitating) to the concerned species. If upper than 1 (red box on the figure), it implies that the biotic environment is overall unfavorable through competition subjected to the species. Spearman rank correlation = 0.67 ; associated p-value : $2.5e^{-68}$



Supplementary material
The distribution of distances to the edge of species coexistence

Mario Desallais, Michel Loreau, and Jean-Francois Arnoldi

July 5, 2024

1 Established theory

Consider a Generalized Lotka-Volterra model written as

$$\frac{1}{N_i}\frac{dN_i}{dt} = r_i - \sum_{j=1}^{S} A_{ij}N_j;\ A_{ii} > 0;\ i = 1, ..., S$$

A growth rate vector r is feasible if the fixed point $N(r) = A^{-1}r$ is interior (all its components are positive). We do not worry about the stability of this state –additional conditions on the interaction matrix A (D-stability) can guarantee that any interior fixed point is automatically stable. Multiplying the vector r by some positive constant does not change feasibility: the latter is determined by the direction of r, not by its magnitude. The feasibility domain $D_f(A)$ is the subspace of all growth rate directions (a hyper-sphere) that are feasible, it is a finite geometrical object more precisely a convex polytope (but we will come back to this). Remarkably, we can easily compute the relative volume of this subspace, which defines a probability measure $\mathbb{P}(D_f) \leq 1$. For that, we note that drawing growth rate values r_i from independent standard normal distributions defines a uniform sampling of the directions of growth rate vectors. Integrating over all feasible configurations amounts to computing the gaussian integral and, after a change of variables, gives

$$\mathbb{P}(D_f) = \frac{1}{\sqrt{2\pi}^S}\int_{N(r)\in\mathbb{R}_+^S} e^{-\frac{||r||^2}{2}} d^S r = \frac{|A|}{\sqrt{2\pi}^S}\int_{\mathbb{R}_+^S} e^{-\frac{||AN||^2}{2}} d^S N$$

where we changed variables to integrate in $N-$space via $N = A^{-1}r$. Thus $\mathbb{P}(D_f)$ is simply the cumulative distribution noted $\Phi_{A^\top A}(0)$ of a multivariate normal distribution centered on 0 and with covariance matrix $C = (A^\top A)^{-1}$. In the absence of interactions $\mathbb{P}(D_f) = 2^{-S}$ so to focus on the effect of interactions it is convenient to define a ratio of probabilities, namely

$$\Omega(A^\top A) = \frac{\mathbb{P}(D_f(A))}{2^{-S}} = 2^S \times \Phi_{A^\top A}(0)$$

As previously mentioned, D_f is a convex polytope: the generalization of a triangle (with S vertices instead of 3, connected by $S(S-1)/2$ edges and drawn on the sphere instead of the plane). This

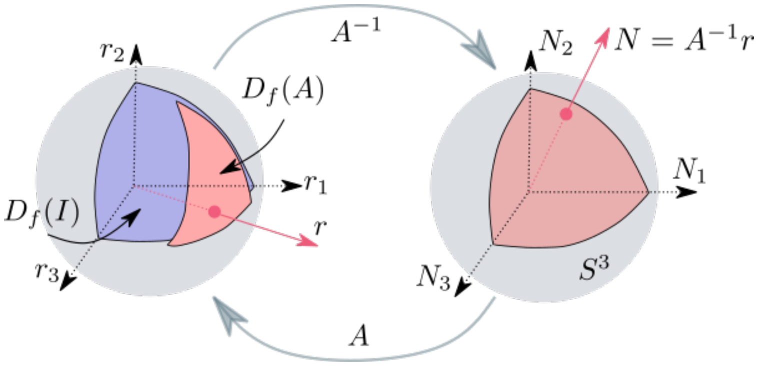


Figure 1: The feasibility domain $D_f(A)$ (in light red on the left) is defined as the subset of growth rate directions that, given a pair-wise interaction matrix A, allows coexistence between all species. It is the intersection of the sphere with the image in $r-$space (via the matrix A) of the positive quadrant in $N-$space (shown on the right). In the absence of interactions, the feasibility domain is the intersection of the positive quadrant and the sphere. The probability of feasibility $\mathbb{P}(D_f)$ is the ratio between the volume of D_f and the volume of that sphere.

can be understood by realizing that the column of A, $r^{(i)} = Ae^{(i)}$ where $e^{(i)} = (0, ..., 1, ..., 0)^\top$, defines the growth rate vector such that species i has unit abundance while all others are exactly at 0. Such points, when represented on the sphere are the S vertices of $D_f(A)$ and any feasible growth rate can be written as a positive linear combination of the extreme vectors $r^{(i)}$. In particular, the path $\phi r^{(i)} + (1-\phi)r^{(j)}$, where $1 \leq \phi \leq 1$, and $i \neq j$, once projected on the sphere draws an edge of the domain (only species i and j have non zero abundance).

These first results are nothing new, but will serve as an introduction to our subsequent analysis, where we want to go beyond the volume of D_f but also describe relevant features of its shape.

2 Motivation

We are interested in the fragility of coexistence, stated in the following way: If a growth rate vector is feasible, how easy is it to 'push it' out of the feasibility domain? This depends on how close to the boundary that vector was in the first place. To deduce a general statement about the interaction matrix we should instead ask: how likely is it that points in $D_f(A)$ are closer to the boundary than some threshold value z? What we will do is determine a *characteristic distance* D_* such that, given a prescribed distance z, z/D_* determines the proportion of points that lie at that distance from the edge of feasibility, and thus the distribution of distances, which completely characterizes both shape and size of the feasibility domain. Let us first see what such a question becomes in the abstract –but simpler– case of triangles drawn in the plane, which are the simplest polytopes.

All we need for now is basic trigonometry. The area of a triangle is $\mathcal{A} = \frac{1}{2}h \times \ell$ where ℓ is the length of a chosen edge L (the base), and h that of the segment orthogonal to L that connects L to a summit. If we choose a point in the triangle, what is the probability that it lies at a distance larger that z from an edge? To answer we need the area $\mathcal{A}'$ of a smaller triangle whose edges lie

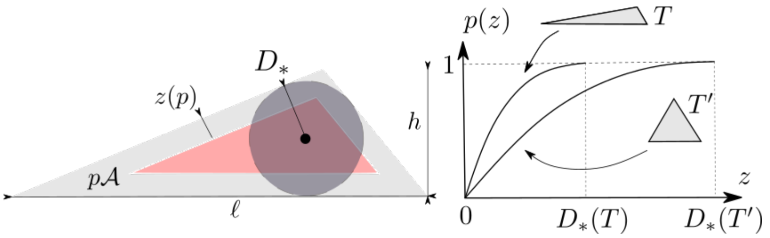


Figure 2: Triangles are parametrized by their base ℓ, height h and area $A = h\ell/2$. $\mathcal{D}$ controls the distribution of normalized distance $\hat{z} = 4.56 \times z/\sqrt{A}$ to the edge which gives the fraction of points p that are closer than some prescribed value z. $\hat{z}_{\text{max}} = 2\mathcal{D}$ is the radius of the largest circle that can be contained in the triangle (its center is *not* the barycenter of the triangle). The optimal shape is those of equilateral triangles for which $\mathcal{D} = 1$. Ω is the ratio of the area of the triangle divided by that of an equilateral triangle with the same perimeter. Its square root is equal to $\mathcal{D}$.

exactly at distance z from the borders of the original one (see Fig. 2). We leave it as an entertaining exercise to show that

$$\frac{\mathcal{A} - \mathcal{A}'}{\mathcal{A}} = 1 - \left(1 - \frac{z}{2\mathcal{A}/\mathcal{P}}\right)^2$$

where $\mathcal{P}$ is the perimeter of the original triangle. The ratio $2\mathcal{A}/\mathcal{P}$ has the dimension of a length, which we can call D_*. It is the radius of the largest disc that the triangle can contain. Its center is called the *incenter*. Thus the proportion $p(z)$ of points that are within distance z from an edge is entirely determined by the ratio z/D_*, since

$$\mathbb{P}(\text{distance} \leq z) = p(z) = 1 - (1 - \frac{z}{D_*})^2$$

The initial slope of $p(z)$ is $2/D_*$ since for $z/D_* \ll 1$, $p(z) \approx 2z/D_*$. The ratio $D_*/\sqrt{\mathcal{A}} = \sqrt{\mathcal{A}}/\mathcal{P}$ determines the relative size of the largest disc contained in the triangle. We can see this number as a shape factor, maximal for equilateral triangles.

3 Ecological theory

To build the ecological theory we need a notion of distance, defined in the space of growth rate vectors. For that we take a perturbative perspective. A perturbation leads to a shift of equilibrium abundance of species i

$$N_i(r + \delta r) - N_i(r) = \left\langle e_i, A^{-1}\delta r \right\rangle = \left\langle A^{-\top} e_i, \delta r \right\rangle = \left\langle v^{(i)}, \delta r \right\rangle$$

where $v^{(i)} = A^{-\top} e_i$ denotes the ith *row* of the inverse interaction matrix, which encodes that species sensitivity to environmental perturbations. Coexistence is lost as soon as one species goes extinct. This implies that $|\left\langle v^{(i)}, \delta r \right\rangle| = N_i(r)$ for some species i. The smallest intensity necessary, where intensity is measured as (we will see why shortly)

$$\text{intensity} = \sqrt{S}\frac{||\delta r||}{||r||}$$

gives us a notion of distance d to the edge of coexistence:

$$\begin{aligned} d = \text{distance}(r, \partial D_f) &= \min\left\{\text{intensity} \mid N_i(r + \delta r) = 0 \text{ for some } i\right\} \\ &= \min\left\{\sqrt{S}\tfrac{||\delta r||}{||r||} \mid |\left\langle v^{(i)}, \delta r\right\rangle| = N_i(r) \text{ for some } i\right\} \\ &= \min\left\{\sqrt{S}\tfrac{||\delta r||}{||r||} \mid ||v^{(i)}|| \times ||\delta r|| = N_i(r) \text{ for some } i\right\} \\ &= \min_i \tfrac{\sqrt{S}}{||r||}\tfrac{N_i(r)}{||v^{(i)}||} \end{aligned}$$

The normalization connects with the gaussian integral view point that allowed us to compute the relative volume of the feasibility domain. Indeed it amounts to enforce that that $||r||^2 = S$, which is indeed the expected norm of growth rate vectors whose components r_i are drawn from independent standard gaussian distributions. With this in hand we can easily compute the incenter r^* and its radius D_*. If we write

$$w_i = ||v^{(i)}||$$

representing a species maximal sensitivity, the point that lies at same distance D_* from all edges of feasibility satisfies

$$\frac{\sqrt{S}}{||r^*||} \times \frac{N_i(r^*)}{w_i} = D_*, \ \forall i$$

thus, if define the vector of species maximal sensitivities $w = (w_i)$, we have[1]

$$N(r^*) = A^{-1}r^* = \frac{||r^*||}{\sqrt{S}}D_* w \Leftrightarrow \sqrt{S}\frac{r^*}{||r^*||} = D_* A w$$

Note that r^* is indeed contained in the feasibility domain because, for all species, $N_i(r^*) = D_* w_i > 0$. All in all this gives us a remarkably simple expression for the maximal distance D_*:

$$D_* = \frac{\sqrt{S}}{||Aw||}$$

With this normalization, $D_* = 1$ in the absence of interactions. Note that

$$r_i^* = \sum_j A_{ij} w_j = \sum_j A_{ij} (A^\top A)_{jj}^{-1/2}$$

is a measure of how effectively *hostile,* the community –as a whole– is to species i. If $r_i^* = 1$, the community has a neutral effect, equal to that of the species on its own. If it is larger than 1, the community is overall hostile for that species. If it is less than 1, the community is benevolent for that species.

To understand edge effects, we now make a slight change of perspective, and go back to a probabilistic approach. Instead of asking growth rate vectors to be strictly normalized. What

---

[1] Note that

$$w_i^2 = \left\langle A^{-\top}e_i, A^{-\top}e_i\right\rangle = \left\langle A^{-\top}e_i, A^{-1}A^{-\top}e_i\right\rangle = \left\langle e_i, (A^\top A)^{-1}e_i\right\rangle$$

so w is the diagonal of $(A^\top A)^{-1/2}$.

we ask is that they are *normalized on average*, meaning that over their distribution, $\mathbb{E}||r||^2 = S$. That being said, recall that $\Phi_{A^\top A}(-x)$ is the gaussian cumulative function associated to A, and whose argument $x = (x_i)$ encodes the lower bounds of integration. The probability that growth rate vectors are feasible and further than a distance z amounts to computing the volume of the set of abundances such that $\{N_i \geq zw_i\}$ which is $\Phi_{A^\top A}(-zw)$, while the conditional probability is $\frac{\Phi_{A^\top A}(-zw)}{\Phi_{A^\top A}(0)}$. Thus the proportion $p(z) = \mathbb{P}(d \leq z)$ of points that are within distance z (measured as minimal perturbation intensity) from the edge of coexistence is

$$\mathbb{P}(d \leq z) = 1 - \frac{\Phi_{A^\top A}(-zw)}{\Phi_{A^\top A}(0)}$$

Now, computing the derivatives of the cumulative functions gives us, if $B = A^\top A$

$$\frac{d}{dz}|_{z=0}\mathbb{P}(d \leq z) = S\sqrt{\frac{2}{\pi}} \times \frac{1}{S}\sum_{i=1}^{S} w_i \sqrt{\frac{|A^\top A|}{|(A^\top A)_{/i}|}}\frac{\Omega_{/i}}{\Omega}$$

where $(A^\top A)_{/i}$ denotes the $(S-1) \times (S-1)$ matrix constructed by removing the ith column and row from the original matrix $A^\top A$, and where we identified the probability ratios $\Omega = 2^S \Phi_{A^\top A}(0)$ and $\Omega_{/i} = 2^{S-1}\Phi_{(A^\top A)_{/i}}(0)$. We deduce that in the absence of interactions $p'(0) = S\sqrt{\frac{2}{\pi}}$. We can define a charactersitic distance that accounts for pure dimensionality effects

$$1/D = \frac{1}{S}\sum_{i=1}^{S} w_i \sqrt{\frac{|A^\top A|}{|(A^\top A)_{/i}|}}\frac{\Omega_{/i}}{\Omega}$$

Thus we have the initial slope (and value, trivially 0) of $p(z)$. We also have that $p(z = D_*) = 1$. So we can get an ansatz that has the characteristics of a cumulative function, mimics the formula for usual triangles, while accounting for dimensionality effects:

$$\mathbb{P}(d \leq z) \approx 1 - \left(1 - \frac{z}{D_*}\right)^{S\sqrt{\frac{2}{\pi}}\frac{D_*}{D}}$$

4 Analytical expression in the mean field case

Consider the simplest non trivial interaction matrix

$$A = \begin{pmatrix} 1 & \mu/S & \dots & \mu/S \\ \mu/S & 1 & & \vdots \\ \vdots & & \ddots & \mu/S \\ \mu/S & \dots & \mu/S & 1 \end{pmatrix} ;\ -1 < \mu < S$$

We introduce some useful parameters:

$$\hat{\mu} = \frac{\mu}{1 - \mu/S};\ \text{and } a = \hat{\mu}(2 + \hat{\mu}) > -\frac{S}{S+1} > -1$$

and note that $1+a=(1+\hat{\mu})^2$. With these notations the relative volume of the feasibility domain is

$$\Omega \equiv \Omega(S,a) = \sqrt{1+a}\sqrt{\frac{2}{\pi}}^S \int_{\mathbb{R}_+^S} \exp\left(-\frac{||x||^2}{2} - \frac{a}{S}\frac{X^2}{2}\right) d^S x;\ X = \sum_{i=1}^S x_i$$

With the convention that $\sqrt{-1} = i$ we will show that

$$\Omega(S,a) = \frac{1}{\sqrt{2\pi}} \int_{\mathbb{R}} e^{-y^2/2} \mathrm{erfc}\left(i\sqrt{\frac{\lambda(a)}{S}}y\right)^S dy;\ \lambda(a) = \frac{1}{2}\frac{a}{1+a}$$

where $\mathrm{erfc}(z)$ is the complementary error function. The above can be approximated, as long as $\frac{\lambda}{S}$ is small enough, as

$$\Omega(S,a) \approx \begin{cases} \exp\left(-\frac{S}{\pi}\frac{a}{1+a}\right) & \text{if } a<0 \text{ (mutualism)} \\ \\ \sqrt{\frac{1+a}{1+Ca}}\exp\left(-\frac{S}{\pi}\frac{a}{1+Ca}\right) & \text{if } a \geq 0 \text{ (competition)} \end{cases};\ \text{where } C = \frac{\pi-2}{\pi} \approx 0.36$$

furthermore, we will show that

$$D_* = \frac{1}{\sqrt{1+a-a/S}};\ D = \frac{\Omega(S,a)}{\Omega(S-1,a-a/S)} \approx \begin{cases} \exp\left(-\frac{a}{\pi}\frac{2+a}{(1+a)^2}\right) & \text{if } a<0 \text{ (mutualism)} \\ \\ \exp\left(-\frac{a}{\pi}\frac{2+Ca}{(1+Ca)^2}\right) & \text{if } a \geq 0 \text{ (competition)} \end{cases}$$

This shows, in particular, that

$$\Omega^{2/S} \approx D$$

4.1 Preparation of $A^\top A$

$P_1 = \frac{|1\rangle\langle 1|}{S}$ the orthonormal projector on the diagonal (direction of total biomass), and $P_1^\perp$ the projector on the orthogonal hyperplane. The spectral decomposition of the mean field interaction matrix takes the form

$$A = (1-\mu/S)\left[P_1^\perp + (1+\hat{\mu})P_1\right]$$

so we can deduce that $|A| = (1-\mu/S)^S(1+\hat{\mu})$ and also

$$A^\top A = (1-\mu/S)^2\left[P_1^\perp + (1+\hat{\mu})^2 P_1\right] = (1-\mu/S)^2\left[\mathbb{I} + \hat{\mu}(2+\hat{\mu})\frac{|1\rangle\langle 1|}{S}\right]$$

We introduce a useful parameter

4.2 Competitive case:

Use the identity

$$\exp\left(-\frac{a}{S}\frac{X^2}{2}\right) = \sqrt{\frac{S}{2\pi a}}\int \exp\left(-\frac{1}{2}\left(S\frac{y^2}{a}+2iXy\right)\right)dy$$

to show that

$$\Omega = \sqrt{\frac{S}{2\pi}}\int_{\mathbb{R}}\sqrt{\frac{1+a}{a}}dy\exp\left(-\frac{S}{2}y^2\frac{1+a}{a}\right)\left(\sqrt{\frac{2}{\pi}}\int_0^{\infty}\exp\left(-\frac{(x+iy)^2}{2}\right)dx\right)^S$$

If we set $\lambda = \lambda(a) = \frac{1}{2}\frac{a}{1+a}$ after a change of variables we get

$$\Omega = \sqrt{\frac{S}{2\pi}}\int_{\mathbb{R}}dy\exp\left(-\frac{S}{2}y^2\right)\left(\sqrt{\frac{2}{\pi}}\int_0^{\infty}\exp\left(-\frac{(x+iy\sqrt{2\lambda})^2}{2}\right)dx\right)^S$$

By contour integration in the complex plane we get

$$\sqrt{\frac{2}{\pi}}\int_{i\sqrt{2\lambda}y}^{i\sqrt{2\lambda}y+\infty}\exp\left(-\frac{x^2}{2}\right)dx = 1-\sqrt{\frac{2}{\pi}}\int_0^{i\sqrt{2\lambda}y}\exp\left(-\frac{x^2}{2}\right)dx$$

and recognize the error function

$$\sqrt{\frac{2}{\pi}}\int_0^{z}\exp\left(-\frac{x^2}{2}\right)dx = \operatorname{erf}\left(\frac{z}{\sqrt{2}}\right)$$

as well as the imaginary error function

$$\operatorname{erfi}(y) = -i\operatorname{erf}(iy)$$

We finally get to an exact expression:

$$\Omega(S,a) = \frac{1}{\sqrt{2\pi}}\int_{\mathbb{R}}dy\exp\left(-\frac{y^2}{2}\right)\left(1-i\operatorname{erfi}\left(\sqrt{\frac{\lambda(a)}{S}}y\right)\right)^S \; ; \; \lambda(a) = \frac{1}{2}\frac{a}{1+a}$$

To get a simpler but approximate formula, we can linearize the error function near 0, which leads to

$$-i\operatorname{erfi}\left(\sqrt{\frac{\lambda}{S}}y\right) \approx -2i\sqrt{\frac{\lambda}{\pi S}}y$$

now $1-i\operatorname{erfi}\left(\sqrt{\frac{\lambda}{S}}y\right) = \rho e^{i\theta}$ with

$$\rho \approx \sqrt{1+\frac{4\lambda}{\pi S}y^2}; \; \theta \approx -2\sqrt{\frac{\lambda}{\pi S}}y - 2\pi k; \; k \in \mathbb{Z}$$

thus

$$S\log\left(1-i\operatorname{erfi}\left(\sqrt{\frac{\lambda}{S}}y\right)\right) \approx \frac{S}{2}\log\left(1+\frac{4\lambda}{\pi S}y^2\right) - i(2\sqrt{\frac{\lambda S}{\pi}}y+2\pi kS)$$

$$\approx \frac{2\lambda}{\pi}y^2 - i(2\sqrt{\frac{\lambda S}{\pi}}y + 2\pi kS)$$

So

$$\Omega \approx \frac{1}{\sqrt{2\pi}}\int_{\mathbb{R}} dy e^{-i2\pi kS} \exp\left(-\frac{1}{2}\left\{y^2(1-\frac{4\lambda}{\pi}) + 4i\sqrt{\frac{\lambda S}{\pi}}y\right\}\right)$$

$$\Omega \approx \exp\left(-S\frac{\frac{2\lambda}{\pi}}{1-\frac{4\lambda}{\pi}}\right)\frac{1}{\sqrt{2\pi}}\int_{\mathbb{R}} dy \exp\left(-\frac{1-\frac{4\lambda}{\pi}}{2}\left\{(y + i\frac{2\sqrt{\frac{S\lambda}{\pi}}}{1-\frac{4\lambda}{\pi}})^2\right\}\right)$$

$$\Omega(S,\lambda) \approx \frac{\exp\left(-S\frac{\frac{2\lambda}{\pi}}{1-\frac{4\lambda}{\pi}}\right)}{\sqrt{1-\frac{4\lambda}{\pi}}} = \sqrt{\frac{1+a}{1+Ca}}\exp\left(-S\frac{a/\pi}{1+Ca}\right)$$

4.3 Mutualistic case:

Use the identity

$$\exp\left(\frac{|a|}{S}\frac{X^2}{2}\right) = \sqrt{\frac{S}{2\pi|a|}}\int \exp\left(-\frac{1}{2}\left(S\frac{y^2}{|a|} + 2Xy\right)\right)dy$$

Using this

$$\Omega = \sqrt{\frac{1+a}{|a|}}\sqrt{\frac{S}{2\pi}}\int_{\mathbb{R}} dy \exp\left(-\frac{S}{2}\frac{y^2}{|a|}\right)\left(\sqrt{\frac{2}{\pi}}\int_0^{\infty}\exp\left(-\frac{(x+y)^2}{2} + \frac{y^2}{2}\right)dx\right)^S$$

$$= \sqrt{\frac{1+a}{|a|}}\sqrt{\frac{S}{2\pi}}\int_{\mathbb{R}} dy \exp\left(-\frac{S}{2}y^2\frac{1+a}{|a|}\right)\left(\sqrt{\frac{2}{\pi}}\int_0^{\infty}\exp\left(-\frac{(x+y)^2}{2}\right)dx\right)^S$$

if we set $\lambda = \frac{1}{2}\frac{|a|}{1+a}$ then

$$\Omega = \sqrt{\frac{S}{2\pi}}\int dy \exp\left(-\frac{S}{2}y^2\right)\left(\sqrt{\frac{2}{\pi}}\int_{\sqrt{2\lambda}y}^{+\infty}\exp\left(-\frac{x^2}{2}\right)dx\right)^S$$

We have that

$$\sqrt{\frac{2}{\pi}}\int_{\sqrt{2\lambda}y}^{+\infty}\exp\left(-\frac{x^2}{2}\right)dx = 1 - \sqrt{\frac{2}{\pi}}\int_0^{\sqrt{2\lambda}y}\exp\left(-\frac{x^2}{2}\right)dx$$

We can recognize the error function

$$\sqrt{\frac{2}{\pi}}\int_0^{z}\exp\left(-\frac{x^2}{2}\right)dx = \operatorname{erf}\left(\frac{z}{\sqrt{2}}\right)$$

and we finally get to an exact expression:

$$\Omega(S,a) = \frac{1}{\sqrt{2\pi}}\int_{\mathbb{R}} dy \exp\left(-\frac{y^2}{2}\right)\left(1 - \operatorname{erf}\left(\sqrt{\frac{\lambda(a)}{S}}y\right)\right)^S ; \lambda = \frac{1}{2}\frac{|a|}{1+a}$$

To get a simpler but approximate formula, we can linearize the error function near 0

$$-\mathrm{erf}\left(\sqrt{\frac{\lambda}{S}}y\right) \approx -2\sqrt{\frac{\lambda}{\pi S}}y$$

thus

$$S\log\left(1-\mathrm{erf}\left(\sqrt{\frac{\lambda}{S}}y\right)\right) \approx -2\sqrt{\frac{\lambda S}{\pi}}y$$

So

$$\Omega \approx \frac{\exp\left(S\frac{2\lambda}{\pi}\right)}{\sqrt{2\pi}}\int_{\mathbb{R}} dy \exp\left(-\frac{1}{2}\left\{(y+2\sqrt{\frac{\lambda S}{\pi}})^2\right\}\right) = \exp\left(S\frac{2\lambda}{\pi}\right) = \exp\left(-\frac{S}{\pi}\frac{a}{1+a}\right)$$

4.4 Computing D

Start from

$$\left\langle x, A^\top A_{/i}x\right\rangle_{S-1} = (1-\mu/S)^2\left(||x||^2_{S-1} + \left(\frac{S-1}{S}\right)a\frac{X^2_{S-1}}{S-1}\right)$$

from which we deduce the spectral decomposition of $A^\top A_{/i}$:

$$A^\top A_{/i} = (1-\mu/S)^2\left(\mathbb{I}_{S-1} + \left(\frac{S-1}{S}\right)a\frac{|1\rangle\langle 1|_{S-1}}{S-1}\right)$$

$$= (1-\mu/S)^2\left(P_1^\perp + (1+a(1-1/S))P_1^\perp\right)$$

Thus

$$\sqrt{|(A^\top A)_{/i}|} = (1-\mu/S)^{(S-1)}\sqrt{1+a(1-1/S)}$$

and so:

$$\Omega_{/i} = \sqrt{(1+a(1-1/S)}\sqrt{\frac{2}{\pi}}^{S-1}\int_{\mathbb{R}_+^{S-1}}\exp\left(-\frac{||x||^2}{2} - (1-1/S)a\frac{X^2}{S-1}\right)d^{S-1}x$$

So the same expression as Ω but with $S \to S-1$ and $a \to a_{/i} = a - a/S$ so

$$\Omega_{/i} = \Omega(S-1, a_{/i})$$

and then

$$D = \frac{\Omega}{\Omega_{/i}} \approx \exp\left(-\frac{a/\pi}{1+Ca}\right)\exp\left(\frac{S}{\pi}(\frac{a_{/i}}{1+Ca_{/i}} - \frac{a}{1+Ca})\right)$$

$$\frac{a_{/i}}{1+Ca_{/i}} = \frac{a}{1+Ca}\frac{1-1/S}{1-\frac{1}{S}\frac{Ca}{1+Ca}} \approx \frac{a}{1+Ca}\left(1-\frac{1/S}{1+Ca}\right)$$

so

$$S(\frac{a_{/i}}{1+Ca_{/i}} - \frac{a}{1+Ca}) \approx -\frac{a}{(1+Ca)^2}$$

and therefore

$$D \approx \exp\left(-\frac{a/\pi}{1+Ca}\left(\frac{2+Ca}{1+Ca}\right)\right) \approx \exp\left(-\frac{a}{\pi}\frac{2+Ca}{(1+Ca)^2}\right)$$

For mutualism the expression is simpler

$$\exp\left(-\frac{1}{\pi}\frac{a_{/i}}{1+a_{/i}}\right)\exp\left(-\frac{S}{\pi}(\frac{a}{1+a}-\frac{a_{/i}}{1+a_{/i}})\right)$$

$$\frac{a_{/i}}{1+a_{/i}} = \frac{a}{1+a}\frac{1-1/S}{1-\frac{1}{S}\frac{a}{1+a}} \approx \frac{a}{1+a}\left(1-\frac{1}{S}\frac{1}{1+a}\right)$$

$$D \approx \exp\left(-\frac{a}{\pi}\frac{2+a}{(1+a)^2}\right)$$