

Unifying mutualism diversity for interpretation and prediction

Feilun Wu¹, Allison J. Lopatkin¹, Daniel A. Needs¹, Charlotte T. Lee², Sayan Mukherjee³, Lingchong You¹,

4, 5*

1. Department of Biomedical Engineering, Duke University, Durham, North Carolina, 27708, USA
 2. Department of Biology, Duke University, Durham, North Carolina, 27708, USA
 3. Departments of Statistical Science, Mathematics, Computer Science, and Bioinformatics & Biostatistics, Duke University, Durham, North Carolina, 27708, USA
 4. Center for Genomic and Computational Biology, Duke University, Durham, North Carolina, 27708, USA
 5. Department of Molecular Genetics and Microbiology, Duke University School of Medicine, NC 27710, USA
- *. Corresponding author. Department of Biomedical Engineering, Duke University, CIEMAS 2355, 101 Science Drive, Box 3382, Durham, North Carolina, 27708, USA
- Tel.: +1 (919)660-8408; Fax: +1 (919)668-0795; E-mail: you@duke.edu

Coarse-grained rules are widely used in chemistry, physics and engineering. In biology, however, such rules are less common and under-appreciated. This gap can be attributed to the difficulty in establishing general rules to encompass the immense diversity and complexity of biological systems. Even when a rule is established, it is often challenging to map it to mechanistic details and to quantify these details. We here address these challenges on a study of mutualism, an essential type of ecological interaction in nature. Using an appropriate level of abstraction, we deduced a general rule that predicts the outcomes of mutualistic systems, including coexistence and productivity. We further developed a standardized calibration procedure to apply the rule to mutualistic systems without the need to fully elucidate or characterize their mechanistic underpinnings. Our approach consistently provides explanatory and predictive power with various simulated and experimental mutualistic systems. Our strategy can pave the way for establishing and implementing other simple rules for biological systems.

How populations interact to generate various outcomes is a key question in biology. Mutualism, where two or more populations provide reciprocal benefit, is an essential type of ecological interaction (1). In marine ecosystems, coral reefs are based on mutualistic interactions between coral and algae that provide ecosystem services for humans and habitats for diverse organisms (2). Plant-bacterial mutualism is estimated to generate 60% of the annual terrestrial nitrogen input (3). Cross-feeding in microbial communities is also a mutualistic interaction that influences community structures and is the cornerstone of various microbial metabolic tasks (4, 5). Although mutualistic coexistence is beneficial in maintaining the biodiversity, function and stability of ecosystems, under some conditions mutualistic systems can collapse, where one or more mutualistic partners is lost. This perturbation can further trigger the extinction or invasion of other populations and alter ecosystem functions (6-9). A framework to interpret and predict mutualistic outcomes is useful to prevent undesirable system behaviors and provide guidance for modulating and engineering synthetic mutualistic systems.

Quantitative rules have been developed to elevate our understanding and provide predictive power for various biological systems (10-13). However, such a framework is not yet available for determining mutualism outcomes. Main barriers in developing such a framework are the diversity of mutualistic interaction mechanisms and the complexity of underlying dynamics. Indeed, even engineered mutualistic systems that are by-design capable of cooperation, may not coexist. For example, it is still difficult to predict *a priori* whether an engineered microbial auxotrophic pair can persist or not (14-16). Previously, theoretical criteria in the form of model parameter inequalities have been developed for specific mutualistic systems such as cross-feeding mutualisms (17), plant-pollinator mutualisms (18), seed-dispersal mutualisms (19), ant-plant mutualisms (20), and plant-mycorrhizal mutualisms (21). These criteria depend on the underlying mechanisms *assumed* in the models and are not applicable to other types of mutualistic systems. General criteria have been developed (22-25), such as the classic criterion which states that intraspecific competition must be greater than mutual benefit for a mutualistic system to be stable (26). However, these criteria

usually describe transitions between stable coexistence and unbounded growth, and fail to address the transitions between coexistence and collapse and other mutualism dynamics (27, 28) (Fig. S1, Supplementary Text section I).

Results

Abstraction reveals a general rule

To reveal any commonality of mutualistic systems, we first summarized the logic of mutualism (Fig. 1a). Mutualism can be defined as the collective action of two or more populations, where each population produces benefit (β) that reduces the other's stress (δ) at a cost (ε) to itself. β , δ , and ε are universal features of mutualistic systems (Table 1). In addition to benefit and cost, which are conventionally considered as the driving forces of mutualistic outcomes (29-31), it has been shown that stress factors such as nutrient limitations (32), rising temperature (33, 34), rising CO₂ levels (35), invasive species (36) are also determining factors of mutualistic outcomes. Although stress is not always explicitly acknowledged in previous models, it has been implemented as decrease in growth rate (37), increase in linear turnover rate (32, 38-41), or increase in intraspecific competition (22, 25). Here we use stress as an umbrella term to capture the effects of stress factors in reducing baseline fitness of individual populations to provide a more complete picture of mutualism (see section II.A of Supplementary Text for the detailed reasoning).

To reflect the diversity of natural mutualistic systems, we systematically generated a total of 52 ordinary differential equation models based on the basic logic of mutualism with various implementation details (see Materials and Methods and section II.B of Supplementary Text for model assumptions). These implementation details are designed to comprehensively cover the various common and plausible forms of kinetic models that have been adopted in previous studies (see section II.C of Supplementary Text for a summary). Specifically, the models all revolve around the logistic growth equation but differ in the locations of β , ε , and δ and in the complexities the models capture, such as competition, partner-density-

dependent cost and asymmetry (see section II. D of Supplementary Text). We only increased the model complexity to an extent that closed-form steady state solutions are obtainable. We derived coexistence criteria for all 52 models and found that these criteria exhibit diverse structures (Fig. 1b; Table S1-S2; Supplementary MATLAB code). The diversity of our derived criteria is consistent with the diversity of criteria that already exist in the literature (17-21). This diversity highlights the need to have a general criterion, since the appropriate model formulation for a specific system is often unknown *a priori* and its selection can also be non-trivial (42).

Despite the diversity, at an appropriate abstraction level, however, all criteria follow a simple general form (Fig. 1c):

$$B(\boldsymbol{\theta}) > \delta \quad (\text{Eq. 1})$$

where $\boldsymbol{\theta}$ denotes model parameters accounting for β and ε . $B(\boldsymbol{\theta})$ represents the effective benefit produced through mutualistic interaction. Quantitatively, $B(\boldsymbol{\theta})$ increases with increasing β and decreasing ε and its structure differs depending on the specific model. δ represents stress; regardless of the model structure, it can always be quantified as $1 - r_m$, where r_m is the growth rate of the population in the absence of its mutualistic partner and normalized by its maximum growth rate. The interpretation of our criterion is intuitive: mutualistic partners can coexist if the effective benefit exceeds stress (Fig. 1d). Note that although alternative forms of the criterion may exist, Eq.1 is the most intuitive and parsimonious structure that decouples the effect of mutualistic interaction ($B(\boldsymbol{\theta})$) from the baseline stress level of single populations without interaction (δ).

Beyond determining qualitative system outcomes (coexistence versus collapse), B/δ defines a general metric that is positively correlated with quantitative mutualistic outcomes (Fig. 1e), such as final population density, probability of coexistence, and resistance to cheater exploitation. This quantitative predictive power of the metric is robustly maintained for both symmetric and asymmetric systems,

including obligate and facultative mutualistic systems (Fig. S2a-f; Supplementary Text section IV). Further, the prediction accuracy of our criterion for qualitative outcomes is also robustly maintained in the presence of noise (Fig. S2g). The generality of the metric indicates that it is a property of mutualism. If so, B is a high-level feature that, along with δ , provides a unifying framework for interpreting and predicting diverse mutualistic systems.

A calibration approach to use the metric

Quantification of both B and δ are required to use the metric. Although δ is easy to measure by the growth of single populations using a model-independent approach (see Fig. S3 for the general procedure), quantification of B , which describes the interactions, can often be challenging. Beyond the difficulties of selecting an appropriate structure of $B(\theta)$, quantification of its underlying parameters often requires non-trivial mechanistic characterizations, such as model parameter fitting and specific biochemical assays. These mechanistic characterizations are especially challenging for cooperative traits, even in well-defined synthetic systems (43-45). Applications of the criterion would thus be difficult for individual systems, let alone enabling streamlined applications for diverse mutualistic systems.

To bypass these challenges, we developed a calibration procedure to use qualitative outcomes to directly quantify B as an empirical function of system context \mathbf{v} , denoted by $B(\mathbf{v})$ (Fig. 2a). \mathbf{v} is comprised of system variables that modulate system behaviors, such as temperature, nutrient availability and genetic variation. \mathbf{v} measurements are often readily available, especially in laboratory settings where they are experimentally controlled independent variables. Thus, using simple measurements, we can approximate the true $B(\theta(\mathbf{v}))$ that describes the diverse and complex interaction mechanisms without characterizing the specific mechanistic details. $B(\mathbf{v})$ along with δ , will serve as the basis for interpretation and prediction beyond initial data. Although the procedure requires initial measurements of *qualitative* outcomes, $B(\mathbf{v})/\delta$ can also provide predictive power for *quantitative* outcomes (Fig. 1e). $B(\mathbf{v})/\delta$ is positively related to the

final density (Fig. 1f; Fig. S2c-d), probability of coexistence (Fig. S2a, e), and cheater resistance (Fig. S2b, f). Further, $B(\mathbf{v})$ can be used to reveal how multiple system variables collectively alter the effectiveness of the interaction, which is a major challenge in studying context dependency of mutualistic outcomes (46).

We first defined the input-output relationship of the calibration procedure (Fig. 2b). Measurements of qualitative outcomes are denoted as $\mathbf{Y} = [y_1, y_2, y_3, \dots, y_n]$ ($y_i = 1$ for coexistence and -1 for collapse; i represents the index of an observation; n represents the total number of observations). Measurements of δ for the same set of observations are denoted as $\boldsymbol{\delta} = [\delta_1, \delta_2, \delta_3, \dots, \delta_n]$. Note that theoretically, quantification of δ for any partner is sufficient. However, choosing the partner with a larger dynamic range of δ is preferable since it can contain more information content. System context is denoted by $\mathbf{v} = [\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3, \dots, \mathbf{v}_n]$, where \mathbf{v}_i is a vector that contains the values of all system variables for observation i . With inputs \mathbf{Y} , $\boldsymbol{\delta}$ and \mathbf{v} , we can establish a smooth boundary between coexistence and collapse described by $F(\boldsymbol{\delta}, \mathbf{v}) = 0$. To ensure $B > \delta$ for coexistence and $B < \delta$ for collapse, we constrain $F(\boldsymbol{\delta}, \mathbf{v}) > 0$ for coexistence and $F(\boldsymbol{\delta}, \mathbf{v}) < 0$ for collapse. Because $B = \delta$ is true at the boundary, we can deduce that $F(B, \mathbf{v}) = 0$. According to the implicit function theorem, if $F(B, \mathbf{v}) = 0$ is continuously differentiable, the output $B(\mathbf{v})$ is implied. A calibrated $B(\mathbf{v})$ can then enable downstream interpretation and prediction.

To implement the calibration, we used support vector machine (SVM), a machine-learning algorithm for supervised classification (see Supplementary MATLAB code of the implementation). Assuming continuity of $B(\mathbf{v})$, we designed kernels that are separable in $\boldsymbol{\delta}$ and \mathbf{v} to train the function $F(\boldsymbol{\delta}, \mathbf{v}) = 0$. We implemented linear, quadratic, cubic and sigmoidal kernels to describe different possible shapes of $B(\mathbf{v})$. Because there are infinite number of $B(\mathbf{v})$ that can provide equivalently high classification accuracy, we ranked the $B(\mathbf{v})$ obtained from different kernel parameters to find the $B(\mathbf{v})$ that are closer to the true $B(\boldsymbol{\theta}(\mathbf{v}))$ (Fig. S4a). The ranking method is established using simulated data where the true $B(\boldsymbol{\theta}(\mathbf{v}))$ is known, so that each $B(\mathbf{v})$ can be evaluated against it by coefficient of determination (R^2). We

found that our procedure consistently optimizes for R^2 (Fig. S4b-c; Supplementary Text section V). The proper sample size for the calibration can be evaluated based on the exponential decay of bias with increasing sample size (47) (Fig. S4d).

Using this procedure, we first tested whether $B(\mathbf{v})/\delta$ can be applied to mutualism models in which no explicit form exists for $B(\boldsymbol{\theta})$. To do so, we constructed an overwhelmingly complex model with competition, partner-density-dependent cost, high Hill coefficient and asymmetric function structures (see Supplementary Text section V.F, Fig S5a). Using an input data set of 100 points, $B(\mathbf{v})/\delta$ correctly predicts coexistence versus collapse for 97.2% of test data beyond the initial 100 data points. As expected, $B(\mathbf{v})/\delta$ provides predictive power for quantitative outcomes including total population size (Fig. 2c), probability of coexistence (Fig. S5b) and resistance to cheater exploitation (Fig. S5c). The confidence of $B(\mathbf{v})$ can be evaluated by the consistency of top $B(\mathbf{v})$ and their relative standard deviation (RSD) (Fig. 5 d, e).

Experimental application of the metric

We next applied our framework to three experimental systems to test its applicability. As the first example, we engineered two synthetic mutualistic partners in Top10F' strain of *Escherichia coli*, denoted by M_1 and M_2 (Fig. 3a, Fig. S6a). In this system, stress is modulated by the concentration of Isopropyl β -D-1-thiogalactopyranoside (IPTG) which induces the expression of CcdB (a toxin). Anhydrotetracycline (aTc) induces quorum sensing (QS) modules in both strains to each produce a unique QS signal that triggers the production of CcdA (the antitoxin of CcdB) in the partner population. The production of aTc-induced expression of the QS module can impose cooperation cost to both strains. Consistent with circuit design, our experimental results demonstrated IPTG-mediated growth suppression and aTc-mediated mutual rescue (Fig. S6b).

We cocultured the two strains starting from the same initial density with different concentrations of IPTG and aTc, which are the two dimensions of \mathbf{v} . The outcomes of coexistence and collapse are evident in the bimodal distribution of optical density (OD) at 32 hours of culturing (Fig. S6c). δ can be quantified by treating monocultures with the same set of [IPTG] and [aTc]. We used δ for M_2 since it has a wider dynamic range (Fig. S6d). Using these data (Fig. 3b), we obtained a calibrated $B(\mathbf{v})$. The confidence of $B(\mathbf{v})$ is evaluated by the consistency of the top 5 $B(\mathbf{v})$ and relative standard deviation of each $B(\mathbf{v})$ (Fig. S6e). Consistent with the circuit logic, $B(\mathbf{v})$ increases with increasing [aTc]. The calibration reveals that [IPTG] also modulates $B(\mathbf{v})$, which indicates unintended system complexities, such as QS cross-talk and unequal fitness of the two populations (Fig. 3c). We used cross-validation to evaluate how well new observations can be predicted. We found that $B(\mathbf{v})/\delta$ provides a cross validation accuracy of 96.8% for coexistence versus collapse and it is also predictive of total final density (Fig. 3d).

We then applied our procedure to data on a pair of *Saccharomyces cerevisiae* auxotrophs that is previously published (32). In this system, one strain cannot produce tryptophan (Trp) and the other cannot produce leucine (Leu). The mutualistic interaction of this system is realized by the exchange of the two amino acids in cocultures (Fig. 3e). Because [Leu] is maintained as 8 times of [Trp], we use [Trp] as one of the dimensions of \mathbf{v} to represent overall concentration of supplemented amino acid. The ratio of initial densities of the two strains composes the other dimension of \mathbf{v} (Fig. 3f, Fig. S7a). All top 5 $B(\mathbf{v})$ reveal that intermediate ratios of initial density and increasing amino acid concentrations elevate $B(\mathbf{v})$ (Fig. 3g). However, at the highest level of supplemented amino acid ([Trp] = 16nM, [Leu] = 128nM), top-ranked $B(\mathbf{v})$ have qualitatively different trends, indicating a low confidence of $B(\mathbf{v})$ at high concentration (Fig. S7b). Interestingly, this high variability coincides with the system transitioning into competitive exclusion (32). Nevertheless, $B(\mathbf{v})/\delta$ is still predictive of final densities (Fig. 3h). Furthermore, we explored using the concentration of supplemented amino acid as a single system variable. $B(\mathbf{v})/\delta$ in this case can also predict the probability of coexistence as the ratio of initial densities varied (Fig. S7c).

In the third example, we applied our framework to previously published measurements of 14 engineered auxotrophic *E. coli* strains that compose 91 pairwise mutualistic systems (48) (Fig. 3i). The genetic context of the two partners vary while the growth environment is kept the same. The classification of coexistence versus collapse is based on the bimodal distribution of total density (Fig. 8a). δ of each auxotroph is determined based on final cell densities of monocultures when supplemented with different concentrations of its corresponding amino acid (Fig. S8b). We sorted the auxotrophs by the number of partners they coexist with to convert categorical indices into an ordinal scale. Thus, \mathbf{v} is composed of ordinal rankings of the two strains and measurements of coexistence versus collapse and δ are both arranged accordingly (Fig. 3j). We used strain 1 as the reference strain for the calibration. The calibrated $B(\mathbf{v})$ generated a cross validation accuracy of 91.8% and we verified that $B(\mathbf{v})/\delta$ is predictive of final total density (Fig. 3k-l, Fig. S8c). We noticed a relatively high level of variability of total density when $B(\mathbf{v})/\delta > 1$, which can be due to system-specific properties that are not fully accounted for by mutualistic interactions.

Generalization of the framework

In nature, mutualism can occur among three or more partners (49). Thus, we tested our framework with simulations and experimental measurements of N-mutualist systems. Here we show the calibration procedure with simulated data from a 5-mutualist system and found that the quality of the calibration results is well-maintained (Fig. 4a; Fig. S9a; Supplementary Text section VI.A). The same 14 auxotrophs (48) presented above were also used to construct all possible three-member systems. Using the same procedure with a 3-dimensional \mathbf{v} , where each dimension represents the genetic context of one strain, we found the predictor $B(\mathbf{v})/\delta$ provides an 89.3% cross validation accuracy and remains predictive of the total density, which indicates the scalability of the framework in experimental settings (Fig. 4b, Supplementary Text section VI.B). We hypothesized that $B(\mathbf{v})$ calibrated for pairwise interactions can be used to directly

construct a metric for 3-member systems, since theoretical analysis shows that n -member $B(\theta)$ can be approximated by pairwise $B(\theta)$ (Table S2). We assume B of a 3-member system is the average of B for all 3 combinations of its underlying 2-population systems and the same is true for δ . The constructed B/δ for 3-member systems explains 80.8% of system outcomes (Fig. S9b). This result suggests the possibility of directly extending B and δ from simple systems to more complex systems without further calibration.

Beside static environments, mutualistic systems can also inhabit dynamic environments where they experience fluctuating physical and chemical cues or cohabitate with other populations. We verified that the theoretical criterion generally holds in both cases (Fig. S10a). However, the transition between collapse and coexistence does not strictly occur at 1, which further advocates for the necessity of the calibration process. With simulated data, we carried out the calibration procedure and verified that the applicability of our framework is well-maintained (Fig. 4c-d, Fig. S10b, Supplementary Text section VI.C-D). The robustness of the framework suggests that it can be used to study microbial communities, of which advancements in both interpretation and prediction are in demand (50).

Discussion

The immense complexity and diversity of biological systems is intriguing and inspires the exploration of mechanistic details. However, these details can distract us from simple rules that emerge at a higher level. By abstracting away from low-level details, many simple rules for biological systems have been developed to enhance our understanding and provide predictive power (51, 52). A classic example is the Hamilton's rule, which states that a cooperative trait will persist if $c/b > r$, where r is the relatedness of the recipient and the actor; b is the benefit gained by the recipient; and c is the cost to the actor. More recent examples include linear correlations underlying cell-size homeostasis in bacteria (53-55), ranking of quorum sensing modules according to their sensing potential (56, 57), and the growth laws resulting from dynamic partitioning of intracellular resources (58, 59). Beyond establishing another simple rule, we also

demonstrated that one can purposefully seek an appropriate abstraction level where a simple unifying rule emerges over system diversity. If this rule anchors in the basic definition of a type of system, it can then be applied to diverse systems of the same type. Beyond microbial systems that we tested, our criterion in principle can also be applied to other systems of larger or smaller scales that share the same logic.

Although simple general rules in biology are powerful tools, their applicability to experimental systems can be limited by the difficulties in associating the abstracted parameters to lower-level mechanistic details and quantifying these details experimentally. This is evident in the application of Hamilton's rule to, for example, to experimental systems (43-45). For simple rules that are dictated by inequalities, our calibration procedure provides a tool to apply these rules directly to experimental systems. If one side of the inequality and some final outcomes can be measured, the other side can be calibrated as an empirical function. Although the procedure cannot further dissect the empirical function into specific mechanistic parameters, the function can serve as an overall summary of the underlying mechanistic details while bypassing the requirement of characterizing them individually. Our approach thus enables the downstream interpretation and prediction by these simple rules with readily-accessible measurements.

References

1. D. H. Boucher, S. James, K. H. Keeler, The Ecology of Mutualism. *Annu Rev Ecol Syst* **13**, 315-347 (1982).
2. F. Moberg, C. Folke, Ecological goods and services of coral reef ecosystems. *Ecol Econ* **29**, 215-233 (1999).
3. H. H. Zahran, Rhizobium-legume symbiosis and nitrogen fixation under severe conditions and in an arid climate. *Microbiol Mol Biol R* **63**, 968-+ (1999).
4. S. Stolyar *et al.*, Metabolic modeling of a mutualistic microbial community. *Molecular Systems Biology* **3**, (2007).
5. J. R. Sieber, M. J. McInerney, R. P. Gunsalus, Genomic Insights into Syntrophy: The Paradigm for Anaerobic Metabolic Cooperation. *Annu Rev Microbiol* **66**, 429-452 (2012).
6. T. M. Palmer *et al.*, Breakdown of an ant-plant mutualism follows the loss of large herbivores from an African Savanna. *Science* **319**, 192-195 (2008).
7. C. E. Christian, Consequences of a biological invasion reveal the importance of mutualism for plant communities. *Nature* **413**, 635-639 (2001).
8. U. Bastolla *et al.*, The architecture of mutualistic networks minimizes competition and increases biodiversity. *Nature* **458**, 1018-1020 (2009).
9. C. E. Aslan, E. S. Zavaleta, B. Tershy, D. Croll, Mutualism Disruption Threatens Global Plant Biodiversity: A Systematic Review. *PloS one* **8**, (2013).
10. W. D. Hamilton, The genetical evolution of social behaviour. I. *Journal of theoretical biology* **7**, 1-16 (1964).
11. H. Ohtsuki, C. Hauert, E. Lieberman, M. A. Nowak, A simple rule for the evolution of cooperation on graphs and social networks. *Nature* **441**, 502-505 (2006).
12. M. A. Nowak, Five rules for the evolution of cooperation. *Science* **314**, 1560-1563 (2006).
13. M. Scott, T. Hwa, Bacterial growth laws and their applications. *Curr Opin Biotech* **22**, 559-565 (2011).
14. M. T. Mee, H. H. Wang, Engineering ecosystems and synthetic ecologies. *Mol Biosyst* **8**, 2470-2483 (2012).
15. E. H. Wintermute, P. A. Silver, Emergent cooperation in microbial metabolism. *Molecular Systems Biology* **6**, (2010).
16. W. Shou, S. Ram, J. M. Vilar, Synthetic cooperation in engineered yeast populations. *Proceedings of the National Academy of Sciences of the United States of America* **104**, 1877-1882 (2007).
17. J. S. Meyer, H. M. Tsuchiya, Dynamics of Mixed Populations Having Complementary Metabolism. *Biotechnology and bioengineering* **17**, 1065-1081 (1975).
18. T. A. Revilla, F. Encinas-Viso, Dynamical Transitions in a Pollination-Herbivory Interaction: A Conflict between Mutualism and Antagonism. *PloS one* **10**, (2015).
19. Y. Harada, Y. Iwasa, Lattice Population-Dynamics for Plants with Dispersing Seeds and Vegetative Propagation. *Res Popul Ecol* **36**, 237-249 (1994).
20. B. Rai, H. I. Freedman, J. F. Addicott, Analysis of 3 Species Models of Mutualism in Predator-Prey and Competitive-Systems. *Math Biosci* **65**, 13-50 (1983).
21. J. D. Hoeksema, M. Kummel, Ecological persistence of the plant-mycorrhizal mutualism: A hypothesis from species coexistence theory. *Am Nat* **162**, S40-S50 (2003).
22. S. Allesina, S. Tang, Stability criteria for complex ecosystems. *Nature* **483**, 205-208 (2012).
23. K. Z. Coyte, J. Schluter, K. R. Foster, The ecology of the microbiome: Networks, competition, and stability. *Science* **350**, 663-666 (2015).
24. R. M. May, Citation Classic - Stability and Complexity in Model-Ecosystems. *Cc/Agr Biol Environ*, 22-22 (1988).

25. J. Bascompte, P. Jordano, J. M. Olesen, Asymmetric coevolutionary networks facilitate biodiversity maintenance. *Science* **312**, 431-433 (2006).
26. B. S. Goh, Stability in Models of Mutualism. *Am Nat* **113**, 261-275 (1979).
27. J. F. Addicott, Stability properties of 2-species models of mutualism: Simulation studies. *Oecologia* **49**, 42-49 (1981).
28. B. Momeni, L. Xie, W. Shou, Lotka-Volterra pairwise modeling fails to capture diverse pairwise microbial interactions. *eLIFE*, (2017).
29. D. H. Boucher, *The Biology of mutualism : ecology and evolution*. (Oxford University Press, New York, 1985), pp. x, 388 p.
30. W. F. Morris, D. P. Vazquez, N. P. Chacoff, Benefit and cost curves for typical pollination mutualisms. *Ecology* **91**, 1276-1285 (2010).
31. K. H. Keeler, in *The biology of mutualism, ecology and evolution*. (1985), pp. 100-127.
32. T. A. Hoek *et al.*, Resource Availability Modulates the Cooperative and Competitive Nature of a Microbial Cross-Feeding Mutualism. *PLoS biology* **14**, (2016).
33. B. E. Brown, Coral bleaching: causes and consequences. *Coral Reefs* **16**, S129-S138 (1997).
34. N. S. Webster, R. E. Cobb, A. P. Negri, Temperature thresholds for bacterial symbiosis with a sponge. *Isme Journal* **2**, 830-842 (2008).
35. M. Holcomb, D. C. McCorkle, A. L. Cohen, Long-term effects of nutrient and CO₂ enrichment on the temperate coral *Astrangia poculata* (Ellis and Solander, 1786). *J Exp Mar Biol Ecol* **386**, 27-33 (2010).
36. D. Cipollini, C. M. Rigsby, E. K. Barto, Microbes as Targets and Mediators of Allelopathy in Plants. *J Chem Ecol* **38**, 714-727 (2012).
37. J. P. Grime, Evidence for Existence of Three Primary Strategies in Plants and Its Relevance to Ecological and Evolutionary Theory. *Am Nat* **111**, 1169-1194 (1977).
38. F. K. Balagadde *et al.*, A synthetic *Escherichia coli* predator-prey ecosystem. *Molecular Systems Biology* **4**, (2008).
39. C. Neuhauser, J. E. Fargione, A mutualism-parasitism continuum model and its application to plant-mycorrhizae interactions. *Ecol Model* **177**, 337-352 (2004).
40. N. Yamamura, Evolution of mutualistic symbiosis: A differential equation model. *Res Popul Ecol* **38**, 211-218 (1996).
41. W. F. Morris, J. L. Bronstein, W. G. Wilson, Three-way coexistence in obligate mutualist-exploiter interactions: The potential role of competition. *Am Nat* **161**, 860-875 (2003).
42. J. B. Johnson, K. S. Omland, Model selection in ecology and evolution. *Trends in ecology & evolution* **19**, 101-108 (2004).
43. J. S. Chuang, O. Rivoire, S. Leibler, Cooperation and Hamilton's rule in a simple synthetic microbial system. *Mol Syst Biol* **6**, 398 (2010).
44. J. Smith, J. D. Van Dyken, P. C. Zee, A Generalization of Hamilton's Rule for the Evolution of Microbial Cooperation. *Science* **328**, 1700-1703 (2010).
45. J. B. Xavier, Social interaction in synthetic and natural microbial communities. *Mol Syst Biol* **7**, 483 (2011).
46. J. L. Bronstein, *Mutualism*. (Oxford University Press, Oxford, United Kingdom, ed. First edition., 2015), pp. xv, 297 pages, 298 unnumbered pages of plates.
47. S. Mukherjee *et al.*, Estimating dataset size requirements for classifying DNA microarray data. *J Comput Biol* **10**, 119-142 (2003).
48. M. T. Mee, J. J. Collins, G. M. Church, H. H. Wang, Syntrophic exchange in synthetic microbial communities. *Proceedings of the National Academy of Sciences of the United States of America* **111**, E2149-2156 (2014).
49. M. L. Stanton, Interacting guilds: Moving beyond the pairwise perspective on mutualisms. *Am Nat* **162**, S10-S23 (2003).

50. S. Widder *et al.*, Challenges in microbial ecology: building predictive understanding of community function and dynamics. *ISME J* **10**, 2557-2568 (2016).
51. S. Klumpp, T. Hwa, Bacterial growth: global effects on gene expression, growth feedback and proteome partition. *Curr Opin Biotech* **28**, 96-102 (2014).
52. C. D. Nadell *et al.*, Cutting through the complexity of cell collectives. *P Roy Soc B-Biol Sci* **280**, (2013).
53. Y. Tanouchi, A. Pai, N. E. Buchler, L. You, Programming stress-induced altruistic death in engineered bacteria. *Mol Syst Biol* **8**, 626 (2012).
54. M. Campos *et al.*, A Constant Size Extension Drives Bacterial Cell Size Homeostasis. *Cell* **159**, 1433-1446 (2014).
55. S. Taheri-Araghi *et al.*, Cell-Size Control and Homeostasis in Bacteria. *Current Biology* **25**, 385-391 (2015).
56. A. Pai, L. C. You, Optimal tuning of bacterial sensing potential. *Molecular Systems Biology* **5**, (2009).
57. A. Pai, Y. Tanouchi, L. C. You, Optimality and robustness in quorum sensing (QS)-mediated regulation of a costly public good enzyme. *Proceedings of the National Academy of Sciences of the United States of America* **109**, 19810-19815 (2012).
58. M. Scott, C. W. Gunderson, E. M. Mateescu, Z. G. Zhang, T. Hwa, Interdependence of Cell Growth and Gene Expression: Origins and Consequences. *Science* **330**, 1099-1102 (2010).
59. M. Scott, S. Klumpp, E. M. Mateescu, T. Hwa, Emergence of robust growth laws from optimal regulation of ribosome synthesis. *Molecular Systems Biology* **10**, (2014).
60. C. E. Christian, M. L. Stanton, Cryptic consequences of a dispersal mutualism: Seed burial, elaiosome removal, and seed-bank dynamics. *Ecology* **85**, 1101-1110 (2004).
61. G. H. Pyke, What Does It Cost a Plant to Produce Floral Nectar. *Nature* **350**, 58-59 (1991).
62. R. Axelrod, D. E. Axelrod, K. J. Pienta, Evolution of cooperation among tumor cells. *Proceedings of the National Academy of Sciences of the United States of America* **103**, 13474-13479 (2006).
63. L. Muscatine, J. W. Porter, Reef Corals - Mutualistic Symbioses Adapted to Nutrient-Poor Environments. *Bioscience* **27**, 454-460 (1977).
64. L. J. McCook, J. Jompa, G. Diaz-Pulido, Competition between corals and algae on coral reefs: a review of evidence and mechanisms. *Coral Reefs* **19**, 400-417 (2001).
65. S. A. Wooldridge, Is the coral-algae symbiosis really 'mutually beneficial' for the partners? *Bioessays* **32**, 615-625 (2010).

Acknowledgements

We thank Tim Hoek and Jeff Gore for providing data used for analysis in Figure 3f-h and Michael Mee and Harris Wang for sharing the data used for analysis in Figure 3j-l. We also thank Yu Tanouchi, Lawrence David, Wenying Shou, Nan Luo, Yangxiaolu Cao, Carolyn Zhang, Ryan Tsoi, Teng Wang, Shangying Wang for constructive inputs.

Funding

This work is partially supported by grants from US National Institutes of Health, National Science Foundation, Office of Naval Research, Army Research Office, and the David and Lucile Packard Foundation.

Author contributions

F.W. conceived the research, designed and performed both modelling and experiments, interpreted the results, and wrote the manuscript. A.J.L. constructed the synthetic circuit, assisted with experimental design and manuscript revisions. D.N. assisted with modeling, results interpretation, and manuscript revisions. C.L. assisted with establishing the general relevance of the criterion and manuscript revisions. S.M. assisted with establishing the calibration process and manuscript revisions. L.Y. conceived the research, assisted in research design, data interpretation, and wrote the manuscript. All authors approved the manuscript.

Competing interests

The authors declare no competing financial interests.

Data and materials availability

All data is available in the manuscript or the supplementary materials.

Tables

Category by function	Mutualistic partners	Benefit	Cost	Stress
Transportation mutualism	Plants	Increased fecundity (60)	Seed consumption and energy loss (60, 61)	Limited spatial range for reproduction
	Seed dispersers or pollinators	Access to nutrient-rich food	Energy loss or by-product mutualism	Starvation
Protection mutualism	Cancer cells	Increased fitness and biochemical cues for proliferation (62)	Energy loss or by-product mutualism	Host defense such as immune response
	Stromal cells	Shared growth factors and altered microenvironments (62)	Energy loss or by-product mutualism	Growth-suppressing factors in the host environment
Nutritional mutualism	Bacterial and archaeal auxotrophs	Increased nutrient availability in the environment	Energy loss (or byproduct mutualism)	Nutrient-poor environments
Nutritional mutualism	Corals	Higher rate of calcification and conservation of nutrient (63)	Reduced cover, growth and fecundity (64)	Nutrient-poor marine environment
	Algae	Better habitat and increased availability of inorganic compound (65)	Energy loss, possible restricted growth by coral (65)	Nutrient-poor marine environment

Table 1: Examples of benefit, cost, and stress in diverse mutualistic systems

Figures and figure legends

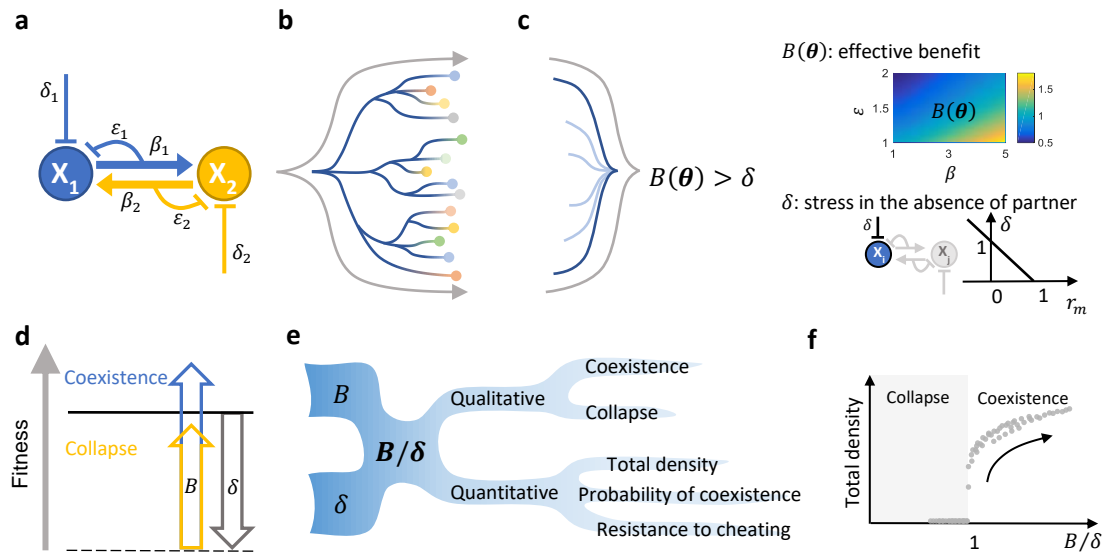


Figure 1: B and δ are two driving forces that determine qualitative and quantitative mutualistic outcomes

- (a) **The basic logic of mutualistic systems.** The two partner populations are denoted by X_1 and X_2 . β_1 and β_2 describe the strength of benefit the partners provide each other. ε_1 and ε_2 describe the cooperation cost of providing benefit. Each population also experience stress δ_1 and δ_2 .
- (b) **Models originating from the basic logic of mutualism yield diverse coexistence criteria.** Each line represents the generation of a model from the basic mutualism logic and branching represents different implementation details and system complexities. The circles represent the models and different colors represent diverse coexistence criteria derived from these models. This process aims to reflect the diversity of mutualistic systems in nature.
- (c) **A simple rule emerges at an appropriate level of abstraction.** The lines represent the abstraction process that establishes $B(\theta) > \delta$ as the common structure shared by diverse criteria in panel b. $B(\theta)$ represents effective benefit and is a complex function of model parameters θ , which include β and ε . In general, B increases with increasing β and decreasing ε . The heatmap is generated using the following model:
- $$\frac{dX_1}{d\tau} = \frac{1}{\varepsilon_1} X_1 (1 - X_1) - \frac{\delta_1}{\beta_2 X_2 + 1} X_1$$
- $$\frac{dX_2}{d\tau} = \frac{1}{\varepsilon_2} X_2 (1 - X_2) - \frac{\delta_2}{\beta_1 X_1 + 1} X_2$$
- δ is the stress experienced by one population in the absence of its partner.
- (d) **Intuitive interpretation of the simple rule.** The effective benefit B must overcome stress δ for the system to coexist. Solid black line represents coexistence boundary and dashed black line represents baseline fitness level in the absence of partner. Blue represents a B that is greater than δ (coexistence) and yellow represents a B that is smaller than δ (collapse).
- (e) **B/δ can predict various system outcomes.** If the two features B and δ are known, many downstream predictions, both qualitative and quantitative, can be made.
- (f) **Quantitative outcomes versus B/δ .** Simulation results show when $B/\delta > 1$, it is predictive of total density. Note that the points do not necessarily lie on a single curve, but a positive trend is well-maintained. Other quantitative outcomes also follow similar positive trends when plotted against B/δ .

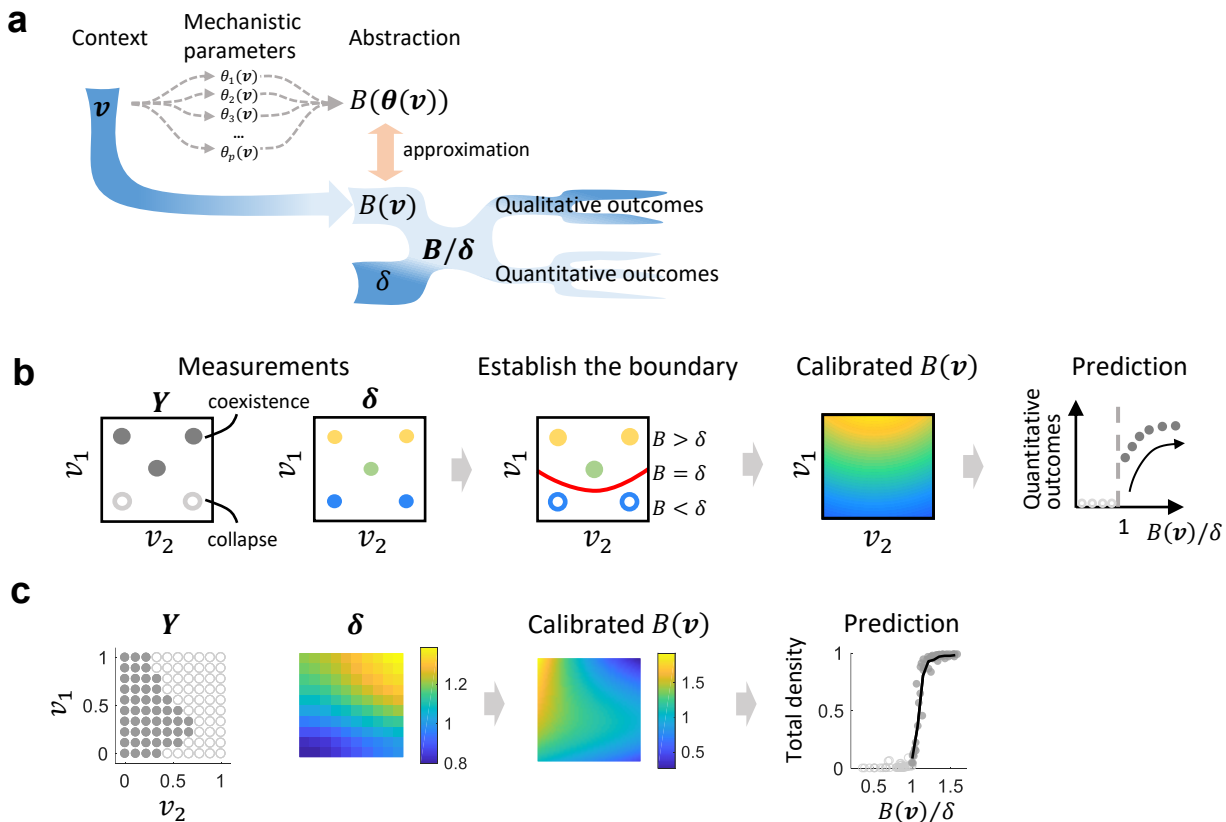


Figure 2: A streamlined approach to calibrate for an empirical $B(v)$.

- (a) **The rationale behind the calibration procedure.** Conventional approaches (denoted by dashed gray arrows) require quantifications of mechanistic parameters as functions of contextual variables $\theta(v)$ and finding the appropriate structure of $B(\theta)$ to construct $B(\theta(v))$. However, both steps are challenging and require case-by-case applications due to the diversity of mutualistic interactions. Instead, using qualitative outcomes of the system, we can distill an empirical function $B(v)$ to approximate the true $B(\theta(v))$. $B(v)/\delta$ can then predict qualitative and quantitative outcomes. Dark blue indicates the data that are relatively easy to measure without requiring mechanistic understanding of the interaction.
- (b) **A schematic demonstrating the mathematical basis of the calibration procedure.** v_1 and v_2 represent two system variables. A circle represents an observation i at a particular $v_i = (v_{1i}, v_{2i})$. 5 observations are shown. Y contains qualitative system outcomes for each observation. Closed circles indicate coexistence and open circles indicate collapse; the same notation scheme is used for all following figures. δ contains the measurement of stress for each observation (lighter colors indicate higher values). Using v , Y and δ , a boundary that separates the two types of outcomes can be established (the red curve). According to our simple rule, $B = \delta$ is true at the boundary; $B > \delta$ is true for coexistence and $B < \delta$ is true for collapse. Using these data and our simple rule, we can calibrate for a $B(v)$ which then enables the interpretation and prediction of system outcomes. Refer to Movie S1 for a 3D visualization of the calibration.
- (c) **Proof of principle using simulated data.** Simulations were performed using a complex mutualism model that does not have an explicit form of $B(\theta)$ (see Supplementary Text section V.F). The input data set contains 100 observations. δ and calibrated $B(v)$ share the same axes with Y (this applies to all following figures). $B(v)/\delta$ correctly classifies 97.2% of 2500 new data points. $B(v)/\delta$ is also predictive of total densities (only 100 data points are shown out of 2500). Black trace in the plot named “Prediction” represents binned averages of total density (this applies to all following figures).

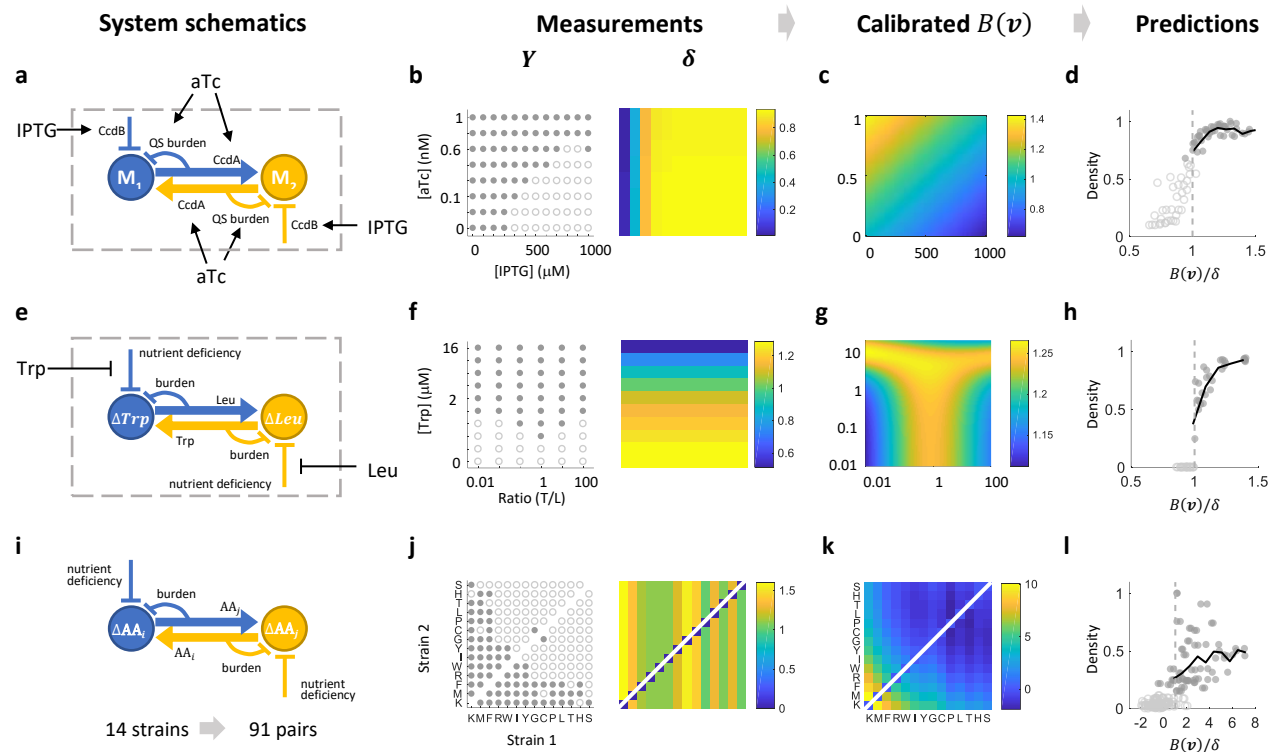
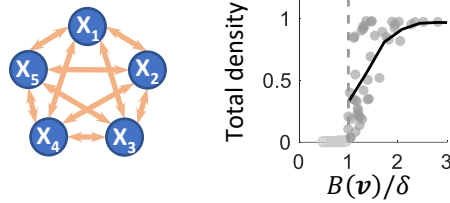


Figure 3: Implementation of the simple rule in experimental systems (Movie S2-4 for 3D visualizations)

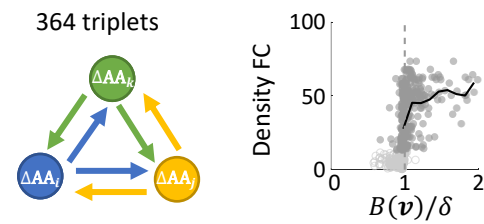
- (a) **The QS-based mutualism system.** IPTG modulates stress and aTc induces QS-mediated mutualistic interaction.
- (b) **Measurements of coexistence and collapse and corresponding δ values.** Coexistence and collapse are measured by co-culturing the two strains starting from the same initial densities. δ is measured by OD of M_2 mono-culture after 32 hours of culturing.
- (c) **Empirical calibration of $B(v)$.** $B(v)$ reveals how [IPTG] and [aTc] together modulate the effectiveness of the interaction.
- (d) **$B(v)/\delta$ is predictive of coexistence versus collapse and total final density.** The x-axis range of [0.5, 1.5] is used to highlight the transition (this also applies to other prediction plots). The trend continues to hold beyond this range. The y axis represents normalized final cell density (the same applies to panel h and l).
- (e) **The pair-wise yeast auxotroph system.** The growth of both auxotrophs are suppressed in monocultures. With increasing Trp and Leu supplemented to the co-culture, the growth suppression can be alleviated.
- (f) **The amount of supplemented amino acid and ratio of initial densities modulate system behavior.** Only [Trp] is shown and [Leu] is 8 times of [Trp]. The total initial density of the two strains are kept constant. Corresponding δ values are measured based on growth yield of ΔLeu monocultures, assuming δ is independent of initial density (Fig. S7a).
- (g) **Optimal effective benefit occurs at an intermediate ratio of initial density.** The calibrated $B(v)$ provides a cross validation accuracy of 95.0%.
- (h) **$B(v)/\delta$ is predictive of normalized total cell number per culture well.**
- (i) **The 91 mutualism systems constructed by 14 engineered *E. coli* auxotrophs.** Growth suppression is evident in their inability to survive individually in minimal medium. However, two auxotrophs can potentially survive through mutualistic interaction in a coculture by exchanging amino acids. The metabolic burden of the exchanging amino acid can be minimal.
- (j) **System outcomes for all 91 pairs and δ for each of the 14 auxotrophs.** Note that for one pair, the calibration is done twice with δ of either strain.
- (k) **The predictive power of $B(v)/\delta$.** The calibrated $B(v)$ and δ measurements provide a 91.8% cross validation accuracy.
- (l) **$B(v)/\delta$ is predictive of the normalized fold change of final total density relative to initial density.**

N-mutualist systems

a

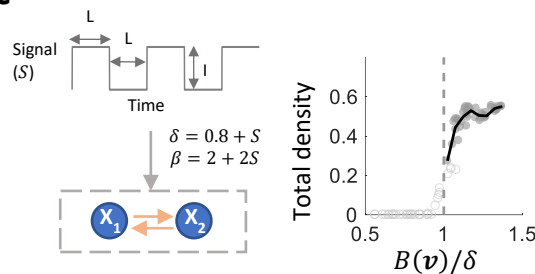


b



Dynamic environments

c



d

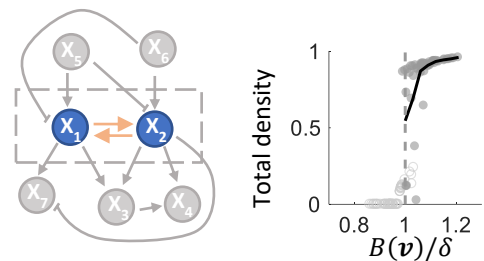


Figure 4: Our framework is generalizable to mutualistic systems with more than two partners and systems that inhabit dynamic environments. Please refer to section VI of Supplementary Text for simulation details.

- (a) **A simulated mutualistic system with 5 partners.** Parameters in the model are functions of two independent variables. Using 100 data points, we obtained a $B(v)$ through calibration. $B(v)/\delta$ successfully predicts coexistence versus collapse for 98.6% of a new set of 2500 data points (100 points are shown) and it is also predictive of total density.
- (b) **Experimental auxotrophic triplets that are comprised of 14 *E. coli* auxotrophs.** The same auxotrophic strains presented in Fig. 3i-k were used to construct three-way mutualistic systems. This experimental validation demonstrates the generality of framework beyond pairwise interactions.
- (c) **A system that is modulated by an oscillatory signal.** The oscillatory signal S is described by I (intensity) and L (duration). The signal temporally modulates δ and β . I and L are the two system variables used in calibration. The procedure achieves a prediction accuracy of 97.3% for new data.
- (d) **A simulated mutualistic system that coinhabits with 5 other populations.** X_1 and X_2 are the mutualistic partners. X_3 to X_7 are bystander populations that either modulate or are modulated by X_1 and X_2 . $B(v)/\delta$ successfully predicts 92.3% of new data in this example.

Supplementary Materials for

Unifying mutualism diversity for interpretation and prediction

Feilun Wu, Allison J. Lopatkin, Daniel A. Needs, Charlotte T. Lee, Sayan Mukherjee, Lingchong You

Correspondence to: you@duke.edu

This PDF file includes:

- Materials and Methods
- Supplementary Text
- Tables S1 to S2
- Figs. S1 to S10
- Captions for Movies S1 to S4

Other Supplementary Materials for this manuscript include the following:

- Movies S1 to S4
- MATLAB code package

Materials and Methods

Model development.

We built mutualism models based on four key assumptions:

- a) Benefit shall increase growth rate or carrying capacity and is positively dependent on partner density.
- b) Cost shall *decrease* growth rate or carrying capacity.
- c) Stress shall produce negative growth of populations at some parameter combinations.
- d) Negative growth of a population shall be potentially counteracted by benefit provided by a partner, but further strengthened by cost.

See SI section II for detailed reasoning and implementation of each assumption.

Criteria derivation.

We calculated the analytical solutions of fixed points of the 52 models using MATLAB R2017a. Then we identified the fixed points that represent stable coexistence. The coexistence criteria are derived by ensuring the fixed points are real positive numbers. We can then rearrange the inequality to have δ on one side. The other side of the inequality is then an expression of other parameters, which is expressed as $B(\theta)$. More details are presented in SI section III. The MATLAB code of the models and the derivation and testing process is included in the Supplementary Materials.

Calibration of benefit landscape using SVM.

We used support vector machine (SVM) algorithms in MATLAB to implement the calibration. The input data are formulated into the following format:

$$\text{Label of coexistence versus collapse:} \quad \mathbf{Y} = [y_1, y_2, \dots, y_n] \quad (\text{Eq. 1})$$

$$\text{System variables:} \quad \mathbf{v} = [v_1, v_2, \dots, v_n] \quad (\text{Eq. 2})$$

$$\text{Stress of the reference population:} \quad \boldsymbol{\delta} = [\delta_1, \delta_2, \dots, \delta_n] \quad (\text{Eq. 3})$$

In Eq. 2-4, n represents total number of observations and each index represents one observation. \mathbf{Y} takes values of 1 or -1 , which represent coexistence versus collapse for each observation. \mathbf{v} contains the

coordinates in the independent variable space where observations are obtained. δ contains the stress levels of the reference population for each observation.

We designed kernels that have additive separability between \mathbf{v} and δ , which can be expressed in a general form:

$$K([\mathbf{v}_1, \delta_1], [\mathbf{v}_2, \delta_2]) = K_v(\mathbf{v}_1, \mathbf{v}_2) + k_\delta(\delta_1 \cdot \delta_2), \quad (\text{Eq. 4})$$

where K_v is the kernel that dictates the shape of the empirical function of B and k_δ is a constant that varies the weight of δ . The predictor trained using SVM can then be expressed as:

$$f([\mathbf{v}, \delta]) = \sum_i \alpha_i y_i K_x(\mathbf{v}_i, \mathbf{v}) + k_\delta \delta \sum_i \alpha_i y_i \delta_i + \lambda_0. \quad (\text{Eq. 5})$$

According to our criterion, we know that $B = \delta$ when $f([\mathbf{v}, \delta]) = 0$. We can then derive from Eq. 6 a function of B in terms of system variables \mathbf{v} :

$$B(\mathbf{v}) = \delta = \frac{-\sum_i \alpha_i y_i K_x(\mathbf{v}_i, \mathbf{v}) - \lambda_0}{k_\delta \sum_i \alpha_i y_i \delta_i}. \quad (\text{Eq. 6})$$

With each set of inputs, we used linear, quadratic, cubic, and sigmoidal kernels with different kernel parameters to train $B(\mathbf{v})$. We then find the $B(\mathbf{v})$ with lowest cross-validation classification loss and bootstrapped variance. The top ranked $B(\mathbf{v})$ is then used along with δ measurements for interpretation and prediction. See section V of SI for detailed calibration method and the Supplementary Materials for the MATLAB code of the calibration process and sample data sets.

QS-based mutualism strains.

The two strains were constructed based on circuit components from a synthetic predator-prey system (66, 67). Both populations carry two plasmids. Briefly, M_1 carries plasmids identical to the predator plasmids, denoted A1 for the module carrying *ccdA* (*tet* promoter (68) driving *luxR* and *lasI* followed by *lux* promoter driving *ccdA*) and B1 for the module carrying *ccdB* (*Lac* promoter (68) upstream of *ccdB* followed by *tet* promoter upstream of *gfp*). To construct M_2 , A1 was used as backbone. To obtain orthogonal communication, KpnI and NotI restriction digest cloning was used to replace *luxR/lasI* genes from A1 with

lasR/luxI genes from the previously published prey plasmid (consisting of pLac *lasRluxI* CcdB (Kan^R, p15A ori)). Reporter plasmid B1 is from (66). To construct B2, enzymes XhoI and KpnI were used to replace the *tet* promoter on prey plasmid with the *ccdB* module from B1. All M₁ and M₂ plasmids were verified using restriction digest and sequencing.

Growth conditions of QS-based synthetic system.

The experiments of QS-based mutualistic system were done in 96-well microtiter plates. PH-buffered M9 medium (M9 salt supplemented with 1mM thiamine, 0.2% casamino acid, 0.4% glucose, 2mM MgSO₄, 0.1mM CaCl₂ and buffered with 100mM MOPS with PH adjusted to 7.0) was used. 50 µg/mL kanamycin and 100 µg/mL chloramphenicol were added to the culture to maintain plasmids.

To measure circuit function, 4 ml LB media in a 14 ml culture tube was inoculated from single colony and incubated overnight at 37 °C at 250 r.p.m. The optical density is adjusted to 0.5 in M9 media (measured at 600 nm with TECAN microplate reader) before use. Cocultures are created by mixing both strains in a 1:1 volume ratio. The culture is then diluted 10⁶ fold and cultured in 200µL batch culture at 30°C in TECAN plate reader to record OD for 32 hours with 10 minutes between each reading. The inducers were added to the media at the beginning with cell culture.

Supplementary Text

Contents

I. Previous mutualism models.....	6
A. The Lotka-Volterra model.....	6
B. Other variants.....	6
II. Building various mutualism models.....	7
A. Incorporation of stress.....	8
B. Model assumptions.....	8
C. Modifications of the logistic growth equation	9
D. Specific implementations of benefit, stress and cost.....	10
III. Derivation of diverse criteria	11
A. An example	11
B. Summary of criteria derived from symmetric models.....	12
C. Criteria from models with additional complexities	12
IV. Theoretical generality of the simple metric	16
A. Establish the general predictive power of \mathbf{B}/δ	17
B. The predictive accuracy is maintained in the presence of noise.....	18
V. Calibration using SVM	18
A. Kernels	18
B. Standardizing input data.....	19
C. From SVM output to calibrated $\mathbf{B}(\mathbf{v})$	19
D. Cross validation and bootstrapping.....	20
E. Twenty sets of simulations to establish and test the calibration procedure	20
F. Calibration with simulated data with unknown $\mathbf{B}(\boldsymbol{\theta})$: an example	21
VI. Framework generality verified by complex mutualistic systems.....	22
A. A 5-mutualist system	22
B. The experimental 3-member mutualistic systems	23
C. A mutualistic system in an oscillatory environment	23
D. A mutualistic system cohabiting with other populations	23
References.....	25

I. Previous mutualism models

Here we briefly summarize some limitations of previously published models in studying the transition between mutualism coexistence and collapse.

A. The Lotka-Volterra model

Consider the following model that follows the basic form of Lotka-Volterra model of mutualism. This formulation is a nondimensionalized form of previous models (69-74). The parameters are renamed according to our parameter assignments.

$$\frac{dX_1}{d\tau} = X_1(1 - \delta_1 X_1) + \beta_1 X_2 X_1, \quad (\text{I. 1})$$

$$\frac{dX_2}{d\tau} = \rho X_2(1 - \delta_2 X_2) + \beta_2 X_1 X_2. \quad (\text{I. 2})$$

Although the model formulation captures the logic of mutualism, it can generate unbounded growth of the two partners, which is not a biologically relevant state. For example, when $\rho = 1$, $\beta_1 = \beta_2 = 2$ and $\delta_1 = \delta_2 = 1$, the two populations are not bounded by their carrying capacities, but both grow exponentially. Importantly, the model does not capture population collapse so it cannot explain the transition between collapse and coexistence. Thus, this model formulation is not suitable for our purpose.

A coexistence criterion for mutualism can be derived using L-V model formulation and is previously demonstrated (69-73):

$$\beta_1 \beta_2 < \rho \delta_1 \delta_2. \quad (\text{I. 3})$$

However, this criterion captures the transition between stable coexistence and unbounded growth. It also suggests that mutualism is destabilizing, since increasing the strength of mutualistic interaction ($\beta_1 \beta_2$) tends to violate the above condition (I. 3). The interpretation of this criterion can be contradictory to empirical observations that mutualism stabilizes community (75-77). Although Lotka-Volterra models are sufficient to answer many questions related to mutualistic systems, it has been proposed that this discrepancy between model dynamics and empirical observations can be attributed to its unrealistic assumptions (78-81).

B. Other variants

We found that general mutualism models that generate unbounded growth usually do not capture the transition between coexistence and collapse. For example, the following three models were established previously to find structurally stable mutualistic models (78). In contrast to the L-V model which implements the interaction as a linear term, the following models present three alternative ways of adding the interaction to the basic logistic growth equation.

$$\begin{aligned} \frac{dN_1}{dt} &= r_1 N_1 \frac{K_1 - N_1 + \alpha_{12} N_2}{K_1 + \alpha_{12} N_2} \\ \frac{dN_2}{dt} &= r_2 N_2 \frac{K_2 - N_2 + \alpha_{21} N_1}{K_2 + \alpha_{21} N_1} \end{aligned} \quad (\text{I. 4})$$

$$\begin{aligned}\frac{dN_1}{dt} &= r_1 N_1 \frac{K_1 - N_1 + \alpha_{12} N_2}{K_1} \\ \frac{dN_2}{dt} &= r_2 N_2 \frac{K_2 - N_2 + \alpha_{21} N_1}{K_2}\end{aligned}\tag{I.5}$$

$$\begin{aligned}\frac{dN_1}{dt} &= r_1 N_1 \frac{\left(1 + \frac{\alpha_{12} N_2}{K_1}\right) (K_1 - N_1)}{K_1} \\ \frac{dN_2}{dt} &= r_2 N_2 \frac{\left(1 + \frac{\alpha_{21} N_1}{K_2}\right) (K_2 - N_2)}{K_2}\end{aligned}\tag{I.6}$$

In all these three cases, some parameter combinations generate unbounded growth. For example, $r_1 = r_2 = 1$; $K_1 = K_2 = 1$; $\alpha_{12} = \alpha_{21} = 2$; $N_{10} = N_{20} = 0.1$ generates unbounded growth for I.4 and I.5 and $r_1 = r_2 = -1$; $K_1 = K_2 = 1$; $\alpha_{12} = \alpha_{21} = 2$; $N_{10} = N_{20} = 2$ generates unbounded growth for I.6. Unbounded growth has been recognized as a limitation of many mutualism models (82-84). In addition, although population collapse can be simulated by $r_1 < 0$ and/or $r_2 < 0$, in all cases, the fitness of the benefit-receiver decreases with increasing partner density, which is contradictory to the basic logic of mutualism.

To avoid generating unbounded growth, one strategy is to introduce saturating benefit (73). However, although preventing unbounded growth, these models may still not generate negative growth which can be potentially countered by the increase of partner density. For example, the following model is adapted from a previous work (81) which falls into this category. In addition, if no decreasing density is captured, this model will stabilize at its coexistence state with any positive initial density for both populations.

$$\begin{aligned}\frac{dX_1}{d\tau} &= r_1 N_1 \left(1 - \delta_1 N_1 + \frac{\beta_{12} N_2}{1/\beta_2 + N_2}\right) \\ \frac{dX_2}{d\tau} &= r_2 N_2 \left(1 - \delta_2 N_2 + \frac{\beta_{21} N_1}{1/\beta_1 + N_1}\right)\end{aligned}\tag{I.7}$$

Due to the limitations of previous mutualism models in studying the transition between coexistence and collapse, a more systematic way of modeling mutualism is required to both ensure the basic logic of mutualism and capture the transition between stable coexistence and collapse.

II. Building various mutualism models

The diversity of mutualistic systems impedes the formulation of a general mutualism model and the general proof of a single criterion. We generated various mutualism models to reflect the diversity of mutualistic systems in nature and examine how diverse the coexistence criteria are. We also aim to investigate whether there exists an invariant form that is preserved regardless of specific model implementations. If such an invariant form exists, it will then reveal a fundamental coexistence criterion that is originated from the basic logic of mutualism.

A. Incorporation of stress

We explicitly define stress as a reduction of growth rate or productivity of biological systems, which is consistent with previous works (85-88). Stress is universal in biology because it is present whenever growth rate is lower than the optimal growth rate. In mutualistic systems, many studies have shown that stress factors alter the basal fitness of individual mutualists. These factors include nutrient limitation (89), rising temperature (90, 91), rising CO₂ levels (92), invasive species (93), etc. Beyond benefit and cost, it is known that these stress factors also determine mutualistic outcomes (94).

To study quantitatively how stress affects mutualistic outcomes, we include stress as a model parameter that reduces the growth rate or carrying capacity of individual mutualists in various biotic/abiotic contexts. In previous models, stress has been included as linear turn-over rate (66, 89, 95-97) or intraspecific competition (69, 70). Although these studies have more specific terms describing the downward pressure they capture, we use “stress” as an appropriate umbrella term.

Stress, thus defined, plays an essential qualitative role in mutualism. First, an absence of stress would mean that populations operate at maximum fitness. Such populations could not benefit from mutualism, because they would already be operating at their optimal level. Second, in mutualistic models the downward pressure can resolve the unrealistic exponential growth of mutualistic partners by imposing an upper limit to mutual benefit (78). Third, using stress can dissect the baseline fitness of individual mutualists caused by biotic/abiotic factors from the effect of mutualistic interaction on fitness.

B. Model assumptions

The key step of generating various mutualism models is to establish the set of assumptions the models should follow. We first start off with the most apparent aspects of mutualism: benefit and cost. These two aspects lead to two assumptions:

- e) Benefit shall *increase* growth rate or carrying capacity and is positively dependent on partner density.
- f) Cost shall *decrease* growth rate or carrying capacity.

To study the transition between coexistence and collapse in mutualism systems, the models must be able to simulate collapse, leading to the third assumption where we explicitly introduce stress to achieve negative growth. In addition to generating negative growth, stress has a physical meaning which is the difference between maximal fitness and baseline fitness in the absence of the partner:

- g) Stress shall produce negative growth of populations with some parameter combinations.

The fourth assumption follows assumption c) to reinforce the effects of benefit and cost in mutualistic interaction:

- h) Negative growth of a population shall be potentially counteracted by benefit provided by a partner, but is further strengthened by cost.

Even when all the above assumptions are satisfied, we still need to verify that the models do not generate unbounded growth with *any* parameter set.

C. Modifications of the logistic growth equation

After establishing a minimal set of assumptions, we then need to establish a systematic way of generating diverse models.

Mathematically, there are infinite possible implementations of a mutualistic system. We attempt to cover the different *common and plausible* forms of kinetic models that have been adopted in previous studies. Previous models have captured benefit, cost and stress in many ways. The following is a short summary of how benefit, stress and cost are modeled previously that serve as building blocks for our own models. Many models use Hill equations to capture saturating benefit (66, 81, 89, 98, 99). Cost is also an aspect that is widely modeled, which can be implemented many ways (81, 100), such as independent or dependent of partner density. Although stress is not often generally discussed, one possible form of stress is self-regulation, also called density dependence or inter-population competition, which often appears in large-scale models (69-71). Linear death rate (often imposed by dilution) is also another common form of stress (66, 89, 95-97).

Inspired by these observations, we first examined different methods to modify a logistic growth equation.

Consider a basic logistic growth equation:

$$\frac{dN}{dt} = \mu N \left(1 - \frac{N}{N_m}\right). \quad (\text{II. 1})$$

This equation only has two parameters, growth rate μ and carrying capacity N_m . If we can derive a non-dimensionalized logistic growth equation, it can be rewritten as:

$$\frac{dX}{d\tau} = X(1 - X) \quad (\tau = \mu t, X = N/N_m). \quad (\text{II. 2})$$

The followings are common modifications of the above equation:

1. The growth rate can be modified by multiplying the right-hand-side with a constant:

$$\frac{dX}{d\tau} = \alpha X(1 - X). \quad (\text{II. 3})$$

This equation modifies the growth rate from 1 to α , where increasing α increases growth rate and leaves the carrying capacity the same. α can take any real number in this case.

2. The carrying capacity can also be modulated, leaving the growth rate unchanged:

$$\frac{dX}{d\tau} = X(1 - \alpha X). \quad (\text{II. 4})$$

The carrying capacity becomes $1/\alpha$ and $\alpha > 0$. This formulation does not generate negative growth.

3. The "1" in the logistic growth equation can also be modified:

$$\frac{dX}{d\tau} = X(\alpha - X). \quad (\text{II. 5})$$

It can be rewritten as the same form of logistic growth equation:

$$\frac{dX}{d\tau} = \alpha X \left(1 - \frac{X}{\alpha}\right). \quad (\text{II. 6})$$

In this case, both growth rate and carrying capacity are scaled by a factor of α . α can take any real number because if the carrying capacity is negative, the growth rate will also be negative and thus the model will only generate bounded growth.

4. A first order term can be added to the equation:

$$\frac{dX}{d\tau} = X(1 - X) - \alpha X. \quad (\text{II. 7})$$

Although this modification is equivalent to the previous one, it is worth noting since this form is commonly used to represent death rate or turnover rate.

The above analysis demonstrates that 1, 3, and 4 are robust modifications that provide consistent model modulations with any real value of α parameter. Following this analysis, we will then model stress, benefit and cost at these three locations in the logistic growth equation to capture the assumptions presented above for mutualistic interactions.

D. Specific implementations of benefit, stress and cost

To capture the model assumptions of a), b) and c), we used the following formulations to modify simple logistic growth equations at location 1, 3 or 4 mentioned above to capture the logic of mutualistic interaction.

- a) *Benefit shall be positively dependent on partner density.*

This assumption can be modeled at all three locations by

$$\frac{\beta' X_2}{X_2 + 1/\beta} \quad (\beta \geq 0, \beta' > 0) \text{ or } \frac{-1}{\beta X_2 + 1} \quad (\beta \geq 0), \quad (\text{II. 8})$$

where X_2 represents partner density (same as below) and β and β' are measures of strength of benefit. In both cases, the benefit function is bounded, which facilitates the stabilization of population densities (73).

- b) *Cost shall decrease growth rate or carrying capacity.*

We use ε as a measure of level of cost. Specifically, this is implemented at location 1 or 3 by multiplication

$$\frac{1}{\varepsilon} \quad (\varepsilon \geq 1) \text{ (density-independent) or } \frac{1}{\varepsilon X_2 + 1} \quad (\varepsilon \geq 0) \text{ (density-dependent).} \quad (\text{II. 9})$$

Cost can also be implemented as a turnover rate at location 4

$$-\varepsilon \quad (\varepsilon \geq 0) \text{ (density-independent) or } -\varepsilon X_2 \quad (\varepsilon \geq 0) \text{ (density-dependent).} \quad (\text{II. 10})$$

- c) *Stress shall produce negative growth of populations with some parameter combinations.*

We use δ as a measure of stress. Specifically, this is implemented by modifying the logistic growth equation at location 1 or 3 by

$$1 - \delta \text{ where } (\delta \geq 0). \quad (\text{II. 11})$$

Stress can also serve as a turnover rate at location 4

$$-\delta. \quad (\text{II. 12})$$

Two or more of the parameters can share one location, so we permuted the three factors in the three locations and systematically generated 81 models (Table S1). The 81 models are then checked against

assumption d) and verify that the models do not generate unbounded growth. See Supplementary Materials for the MATLAB code that investigates the model assumption d) and unbounded growth behavior. Out of the 81 models, 48 satisfy all 4 assumptions and do not generate unbounded growth with positive parameter values and positive initial densities.

III. Derivation of diverse criteria

This section demonstrates the diversity of criteria with various model formulations. We first use an example model to serve as the base model to demonstrate a detailed process to derive coexistence criteria. We then show the process of building various mutualism models with more complexity and deriving or approximating the corresponding criterion.

A. An example

Consider the following model as our base model for the analytical analyses:

$$\frac{dN_1}{dt} = \frac{\mu}{\varepsilon} N_1 \left(1 - \frac{N_1}{N_m}\right) - \frac{d K}{N_2 + K} N_1, \quad (\text{III. 1})$$

$$\frac{dN_2}{dt} = \frac{\mu}{\varepsilon} N_2 \left(1 - \frac{N_2}{N_m}\right) - \frac{d K}{N_1 + K} N_2. \quad (\text{III. 2})$$

The non-dimensionalized version of the model is (model 21 in Table S1):

$$\frac{dX_1}{d\tau} = \frac{1}{\varepsilon} X_1 (1 - X_1) - \frac{\delta}{\beta X_2 + 1} X_1, \quad (\text{III. 3})$$

$$\frac{dX_2}{d\tau} = \frac{1}{\varepsilon} X_2 (1 - X_2) - \frac{\delta}{\beta X_1 + 1} X_2. \quad (\text{III. 4})$$

Where $\tau = \mu t$, $X_1 = \frac{N_1}{N_m}$, $X_2 = \frac{N_2}{N_m}$, $\delta = \frac{d}{\mu}$, $\beta = \frac{N_m}{K}$. The ranges of the parameters are: $\varepsilon \geq 1$, $\delta \geq 0$, $\beta \geq 0$

This system has 5 fixed points:

$$(0, 0), \quad (\text{III. 5})$$

$$(1 - \varepsilon\delta, 0), \quad (\text{III. 6})$$

$$(0, 1 - \varepsilon\delta), \quad (\text{III. 7})$$

$$\left(\frac{\beta - 1 + \sqrt{(\beta + 1)^2 - 4\beta\varepsilon\delta}}{4\beta}, \frac{\beta - 1 + \sqrt{(\beta + 1)^2 - 4\beta\varepsilon\delta}}{4\beta} \right), \quad (\text{III. 8})$$

$$\left(\frac{\beta - 1 - \sqrt{(\beta + 1)^2 - 4\beta\varepsilon\delta}}{4\beta}, \frac{\beta - 1 - \sqrt{(\beta + 1)^2 - 4\beta\varepsilon\delta}}{4\beta} \right), \quad (\text{III. 9})$$

By calculating the Jacobian, we found that the 4th fixed point (III.8) represents stable coexistence. The 5th fixed point is a saddle point and is unstable. For the 4th fixed point to be in the first quadrant of the Real domain, 2 conditions must be satisfied at the same time:

$$(\beta + 1)^2 - 4\beta\varepsilon\delta \geq 0, \quad (\text{III. 10})$$

$$\frac{\beta - 1 + \sqrt{(\beta + 1)^2 - 4\beta\epsilon\delta}}{4\beta} > 0. \quad (\text{III. 11})$$

The first condition (III.10) derives the coexistence criterion $\frac{(\beta+1)^2}{4\beta\epsilon} \geq \delta$ and when this criterion is satisfied, the second condition is automatically satisfied if $\beta > 1$. When $\beta \leq 1$, only facultative mutualism can be captured. Bifurcation analysis of the system shows Allee effect and how final density increases with increasing β , decreasing δ or decreasing ϵ .

B. Summary of criteria derived from symmetric models

Symmetric models in general generate interpretable analytical solutions. In Table S1, all the models are assumed to be symmetric between two populations. We found that these symmetric models already generate a diverse set of coexistence criteria.

Most of these criteria are derived using the same logic used for the base model presented above, except for the criteria of model 1, 4, 28, 31, 55 and 58 in Table S1, which are derived from setting the analytical solution of the saddle point lower than the initial density X_0 . This is because the steady states that represent coexistence for these models are constants, whereas the saddle points are modulated by model parameters. For a detailed example of the derivation, refer to section III.C.2.

In natural mutualism systems, cost can also scale with its partner's density (100). In general, we observed that inclusion of density-dependent cost increases the complexity of the criteria. In addition, mutualists can also compete in within the same niche (101). In this case, the effective benefit term in general decreases from the model without competition, indicating the criteria also lumps the effect of competition.

C. Criteria from models with additional complexities

To relax some additional assumptions, such as complete symmetry, we also included additional levels of complexities to the base model by including asymmetry and turnover rate:

$$\frac{dN_1}{dt} = \frac{\mu_1}{\epsilon_1} N_1 \left(1 - \frac{N_1 + \alpha N_2}{N_m} \right) - \frac{d_1 K_2}{N_2 + K_2} N_1 - d_{01} N_1, \quad (\text{III. 12})$$

$$\frac{dN_2}{dt} = \frac{\mu_2}{\epsilon_2} N_2 \left(1 - \frac{\alpha N_1 + N_2}{N_m} \right) - \frac{d_2 K_1}{N_1 + K_1} N_2 - d_{02} N_2. \quad (\text{III. 13})$$

After non-dimensionalization, the model becomes:

$$\frac{dX_1}{d\tau} = \frac{1}{\epsilon_1} X_1 (1 - X_1 - \alpha X_2) - \frac{\delta_1}{\beta_2 X_2 + 1} X_1 - \delta_{01} X_1, \quad (\text{III. 14})$$

$$\frac{dX_2}{d\tau} = \frac{1}{\epsilon_2} \rho X_2 (1 - \alpha X_1 - X_2) - \frac{\delta_2}{\beta_1 X_1 + 1} X_2 - \delta_{02} X_2, \quad (\text{III. 15})$$

where $\tau = \mu_1 t$, $X_1 = \frac{N_1}{N_m}$, $\rho = \frac{\mu_2}{\mu_1}$, $\delta_1 = \frac{d_1}{\mu_1}$, $\beta_1 = \frac{N_m}{K_1}$, $X_2 = \frac{N_2}{N_m}$, $\delta_2 = \frac{d_2}{\mu_1}$, $\beta_2 = \frac{N_m}{K_2}$.

1. Criterion with two populations competing for resources

Our base model assumes the two populations having separate carrying capacities. However, in natural settings, mutualists often share resources (83). We can impose competition between them by adding $-\alpha X_2$ or $-\alpha X_1$ to location 3 in logistic growth equations of X_1 and X_2 , respectively (number 82 in Table S2):

$$\frac{dX_1}{dt} = \frac{1}{\varepsilon} X_1 (1 - X_1 - \alpha X_2) - \frac{\delta}{\beta X_2 + 1} X_1, \quad (\text{III. 16})$$

$$\frac{dX_2}{dt} = \frac{1}{\varepsilon} X_2 (1 - \alpha X_1 - X_2) - \frac{\delta}{\beta X_1 + 1} X_2. \quad (\text{III. 17})$$

Using the same method described in section III.A, we derived the criterion of this model: $\frac{(\beta+\alpha+1)^2}{4\beta(\alpha+1)\varepsilon} > \delta$.

2. Criterion considering initial density

For this analysis, we used a symmetric model (number 83 in Table S2):

$$\frac{dX_1}{d\tau} = \frac{1}{\varepsilon} X_1 (1 - X_1) - \frac{\delta}{\beta X_2 + 1}, \quad (\text{III. 18})$$

$$\frac{dX_2}{d\tau} = \frac{1}{\varepsilon} X_2 (1 - X_2) - \frac{\delta}{\beta X_1 + 1}. \quad (\text{III. 19})$$

Because most mutualism models generate Allee effect, coexistence can also be affected by initial density (x_0). However, we can derive a benefit term that depends on the initial density x_0 (for both X_1 and X_2) to predict deterministically whether the two populations coexist or not. To coexist, x_0 needs to be greater than the saddle point:

$$x_0 > \frac{\beta - 1 - \sqrt{(\beta + 1)^2 - 4\beta\varepsilon\delta}}{4\beta}. \quad (\text{III. 20})$$

If we rearrange this inequality, we get

$$\frac{(1 - x_0)x_0\beta + (1 - x_0)}{\varepsilon} > \delta. \quad (\text{III. 21})$$

3. Criterion with asymmetric growth rate, cost, stress, and benefit

Asymmetric parameters of the mutualism are more realistic in capturing real-world mutualism systems, so we relaxed the assumptions that cost (ε), rescue strength (β) and stress (δ) are symmetric for both populations. The asymmetry of growth rate is captured by ρ . In the following model (number 84 in Table S2), we assumed X_1 and X_2 share the same carrying capacity. We found that separated carrying capacity also yields similar results.

$$\frac{dX_1}{d\tau} = \frac{1}{\varepsilon_1} X_1 (1 - X_1 - X_2) - \frac{\delta_1}{\beta_2 X_2 + 1} X_1 \quad (\text{III. 22})$$

$$\frac{dX_2}{d\tau} = \frac{1}{\varepsilon_2} \rho X_2 (1 - X_1 - X_2) - \frac{\delta_2}{\beta_1 X_1 + 1} X_2 \quad (\text{III. 23})$$

This model has 5 fixed points:

$$(0, 0), \quad (\text{III. 24})$$

$$(1 - \varepsilon_1 \delta_1, 0), \quad (\text{III. 25})$$

$$\left(0, 1 - \frac{\varepsilon_2 \delta_2}{\rho}\right), \quad (\text{III. 26})$$

$$\left(\frac{\frac{\varepsilon_2 \delta_2}{\rho} \sqrt{A} - \beta_2 \varepsilon_2 \delta_2 + \beta_1 \varepsilon_2 \delta_2 - 2\beta_1 \varepsilon_1 \delta_1 \rho + \beta_1 \beta_2 \varepsilon_2 \delta_2}{2\beta_1(\beta_1 \varepsilon_1 \delta_1 \rho + \beta_2 \varepsilon_2 \delta_2)}, \frac{\varepsilon_1 \delta_1 \sqrt{A} - \beta_1 \varepsilon_1 \delta_1 \rho + \beta_2 \varepsilon_1 \delta_1 \rho - 2\beta_2 \varepsilon_2 \delta_2 + \beta_1 \beta_2 \varepsilon_1 \delta_1 \rho}{2\beta_2(\beta_1 \varepsilon_1 \delta_1 \rho + \beta_2 \varepsilon_2 \delta_2)}\right), \quad (\text{III. 27})$$

$$\left(\frac{-\frac{\varepsilon_2 \delta_2}{\rho} \sqrt{A} - \beta_2 \varepsilon_2 \delta_2 + \beta_1 \varepsilon_2 \delta_2 - 2\beta_1 \varepsilon_1 \delta_1 \rho + \beta_1 \beta_2 \varepsilon_2 \delta_2}{2\beta_1(\beta_1 \varepsilon_1 \delta_1 \rho + \beta_2 \varepsilon_2 \delta_2)}, \frac{-\varepsilon_1 \delta_1 \sqrt{A} - \beta_1 \varepsilon_1 \delta_1 \rho + \beta_2 \varepsilon_1 \delta_1 \rho - 2\beta_2 \varepsilon_2 \delta_2 + \beta_1 \beta_2 \varepsilon_1 \delta_1 \rho}{2\beta_2(\beta_1 \varepsilon_1 \delta_1 \rho + \beta_2 \varepsilon_2 \delta_2)}\right), \quad (\text{III. 28})$$

where $A = \rho^2(\beta_1 \beta_2 + \beta_1 + \beta_2)^2 - 4\beta_1 \beta_2 \rho(\beta_1 \varepsilon_1 \delta_1 \rho + \beta_2 \varepsilon_2 \delta_2)$.

Use the same logic presented in section III.A, the following criterion is a necessary condition for the 4th fixed point to be present in the positive Real domain:

$$\frac{\rho(\beta_1 \beta_2 + \beta_1 + \beta_2)^2}{4\beta_1 \beta_2} \geq \beta_1 \varepsilon_1 \delta_1 \rho + \beta_2 \varepsilon_2 \delta_2. \quad (\text{III. 29})$$

This criterion can also be rewritten as $\frac{\rho(\beta_1 + \frac{1}{\theta_1} + 1)^2}{4\beta_1 \varepsilon_1 (\frac{\rho}{\theta_1} + \theta_2)} \geq \delta_1$ where the asymmetry is captured by $\theta_1 = \frac{\beta_2}{\beta_1}$ and $\theta_2 = \frac{\varepsilon_2 \delta_2}{\varepsilon_1 \delta_1}$. Note that we specifically used X_1 as the reference population.

4. Criterion with turnover rate and asymmetric growth rate

If we consider both the asymmetry in growth rate (ρ) and a turnover rate δ_0 , the model becomes (number 85 in Table S2):

$$\frac{dX_1}{d\tau} = \frac{1}{\varepsilon} X_1(1 - X_1 - X_2) - \frac{\delta}{\beta X_2 + 1} X_1 - \delta_0 X_1, \quad (\text{III. 30})$$

$$\frac{dX_2}{d\tau} = \frac{1}{\varepsilon} \rho X_2(1 - X_1 - X_2) - \frac{\delta}{\beta X_1 + 1} X_2 - \delta_0 X_2. \quad (\text{III. 31})$$

The analytical solution of this model is complex and involves solving 4th order polynomials. However, we can introduce the concept of correction terms to approximate the criterion. When we only add ρ in the model and assume $\delta_0 = 0$, we get the following criterion:

$$\frac{(\beta + 2)^2}{4\beta \varepsilon} \left(\frac{\rho}{\rho + 1}\right) \geq \delta. \quad (\text{III. 32})$$

In addition, when we only add the δ_0 term in the model and assume $\rho = 1$, we get

$$\frac{(\beta(1 - \varepsilon \delta_0) + 2)^2}{4\beta \varepsilon} \geq \delta. \quad (\text{III. 33})$$

We hypothesized that the criterion with both correction terms $\left(\frac{\rho}{\rho + 1}\right)$ and $(1 - \varepsilon \delta_0)$ can approximate the criterion for model III.30-31. The criterion for $\delta_0 > 0$ is:

$$\frac{(\beta(1 - \varepsilon \delta_0) + 2)^2}{4\beta \varepsilon} \left(\frac{\rho}{\rho + 1}\right) \geq \delta \quad (\text{III. 34})$$

The accuracy of this criterion is evaluated by substitution of variables:

$$\begin{cases} z_1 = (1 - X_1^* - X_2^*) - \delta_0 \varepsilon \\ z_2 = X_2^* + \frac{1}{\beta} \end{cases} \quad (\text{III. 35})$$

where X_1^* and X_2^* represent fixed points of the model. $\frac{dX_1}{d\tau} = 0$ and $\frac{dX_2}{d\tau} = 0$ can then be written as

$$\begin{cases} z_2 = \frac{\delta \varepsilon}{\beta z_1} \\ \rho z_1 + (\rho - 1)\delta_0 \varepsilon + \frac{\delta \varepsilon}{\beta(z_1 + z_2) - (\beta(1 - \delta_0 \varepsilon) + 2)} = 0 \end{cases} \quad (\text{III. 36})$$

After substitution of z_2 with z_1 in the second equation, we get

$$\beta z_1^2 - (\beta(1 - \delta_0 \varepsilon) + 2)z_1 + \frac{\rho + 1}{\rho} \delta \varepsilon = \frac{\frac{\rho - 1}{\rho} \delta_0 \delta \varepsilon^2}{\rho z_1 + (\rho - 1)\delta_0 \varepsilon} \quad (\text{III. 37})$$

The left-hand side is a quadratic equation of z_1 , which has $\Delta = (\beta(1 - \delta_0 \varepsilon) + 2)^2 - 4 \frac{\rho + 1}{\rho} \beta \delta \varepsilon$.

$$\Delta \geq 0 \Leftrightarrow \frac{(\beta(1 - \varepsilon \delta_0) + 2)^2}{4\beta \varepsilon} \left(\frac{\rho}{\rho + 1} \right) \geq \delta \quad (\text{III. 38})$$

Thus, we know that the criterion is more accurate when $\frac{\frac{\rho - 1}{\rho} \delta_0 \delta \varepsilon^2}{\rho z_1 + (\rho - 1)\delta_0 \varepsilon} \rightarrow 0$. The approximated criterion is accurate when $\rho = 1$. When $\rho \rightarrow +\infty$ and $\delta_0 \rightarrow 0$, the criterion will also provide good approximations.

5. The criterion for N-mutualist systems

The following model is used to determine the criterion with N-mutualist model:

$$\frac{dX_i}{d\tau} = \frac{1}{\varepsilon} X_i \left(1 - \sum_{k=1}^n X_k \right) - \frac{\delta}{\beta \sum_{k \neq i}^n X_k + 1} X_i. \quad (\text{III. 39})$$

The non-trivial steady state can be solved by solving the following equation for X^* :

$$\frac{n}{\varepsilon} X^* (1 - nX^*) - \frac{n\delta}{\beta(n - 1)X^* + 1} X^* = 0 \quad (\text{III. 40})$$

Using the same strategy in III.A, we derive the criterion for coexistence:

$$\frac{(\beta + \frac{n}{n-1})^2}{4\beta \varepsilon \frac{n}{n-1}} > \delta. \quad (\text{III. 41})$$

where benefit becomes a function of both β and n . This result suggests that mutualism system can tolerate higher stress levels with increasing number of mutualists.

6. Mutualism model including cheater exploitation

The mutualism model including cheaters is adapted from III.22-23:

$$\frac{dX_1}{d\tau} = \frac{1}{\varepsilon} X_1 (1 - X_1 - X_2 - X_{1c} - X_{2c}) - \frac{\delta}{\beta X_2 + 1} X_1 - \varphi X_1, \quad (\text{III. 42})$$

$$\frac{dX_2}{d\tau} = \frac{1}{\varepsilon} X_2 (1 - X_1 - X_2 - X_{1c} - X_{2c}) - \frac{\delta}{\beta X_1 + 1} X_2 - \varphi X_2, \quad (\text{III. 43})$$

$$\frac{dX_{1c}}{d\tau} = X_{1c} (1 - X_1 - X_2 - X_{1c} - X_{2c}) - \frac{\delta}{\beta X_2 + 1} X_{1c} + \varphi X_1, \quad (\text{III. 44})$$

$$\frac{dX_{2c}}{d\tau} = X_{2c} (1 - X_1 - X_2 - X_{1c} - X_{2c}) - \frac{\delta}{\beta X_1 + 1} X_{2c} + \varphi X_2. \quad (\text{III. 45})$$

We assume 1) cheaters and cooperators share the same carrying capacity, 2) cheaters accept benefit produced by cooperators but do not provide benefit, thus do not experience cost, and 3) there is a constant transfer rate (φ) from cooperator to cheater, representing mutation from the cooperator phenotype to cheater phenotype. Although this model does not generate stable coexistence of the cooperators, the time it takes for cheaters to take over the cooperators can serve as a metric for how stable the mutualistic system is.

7. Model structures that generate lower boundary for stress

We notice that to coexist, some model formulations not only require an upper boundary of δ , but also require a lower boundary as well. If δ is lower than a threshold, the system can be dominated by the population that is more fit. However, this lower boundary occurs due to the system dynamic shifting from a mutualism-dominated mode to a competition-dominated mode.

There are two types of system dynamics that can lead to a loss of partner. One is due to high stress where the weaker partner will go extinct and the stronger partner will suffer from the loss of its partner. The other is due to lack of stress where the system shifts to a competition-dominated interaction where the fitter partner survives better by excluding its partner. In our study, we only focus on the first type of partner loss since the second type is a dynamic of competitive systems instead of mutualistic systems.

In general, we observed that a model needs to be both asymmetric and potentially competitive to create a scenario where the fitter population excludes the weaker population. The lower boundary for δ decreases when the weaker population gives out more benefit to its partner or reduction of competition. This observation can potentially explain why some mutualistic systems collapse when external stress is reduced (89, 102).

IV. Theoretical generality of the simple metric

To establish the generality of the theoretical criterion, we verified that it is applicable to both symmetric and asymmetric mutualistic systems. We also want to test that the predictive power of the predictor is maintained when the partners are obligatory or facultative. Because mutualists can compete for the same resources in nature, we also verified that coexistence of partners that compete for resources also can be depicted by the criterion. We show some typical results of the predictive power of the criterion in these cases with Fig. S2.

A. Establish the general predictive power of B/δ

Fig. 1f is generated using the model presented by equation III.16-17, where $\varepsilon \in [1, 1.2]$, $\delta \in [0, 2]$, $\beta \in [2, 5]$ and $a = 1$. The positive trend holds even when the mutualism is facultative ($\delta < 1$).

Fig. S2a is also generated using the model presented by equation III.16-17. Both are generated with $\beta \in [2, 10]$, $\delta \in [1.2, 2]$, $\varepsilon \in [1, 1.2]$ and $a = 1$. The panel on the left assumes X_1 and X_2 have the same initial density which is a uniformly distributed between 0 and 0.5. The panel on the right assumes changing of individual initial densities in a linear fashion while maintaining the sum of initial density of X_1 and X_2 at 1.

Fig. S2b is generated with the model presented above in equation III.42-III.45, where $\varphi = 10^{-3}$, $\beta \in [5, 20]$, $\delta \in [1.2, 1.3]$, and $\varepsilon \in [1.2, 1.3]$. The initial densities are $X_{10} = 0.1$, $X_{1c} = 0$ and $X_{2c} = 0$. The time to cheater exploitation is quantified by the first time point where the total density of cooperators drops below their initial total density due to overwhelming competitions from the cheater populations. Other time to cheater exploitation in this study is quantified using the same method.

Fig. S2c is generated by

$$\frac{dX_1}{d\tau} = \frac{1}{\varepsilon_1} X_1 (1 - X_1) - \frac{\delta_1}{\beta_2 X_2 + 1} X_1 \quad (\text{IV. 1})$$

$$\frac{dX_2}{d\tau} = \frac{1.2}{\varepsilon_2} X_2 (1 - X_2) - \frac{\delta_2}{\beta_1 X_1 + 1} X_2 \quad (\text{IV. 2})$$

Where the left panel uses parameter values of $\beta_1 = 2$, $\beta_2 \in [2, 5]$, $\delta_1 = 1.2$, $\delta_2 \in [0.5, 2]$, $\varepsilon_1 = 1.1$, and $\varepsilon_2 \in [1, 1.2]$. The initial densities are $X_{10} = 1$ and $X_{20} = 1$. The right panel uses parameter values of $\beta_1 = 2$, $\beta_2 \in [2, 5]$, $\delta_1 = 0.8$, $\delta_2 \in [1, 2]$, $\varepsilon_1 = 1.1$, and $\varepsilon_2 \in [1, 1.2]$. The initial densities are $X_{10} = 1$ and $X_{20} = 1$. This panel specifically tests the cases where the survival of the partners is fully dependent on each other.

Fig. S2d is generated by III.22-III.23, where the left panel uses parameter values of $\rho = 1.2$, $\beta_1 = 2$, $\beta_2 \in [5, 10]$, $\delta_1 = 1.2$, $\delta_2 \in [0.5, 2]$, $\varepsilon_1 = 1.1$, and $\varepsilon_2 \in [1, 1.2]$. The initial densities are $X_{10} = 1$ and $X_{20} = 1$. The right panel uses parameter values of $\beta_1 = 2$, $\beta_2 \in [5, 10]$, $\delta_1 = 0.8$, $\delta_2 \in [1, 2]$, $\varepsilon_1 = 1.1$, and $\varepsilon_2 \in [1, 1.2]$. The initial densities are $X_{10} = 1$ and $X_{20} = 1$. This panel specifically tests the cases where one mutualist's survival is not fully dependent on the other.

Fig. S2e is generated using the same model as the above. Where the parameter values are $\beta_1 = 5$, $\beta_2 \in [2, 10]$, $\delta_1 = 1.2$, $\delta_2 \in [1.2, 2]$, $\varepsilon_1 = 1.1$, and $\varepsilon_2 \in [1, 1.2]$. The sum of initial densities is kept at 0.3 and the values of X_{10} and X_{20} are changed in a linear fashion. 1000 different combinations of parameters are used and 50 different $X_{10}:X_{20}$ simulations are performed with each parameter combination to calculate the probability of coexistence.

Fig. S2f is generated using the following model, which is based on the symmetric model presented in equation III.42-45, while adding asymmetry:

$$\frac{dX_1}{d\tau} = \frac{1}{\varepsilon_1} X_1 (1 - X_1 - X_2 - X_{1c} - X_{2c}) - \frac{\delta_1}{\beta_2 X_2 + 1} X_1 - \varphi X_1 \quad (\text{IV. 3})$$

$$\frac{dX_2}{d\tau} = \frac{1.3}{\varepsilon_2} X_2 (1 - X_1 - X_2 - X_{1c} - X_{2c}) - \frac{\delta_2}{\beta_1 X_1 + 1} X_2 - \varphi X_2 \quad (\text{IV. 4})$$

$$\frac{dX_{1c}}{d\tau} = X_{1c} (1 - X_1 - X_2 - X_{1c} - X_{2c}) - \frac{\delta_1}{\beta_2 X_2 + 1} X_{1c} + \varphi X_1 \quad (\text{IV. 5})$$

$$\frac{dX_{2c}}{d\tau} = 1.3 X_{2c} (1 - X_1 - X_2 - X_{1c} - X_{2c}) - \frac{\delta_2}{\beta_1 X_1 + 1} X_{2c} + \varphi X_2 \quad (\text{IV. 6})$$

Where $\varphi = 10^{-3}$, $\beta_1 = 4$, $\beta_2 \in [40, 60]$, $\delta_1 = 1.8$, $\delta_2 \in [1.4, 1.5]$, $\varepsilon_1 = 1.4$, $\varepsilon_2 \in [1, 1.2]$. The initial densities are $X_{10} = 0.05$, $X_{10} = 0.05$, $X_{1c} = 0$ and $X_{2c} = 0$.

B. The predictive accuracy is maintained in the presence of noise

In addition to investigating the probability of coexistence when initial density is randomly distributed (Fig. S2a, e), we also explicitly modeled noise to test the effect of noise on the prediction accuracy of the criterion. We added Gaussian noise to our base model (η):

$$\frac{dX_1}{d\tau} = \frac{1}{\varepsilon} X_1 (1 - X_1) - \frac{\delta}{\beta X_2 + 1} X_1 + \eta, \quad (\text{IV. 7})$$

$$\frac{dX_2}{d\tau} = \frac{1}{\varepsilon} X_2 (1 - X_2) - \frac{\delta}{\beta X_1 + 1} X_2 + \eta. \quad (\text{IV. 8})$$

In Fig. S2g, we used the model above and the criterion tested is III.19. Multiple sets of initial densities are tested. Although we observed a decreasing accuracy with increasing noise, the prediction accuracy is robustly maintained (above 90%) even when the standard deviation of noise reaches 50% of signal strength. The high maintenance of prediction accuracy is due to the stability of the mutualistic model. Because the two stable steady states exist on opposite sides of the separatrix, only when noise pushes densities across the separatrix (against the vector field), the outcome of the system will be altered. Otherwise, noise will not influence the final steady state.

V. Calibration using SVM

We chose SVM because it is a well-established algorithm and it only requires information of the support vectors which are usually data points along the boundary to obtain a classification boundary. Note that other algorithms can also be implemented for the same purpose. For example, when predicting probability of coexistence, logistic regression can be more suitable to directly predict probability of coexistence. See the Supplementary Materials for our MATLAB program which performs the calibration procedure using SVM.

A. Kernels

We used four types of kernels to cover different general shapes of the B landscape:

$$\text{Linear kernel: } K_v(\mathbf{v}_1, \mathbf{v}_2) = \mathbf{v}_1 \cdot \mathbf{v}_2 + k_\delta(\delta_1 \cdot \delta_2) \quad (\text{V. 1})$$

$$\text{Quadratic kernel: } K_v(\mathbf{v}_1, \mathbf{v}_2) = (\mathbf{v}_1 \cdot \mathbf{v}_2 + k_x)^2 + k_\delta(\delta_1 \cdot \delta_2) \quad (\text{V. 2})$$

$$\text{Cubic kernel: } K_v(\mathbf{v}_1, \mathbf{v}_2) = (\mathbf{v}_1 \cdot \mathbf{v}_2 + k_x)^3 + k_\delta(\delta_1 \cdot \delta_2) \quad (\text{V. 3})$$

$$\text{Sigmoidal kernel: } K_v(\mathbf{v}_1, \mathbf{v}_2) = \frac{(\mathbf{v}_1 \cdot \mathbf{v}_2)^2}{\|\mathbf{v}_1 - \mathbf{v}_2\|^2 + k_1} + k_\delta(\delta_1 \cdot \delta_2) \quad (\text{V. 4})$$

Where in equation V.4, $\|\mathbf{v}_1 - \mathbf{v}_2\|^2 = \sum_i (v_{1i} - v_{2i})^2$.

These four kernels follow the same structure:

$$K([\mathbf{v}_1, \delta_1], [\mathbf{v}_2, \delta_2]) = K_v \langle \mathbf{v}_1, \mathbf{v}_2 \rangle + k_\delta (\delta_1 \cdot \delta_2),$$

Our procedure allows any kernel structure that is supplied by the user, so other customized kernels can also be used for the calibration.

B. Standardizing input data

Standardizing input data is essential for robust learning of $B(\mathbf{v})$. Before training the model with SVM, we first standardize the system variables v and stress δ to mean 0 and variance 1.

$$v^{norm} = \frac{v - \text{mean}(v)}{\text{var}(v)} \quad (\text{V.5})$$

$$\delta^{norm} = \frac{\delta - \text{mean}(\delta)}{\text{var}(\delta)} \quad (\text{V.6})$$

C. From SVM output to calibrated $B(\mathbf{v})$

SVM will output a predictor that separates coexistence and collapse indicated in the \mathbf{Y} vector. Our program requires that the two classes are represented by 1 and -1 in the \mathbf{Y} vector. Using kernels V.1-4, we can write the predictor in the general form:

$$f([\mathbf{v}, \delta]) = \sum_i \alpha_i y_i K_x \langle \mathbf{v}_i, \mathbf{v} \rangle + k_\delta \delta \sum_i \alpha_i y_i \delta_i + \lambda_0 \quad (\text{V.7})$$

where k_δ is a kernel parameter; α_i represents coefficients associated with each input observation that are optimized by the SVM algorithm; λ_0 is the bias that is optimized by the algorithm. If we impose $B = \delta$ at $f([\mathbf{v}, \delta]) = 0$, we can obtain a function of B in terms of input data:

$$B(\mathbf{v}) = \frac{-\sum_i \alpha_i y_i K_x \langle \mathbf{v}_i, \mathbf{v} \rangle - \lambda_0}{k_\delta \sum_i \alpha_i y_i \delta_i} \quad (\text{V.8})$$

Using this function, we can obtain a value of B with any system variable dictated by \mathbf{v} based on the limited number of observations used as inputs.

$B(\mathbf{v})$ can sometimes have wrong directionality, meaning that $B/\delta > 1$ is associated with collapse and $B/\delta < 1$ is associated with coexistence. We identify these cases by calculating

$$\sum_1^n y_i \cdot \text{sign}(B_i - \delta_i) \quad (\text{V.9})$$

If the above expression is negative (when the trained boundary have a learning accuracy greater than 50%), the calibrated $B(\mathbf{v})$ has a wrong directionality. Flipping of the B landscape is then required, and it is done by assuming:

$$-\left(\frac{B}{\delta} - 1\right) = \frac{B^{flipped}}{\delta} - 1 \quad (\text{V.10})$$

Thus,

$$B^{flipped} = 2\delta - B \quad (\text{V.11})$$

We name the output B landscape with properly adjusted directionality B' . This landscape B' describes the shape of the final calibrated $B(\mathbf{v})$ but still needs to be rescaled for prediction purpose. The final $B(\mathbf{v})$ is calculated by rescaling it back according to the mean and variance of the δ measurements.

$$B = B' \cdot \text{var}(\delta) + \text{mean}(\delta) \quad (\text{V. 12})$$

D. Cross validation and bootstrapping

All the cross validations in this work are 10-fold cross validations. The cross-validation accuracy is represented by the average values. Bootstrapping is used to evaluate the degree of variation of quantified B . The same number of data points as the input data is randomly sampled from the input data with replacement. We performed bootstrapping for 500 times. The variance and relative standard deviation (RSD) are then calculated based on the 500 bootstrapped $B(\mathbf{v})$ quantified with 500 sets of sampled training data. The mean cross validation accuracy, the bootstrapped variance and relative standard deviation are then used to evaluate the accuracy of the $B(\mathbf{v})$ outputs.

E. Twenty sets of simulations to establish and test the calibration procedure

To establish the calibration process, we first developed the method with simulated data where the true $B(\mathbf{v})$ are known. This allows us to evaluate calibration results against the true function. We used the model presented in equation III.22-23 to conduct this set of analyses. Using X_1 as the reference population, the $B(\mathbf{v})$ can be expressed as:

$$B(\boldsymbol{\theta}) = \frac{\rho \left(\beta_1 + \frac{1}{\theta_1} + 1 \right)^2}{4\beta_1\epsilon_1 \left(\frac{\rho}{\theta_1} + \theta_2 \right)}, \theta_1 = \frac{\beta_2}{\beta_1}, \theta_2 = \frac{\epsilon_2\delta_2}{\epsilon_1\delta_1} \quad (\text{V. 13})$$

However, depending on how system variables change the model parameters ($\boldsymbol{\theta}(\mathbf{v})$) in the system variable space, the true $B(\boldsymbol{\theta}(\mathbf{v}))$ can exhibit diverse shapes. With $v_1, v_2 \in [0,1]$, we constructed 20 arbitrary set of equations with different underlying functions. We also made sure the generated data have around 50:50 split of coexistence and collapse in the simulated results to reduce bias in the input data. This 50:50 split constraint also applies to the all other simulated data. Note that these functions are arbitrary and only serve the purpose of generating true $B(\mathbf{v})$ that has various underlying functions. The data generated using these 20 models are included in the MATLAB code in the Supplementary Materials.

The 20 sets of $\boldsymbol{\theta}(\mathbf{v})$ are grouped into 5 types with each type containing 4 different examples. The 5 types of functions that describe $\boldsymbol{\theta}(\mathbf{v})$ include linear, quadratic, cubic, and Hill equation and another type of functions that are a mixture of the previous four types. As examples, the following are four sets of $\boldsymbol{\theta}(\mathbf{v})$ used to generate simulation results shown in Fig. S4b.

	Linear	Quadratic
$\beta_1 =$	$10v_1 - 2v_2 + 7$	$-3.5(v_1 + 0.2)^2 - 5(v_2 - 0.2)^2 + 9$
$\beta_2 =$	$-5v_1 + 18v_2 + 6$	$4(v_1 - 0.2)^2 + 5(v_2 - 0.2)^2 + 6.25$
$\delta_1 =$	$0.2v_1 + 0.4v_2 + 1.0$	$0.07(v_1 - 0.2)^2 + 0.05(v_2 + 0.2)^2 + 1.09$
$\delta_2 =$	$-0.5v_1 + 2v_2 + 1.7$	$0.05(v_1 + 0.1)^2 + 0.1(v_2 + 0.2)^2 + 1.3$

$\varepsilon_1 =$	$0.2v_1 + 0.1v_2 + 1$	$0.05(v_1 + 0.5)^2 + 0.2(v_2)^2 + 1.1$
$\varepsilon_2 =$	$0.05v_1 + 0.1v_2 + 1$	$0.18(v_1 + 0.2)^2 + 0.6(v_2 + 0.1)^2 + 0.9$

	Cubic
$\beta_1 =$	$0.2v_1^3 - 5(v_1 - 1)^2 + 5 + 0.2v_2^3 - 5(v_2 + 0.2)^2 + 10$
$\beta_2 =$	$0.16v_1^3 + 4(v_1 + 0.5)^2 - 0.1v_2^3 - 0.25(v_2)^2 + 3.5$
$\delta_1 =$	$0.28v_1^3 + 0.14(v_1)^2 - 0.12v_2^3 - 0.06(v_1 - 0.2)^2 + 1.2$
$\delta_2 =$	$0.04v_1^3 + 0.12(v_1 + 0.1)^2 + 0.01v_2^3 - 0.25(v_2 + 0.2)^2 + 4.95$
$\varepsilon_1 =$	$0.16v_1^3 + 0.02(v_1 - 0.3)^2 + 0.003v_2^3 + 0.06(v_2 + 0.2)^2 + 1.22$
$\varepsilon_2 =$	$-0.28v_1^3 + 0.14(v_1)^2 + 0.28 + 0.12v_2^3 - 0.06(v_2 - 0.2)^2 + 0.9$

	Hill equation
$\beta_1 =$	$10 \frac{v_1^2}{v_1^2 + 0.5^2} - 6 \frac{v_2}{v_2 + 1} + 12$
$\beta_2 =$	$2 \frac{v_1}{v_1 + 1} - 5 \frac{v_2}{v_2 + 1} + 9$
$\delta_1 =$	$0.4 \frac{v_1^3}{v_1^3 + 0.5^3} + 0.2 \frac{v_2}{v_2 + 1} + 1.01$
$\delta_2 =$	$0.1 \frac{v_1^2}{v_1^2 + 1^2} + 0.05 \frac{v_2^2}{v_2^2 + 0.3^2} + 1.2$
$\varepsilon_1 =$	$0.1 \frac{v_1^5}{v_1^5 + 0.8^5} + 0.2 \frac{v_2^4}{v_2^4 + 1^4} + 1$
$\varepsilon_2 =$	$0.2 \frac{v_1^5}{v_1^5 + 0.5^5} + 0.1 \frac{v_2^2}{v_2^2 + 0.2^2} + 1$

The training data are 10×10 on the v_1, v_2 space. R^2 between calibrated B landscape (B_c) and the true landscape (B_t) is calculated by first performing a least square linear fit of B_t with B_c . This fitting process aims to get a linear transformation of B_c that conforms with the scale of B_t while maintaining the shape of B_c . The absolute scale of the calibrated $B(v)$ is less crucial than its shape, and an absolute scale is also challenging to obtain.

$$B'_c = aB_c + b \quad (\text{V. 14})$$

Then R^2 is then calculated by:

$$1 - \frac{\sum(B'_c - B_t)^2}{\sum(B_t - \bar{B}_t)^2} \quad (\text{V. 15})$$

F. Calibration with simulated data with unknown $B(\theta)$: an example

We want to test whether we can apply the calibration procedure to data that are generated by an arbitrary mutualism model. The specific model structure we used to generate data in Fig. 2c is:

$$\frac{dX_1}{d\tau} = \frac{1}{\varepsilon_1 \left(1 + \frac{X_2}{X_2 + 0.2}\right)} X_1(1 - X_1 - X_2) - \frac{\delta_1}{\beta_2^3 X_2^3 + 1} X_1 \quad (\text{V. 16})$$

$$\frac{dX_2}{d\tau} = \frac{1.5}{\varepsilon_2} X_2(1 - 0.8X_1 - X_2) - \frac{\delta_2}{\beta_1^2 X_1^2 + 1} X_2 \quad (\text{V. 17})$$

The analytical solution of this model cannot be expressed in a simple and explicit form. However, using simulated data, we can obtain an empirical function of B on the system variable space that allow further prediction in the system variable space. The simulations are done using initial densities of $[0.01, 0.01]$ for 10 unit-time.

To get the probability of coexistence, we ran the model 100 times with varying ratios of initial density while keeping the total initial density constant as 0.02. The density for a population is varied in a linear fashion from 0 to 0.02 and we terminated the simulation after 10 unit-time.

To calculate how well the systems can resist cheater exploitation, we modified the equations V. 16 – 17 to account for the emergence of cheaters:

$$\frac{dX_1}{d\tau} = \frac{1}{\varepsilon_1 \left(1 + \frac{X_2}{X_2 + 0.2}\right)} X_1(1 - X_1 - X_2 - X_{1c} - X_{2c}) - \frac{\delta_1}{\beta_2^3 X_2^3 + 1} X_1 - \varphi X_1 \quad (\text{V. 18})$$

$$\frac{dX_2}{d\tau} = \frac{1.5}{\varepsilon_2} X_2(1 - 0.8X_1 - X_2 - X_{1c} - X_{2c}) - \frac{\delta_2}{\beta_1^2 X_1^2 + 1} X_2 - \varphi X_2 \quad (\text{V. 19})$$

$$\frac{dX_{1c}}{d\tau} = X_1(1 - X_1 - X_2 - X_{1c} - X_{2c}) - \frac{\delta_1}{\beta_2^3 X_2^3 + 1} X_1 + \varphi X_1 \quad (\text{V. 20})$$

$$\frac{dX_{2c}}{d\tau} = 1.5 X_2(1 - 0.8X_1 - X_2 - X_{1c} - X_{2c}) - \frac{\delta_2}{\beta_1^2 X_1^2 + 1} X_2 + \varphi X_2 \quad (\text{V. 21})$$

The simulations were done with $\varphi = 10^{-3}$. The initial densities are $[0.01, 0.01, 0, 0]$ for X_1, X_2, X_{1c} , and X_{2c} , respectively. The simulations were terminated after 1300 unit-time.

VI. Framework generality verified by complex mutualistic systems

A. A 5-mutualist system

The data presented in Fig. 4a and Fig. S9a are generated with the following mutualism model based on the general model structure presented in equation III.39:

$$\frac{dX_1}{d\tau} = \frac{1}{\varepsilon_1} X_1 \left(1 - \sum_{k=1}^5 X_k\right) - \frac{\delta_1}{\sum_{k \neq 1}^n \beta_k X_k + 1} X_1 \quad (\text{VI. 1})$$

$$\frac{dX_2}{d\tau} = \frac{1.1}{\varepsilon_2} X_2 \left(1 - \sum_{k=1}^5 X_k\right) - \frac{\delta_2}{\sum_{k \neq 2}^n \beta_k X_k + 1} X_2 \quad (\text{VI. 2})$$

$$\frac{dX_3}{d\tau} = \frac{1.2}{\varepsilon_3} X_3 \left(1 - \sum_{k=1}^5 X_k\right) - \frac{\delta_3}{\sum_{k \neq 3}^n \beta_k X_k + 1} X_3 \quad (\text{VI. 3})$$

$$\frac{dX_4}{d\tau} = \frac{1.3}{\varepsilon_4} X_4 \left(1 - \sum_{k=1}^5 X_k \right) - \frac{\delta_4}{\sum_{k \neq i}^n \beta_k X_k + 1} X_4 \quad (\text{VI. 4})$$

$$\frac{dX_5}{d\tau} = \frac{1.4}{\varepsilon_5} X_5 \left(1 - \sum_{k=1}^5 X_k \right) - \frac{\delta_5}{\sum_{k \neq i}^n \beta_k X_k + 1} X_5 \quad (\text{VI. 5})$$

All the parameters are linear combinations of $v_1, v_2 \in [0,1]$ and the coefficients are randomly generated. The initial densities are $[0.001, 0.001, 0.001, 0.001, 0.001]$ for X_1 to X_5 respectively. The simulations were terminated after 5 unit-time.

B. The experimental 3-member mutualistic systems

The calibration shown in Fig. 4b is done on all 384 systems with triplicates that alternatively use each strain in a system as the reference strain. The δ measurements used in this calibration is the same as the measurements presented in Fig. 3j and Fig. S8b with the pairwise systems.

C. A mutualistic system in an oscillatory environment

The oscillatory signal is implemented as square pulses, where

$$S = \begin{cases} I, & [\tau/L] \text{ is odd} \\ 0, & [\tau/L] \text{ is even} \end{cases} \quad (\text{VI. 6})$$

τ is time; L is the duration of the pulses and the duration between the end of one pulse and the start of the next pulse; I is the intensity of the signal. Throughout the simulation, δ_1 , δ_2 , β_1 , and β_2 are affected by this oscillating signal, where

$$\delta_1 = \delta_2 = 0.8 + S \quad (\text{VI. 7})$$

$$\beta_1 = \beta_2 = 2(1 + S) \quad (\text{VI. 8})$$

To simulate the data, we used a logarithm scale that vary L from 10^{-2} to 10^2 and vary I from 10^{-1} to 10^1 . The model used to generate the set of data in Fig. 4c and the left column of Fig. S10 is presented in equation III.22-III.23, where $\rho = 1.5$. The initial densities used are $[0.2, 0.2]$ for X_1 and X_2 respectively. The simulations are performed for 500 unit-time. The theoretical $B(\theta(\mathbf{v}))$ is calculated using III.32 with $\delta_1 = \delta_2 = 0.8 + I$ and $\beta_1 = \beta_2 = 2(1 + I)$.

D. A mutualistic system cohabiting with other populations

To generate simulated data presented in Fig. 4d and the right column of Fig. S10, we used the following arbitrary 7-population model, where X_1 and X_2 are mutualistic and their population dynamics are modulated by the other 5 populations.

$$\frac{dX_1}{d\tau} = \frac{1}{\varepsilon} X_1 \left(1 - \sum X_i \right) - \frac{\delta}{\beta X_2 + \beta_{51} X_5 + 1} X_1 - X_6 X_1 \quad (\text{VI. 9})$$

$$\frac{dX_2}{d\tau} = \rho \frac{1}{\varepsilon} X_2 \left(1 - \sum X_i \right) - \frac{\delta}{\beta X_1 + \beta_{62} X_6 + 1} X_2 - X_5 X_2 \quad (\text{VI. 10})$$

$$\frac{dX_3}{d\tau} = X_3 \left(1 - \sum X_i \right) - \frac{\delta_3}{\beta_{13} X_1 + \beta_{23} X_2 + 1} X_3 \quad (\text{VI. 11})$$

$$\frac{dX_4}{d\tau} = X_4 \left(1 - \sum X_i \right) - \frac{\delta_4}{\beta_{34}X_3 + \beta_{23}X_2 + 1} X_4 \quad (\text{VI. 12})$$

$$\frac{dX_5}{d\tau} = X_5 \left(1 - 2 \sum X_i \right) \quad (\text{VI. 13})$$

$$\frac{dX_6}{d\tau} = X_6 \left(1 - 2 \sum X_i \right) \quad (\text{VI. 14})$$

$$\frac{dX_7}{d\tau} = X_7 \left(1 - 3.3 \sum X_i \right) + \frac{X_1}{X_1 + 0.5} X_7 - X_2 X_7 \quad (\text{VI. 15})$$

In the simulation, only δ and β are modulated by system variables $v_1, v_2 \in [0,1]$, while other parameters were kept constants.

$$\delta = 0.5v_1 + 1.2v_2 + 0.8 \quad (\text{VI. 16})$$

$$\beta = -15v_1 - 10v_2 + 30 \quad (\text{VI. 17})$$

The initial densities are [1, 1, 1, 1, 1, 1, 1] for X_1 to X_7 respectively. The simulations were performed for 5000 unit-time. The theoretical $B(\theta(\mathbf{v}))/\delta$ in the right panel of Fig. S10a is directly calculated using III.32.

References

66. F. K. Balagadde *et al.*, A synthetic Escherichia coli predator-prey ecosystem. *Molecular Systems Biology* **4**, (2008).
67. H. Song, S. Payne, M. Gray, L. C. You, Spatiotemporal modulation of biodiversity in a synthetic chemical-mediated ecosystem. *Nature Chemical Biology* **5**, 929-935 (2009).
68. R. Lutz, H. Bujard, Independent and tight regulation of transcriptional units in Escherichia coli via the LacR/O, the TetR/O and AraC/I-1-I-2 regulatory elements. *Nucleic acids research* **25**, 1203-1210 (1997).
69. J. Bascompte, P. Jordano, J. M. Olesen, Asymmetric coevolutionary networks facilitate biodiversity maintenance. *Science* **312**, 431-433 (2006).
70. S. Allesina, S. Tang, Stability criteria for complex ecosystems. *Nature* **483**, 205-208 (2012).
71. K. Z. Coyte, J. Schluter, K. R. Foster, The ecology of the microbiome: Networks, competition, and stability. *Science* **350**, 663-666 (2015).
72. R. M. May, Citation Classic - Stability and Complexity in Model-Ecosystems. *Cc/Agr Biol Environ*, 22-22 (1988).
73. D. L. DeAngelis, W. M. Post, C. C. Travis, *Positive feedback in natural systems*. Biomathematics (Springer-Verlag, Berlin ; New York, 1986), pp. xii, 290 p.
74. A. Mougi, M. Kondoh, Diversity of Interaction Types and Ecological Community Stability. *Science* **337**, 349-351 (2012).
75. E. Thebault, C. Fontaine, Stability of Ecological Communities and the Architecture of Mutualistic and Trophic Networks. *Science* **329**, 853-856 (2010).
76. J. J. Stachowicz, Mutualism, facilitation, and the structure of ecological communities. *Bioscience* **51**, 235-246 (2001).
77. U. Bastolla *et al.*, The architecture of mutualistic networks minimizes competition and increases biodiversity. *Nature* **458**, 1018-1020 (2009).
78. J. F. Addicott, Stability properties of 2-species models of mutualism: Simulation studies. *Oecologia* **49**, 42-49 (1981).
79. B. Momeni, L. Xie, W. Shou, Lotka-Volterra pairwise modeling fails to capture diverse pairwise microbial interactions. *eLIFE*, (2017).
80. J. N. Holland, D. L. DeAngelis, Interspecific population regulation and the stability of mutualism: fruit abortion and density-dependent mortality of pollinating seed-eating insects. *Oikos* **113**, 563-571 (2006).
81. J. N. Holland, D. L. DeAngelis, J. L. Bronstein, Population dynamics and mutualism: Functional responses of benefits and costs. *Am Nat* **159**, 231-244 (2002).
82. J. Garcia-Algarra, J. Galeano, J. M. Pastor, J. M. Iriando, J. J. Ramasco, Rethinking the logistic approach for population dynamics of mutualistic interactions. *Journal of theoretical biology* **363**, 332-343 (2014).
83. J. N. Holland, D. L. DeAngelis, A consumer-resource approach to the density-dependent population dynamics of mutualism. *Ecology* **91**, 1286-1295 (2010).
84. M. L. Stanton, Interacting guilds: Moving beyond the pairwise perspective on mutualisms. *Am Nat* **162**, S10-S23 (2003).
85. R. Munns, Comparative physiology of salt and water stress. *Plant Cell and Environment* **25**, 239-250 (2002).
86. J. Schimel, T. C. Balser, M. Wallenstein, Microbial stress-response physiology and its implications for ecosystem function. *Ecology* **88**, 1386-1394 (2007).
87. O. Hoegh-Guldberg *et al.*, Coral reefs under rapid climate change and ocean acidification. *Science* **318**, 1737-1742 (2007).
88. T. Finkel, N. J. Holbrook, Oxidants, oxidative stress and the biology of ageing. *Nature* **408**, 239-247 (2000).

89. T. A. Hoek *et al.*, Resource Availability Modulates the Cooperative and Competitive Nature of a Microbial Cross-Feeding Mutualism. *PLoS biology* **14**, (2016).
90. B. E. Brown, Coral bleaching: causes and consequences. *Coral Reefs* **16**, S129-S138 (1997).
91. N. S. Webster, R. E. Cobb, A. P. Negri, Temperature thresholds for bacterial symbiosis with a sponge. *Isme Journal* **2**, 830-842 (2008).
92. M. Holcomb, D. C. McCorkle, A. L. Cohen, Long-term effects of nutrient and CO₂ enrichment on the temperate coral *Astrangia poculata* (Ellis and Solander, 1786). *J Exp Mar Biol Ecol* **386**, 27-33 (2010).
93. D. Cipollini, C. M. Rigsby, E. K. Barto, Microbes as Targets and Mediators of Allelopathy in Plants. *J Chem Ecol* **38**, 714-727 (2012).
94. J. L. Bronstein, *Mutualism*. (Oxford University Press, Oxford, United Kingdom, ed. First edition., 2015), pp. xv, 297 pages, 298 unnumbered pages of plates.
95. C. Neuhauser, J. E. Fargione, A mutualism-parasitism continuum model and its application to plant-mycorrhizae interactions. *Ecol Model* **177**, 337-352 (2004).
96. N. Yamamura, Evolution of mutualistic symbiosis: A differential equation model. *Res Popul Ecol* **38**, 211-218 (1996).
97. W. F. Morris, J. L. Bronstein, W. G. Wilson, Three-way coexistence in obligate mutualist-exploiter interactions: The potential role of competition. *Am Nat* **161**, 860-875 (2003).
98. A. Kerner, J. Park, A. Williams, X. N. Lin, A programmable *Escherichia coli* consortium via tunable symbiosis. *PloS one* **7**, e34032 (2012).
99. R. P. Rohr, S. Saavedra, J. Bascompte, On the structural stability of mutualistic systems. *Science* **345**, 416-+ (2014).
100. W. F. Morris, D. P. Vazquez, N. P. Chacoff, Benefit and cost curves for typical pollination mutualisms. *Ecology* **91**, 1276-1285 (2010).
101. M. D. Bertness, G. H. Leonard, The role of positive interactions in communities: Lessons from intertidal habitats. *Ecology* **78**, 1976-1989 (1997).
102. T. M. Palmer *et al.*, Breakdown of an ant-plant mutualism follows the loss of large herbivores from an African Savanna. *Science* **319**, 192-195 (2008).

Effect of cost (ε) is independent of partner density

$\frac{dX_1}{dt} = (1 - \delta_1)X_1(1 - X_1)$		
1. $\frac{dX_1}{dt} = (1 - \delta\varepsilon \frac{1}{\beta X_2 + 1})X_1(1 - X_1)$ $\frac{\beta X_0 + 1}{\varepsilon} > \delta$	2. $\frac{dX_1}{dt} = (1 - \delta) \frac{1}{\varepsilon} X_1 \left(1 + \frac{\beta' X_2}{X_2 + 1/\beta} - X_1 \right)$ Violates assumption d) and unbounded growth	3. $\frac{dX_1}{dt} = (1 - \delta\varepsilon)X_1(1 - X_1) + \frac{\beta' X_2}{X_2 + 1/\beta} X_1$ Unbounded growth
4. $\frac{dX_1}{dt} = (1 - \delta \frac{1}{\beta X_2 + 1})X_1 \left(\frac{1}{\varepsilon} - X_1 \right)$ $\beta X_0 + 1 > \delta$	5. $\frac{dX_1}{dt} = (1 - \delta)X_1 \left(1 + \frac{1}{\varepsilon} \frac{\beta' X_2}{X_2 + 1/\beta} - X_1 \right)$ Violates assumption d) and unbounded growth	6. $\frac{dX_1}{dt} = (1 - \delta)X_1 \left(\frac{1}{\varepsilon} - X_1 \right) + \frac{\beta' X_2}{X_2 + 1/\beta} X_1$ Violates assumption d) and unbounded growth
7. $\frac{dX_1}{dt} = (1 - \delta \frac{1}{\beta X_2 + 1})X_1(1 - X_1) - \varepsilon X_1$ $(\sqrt{\beta + 1} - \sqrt{\beta\varepsilon})^2 > \delta$	8. $\frac{dX_1}{dt} = (1 - \delta)X_1 \left(1 + \frac{\beta' X_2}{X_2 + 1/\beta} - X_1 \right) - \varepsilon X_1$ Violates assumption d)	9. $\frac{dX_1}{dt} = (1 - \delta)X_1(1 - X_1) + \frac{1}{\varepsilon} \frac{\beta' X_2}{X_2 + 1/\beta} X_1$ Unbounded growth
$\frac{dX_1}{dt} = X_1((1 - \delta_1) - X_1)$		
10. $\frac{dX_1}{dt} = \frac{1}{\varepsilon} \frac{\beta' X_2}{X_2 + 1/\beta} X_1((1 - \delta) - X_1)$ Violates assumption d)	11. $\frac{dX_1}{dt} = \frac{1}{\varepsilon} X_1 \left((1 - \delta \frac{1}{\beta X_2 + 1}) - X_1 \right)$ $\frac{(\beta + 1)^2}{4\beta} > \delta$	12. $\frac{dX_1}{dt} = \frac{1}{\varepsilon} X_1(1 - \delta - X_1) + \frac{\beta' X_2}{X_2 + 1/\beta} X_1$ Violates assumption d)
13. $\frac{dX_1}{dt} = \frac{\beta' X_2}{X_2 + 1/\beta} X_1 \left((1 - \delta) \frac{1}{\varepsilon} - X_1 \right)$ Violates assumption d)	14. $\frac{dX_1}{dt} = X_1 \left((1 - \delta \frac{1}{\beta X_2 + 1}) \frac{1}{\varepsilon} - X_1 \right)$ Violates assumption d)	15. $\frac{dX_1}{dt} = X_1 \left(\frac{1}{\varepsilon} - \delta - X_1 \right) + \frac{\beta' X_2}{X_2 + 1/\beta} X_1$ $\frac{(\sqrt{\beta'\varepsilon} - \sqrt{\varepsilon})^2 + \beta}{\beta\varepsilon} > \delta$
16. $\frac{dX_1}{dt} = \frac{\beta' X_2}{X_2 + 1/\beta} X_1((1 - \delta) - X_1) - \varepsilon X_1$ Violates assumption d)	17. $\frac{dX_1}{dt} = X_1 \left((1 - \delta \frac{1}{\beta X_2 + 1}) - X_1 \right) - \varepsilon X_1$ $\frac{(\beta - \beta\varepsilon + 1)^2}{4\beta} > \delta$	18. $\frac{dX_1}{dt} = X_1(1 - \delta - X_1) + \frac{1}{\varepsilon} \frac{\beta' X_2}{X_2 + 1/\beta} X_1$ $\frac{(\sqrt{\beta'\varepsilon} - \sqrt{\varepsilon})^2 + \beta\varepsilon}{\beta\varepsilon} > \delta$
$\frac{dX_1}{dt} = X_1(1 - X_1) - \delta_1 X_1$		
19. $\frac{dX_1}{dt} = \frac{1}{\varepsilon} \frac{\beta' X_2}{X_2 + 1/\beta} X_1(1 - X_1) - \delta X_1$ $\frac{2\beta' - 2\beta'\sqrt{\beta + 1} + \beta\beta'}{\beta\varepsilon} > \delta$	20. $\frac{dX_1}{dt} = \frac{1}{\varepsilon} X_1 \left(\frac{\beta' X_2}{X_2 + 1/\beta} - X_1 \right) - \delta X_1$ $\frac{(\sqrt{\beta'\varepsilon} - 1)^2 + \beta}{\beta\varepsilon} > \delta$	21. $\frac{dX_1}{dt} = \frac{1}{\varepsilon} X_1(1 - X_1) - \delta \frac{1}{\beta X_2 + 1} X_1$ $\frac{(\beta + 1)^2}{4\beta\varepsilon} > \delta$
22. $\frac{dX_1}{dt} = \frac{\beta' X_2}{X_2 + 1/\beta} X_1 \left(\frac{1}{\varepsilon} - X_1 \right) - \delta X_1$ $\frac{2\beta'\varepsilon - 2\beta'\sqrt{\beta\varepsilon + \varepsilon^2} + \beta\beta'}{\beta\varepsilon} > \delta$	23. $\frac{dX_1}{dt} = X_1 \left(\frac{1}{\varepsilon} \frac{\beta' X_2}{X_2 + 1/\beta} - X_1 \right) - \delta X_1$ $\frac{(\sqrt{\beta'\varepsilon} - \sqrt{\varepsilon})^2 + \beta\varepsilon}{\beta\varepsilon} > \delta$	24. $\frac{dX_1}{dt} = X_1 \left(\frac{1}{\varepsilon} - X_1 \right) - \delta \frac{1}{\beta X_2 + 1} X_1$ $\frac{(\beta + \varepsilon)^2}{4\beta\varepsilon^2} > \delta$
25. $\frac{dX_1}{dt} = \frac{\beta' X_2}{X_2 + 1/\beta} X_1(1 - X_1) - \delta\varepsilon X_1$ $\frac{2\beta' - 2\beta'\sqrt{\beta + 1} + \beta\beta'}{\beta\varepsilon} > \delta$	26. $\frac{dX_1}{dt} = X_1 \left(\frac{\beta' X_2}{X_2 + 1/\beta} - X_1 \right) - \delta\varepsilon X_1$ $\frac{(\sqrt{\beta'\varepsilon} - 1)^2 + \beta}{\beta\varepsilon} > \delta$	27. $\frac{dX_1}{dt} = X_1(1 - X_1) - \delta\varepsilon \frac{1}{\beta X_2 + 1} X_1$ $\frac{(\beta + 1)^2}{4\beta\varepsilon} > \delta$

Sharing carrying capacity

$\frac{dX_1}{dt} = (1 - \delta_1)X_1(1 - X_1)$		
28. $\frac{dX_1}{dt} = (1 - \delta\epsilon \frac{1}{\beta X_2 + 1})X_1(1 - X_1 - X_2)$ $\frac{\beta X_0 + 1}{\epsilon} > \delta$	29. $\frac{dX_1}{dt} = (1 - \delta)\frac{1}{\epsilon}X_1\left(\frac{\beta'X_2}{X_2 + 1/\beta} - X_1 - X_2\right)$ Violates assumption d) and unbounded growth	30. $\frac{dX_1}{dt} = (1 - \delta\epsilon)X_1(1 - X_1 - X_2) + \frac{\beta'X_2}{X_2 + 1/\beta}X_1$ Unbounded growth
31. $\frac{dX_1}{dt} = (1 - \delta\frac{1}{\beta X_2 + 1})X_1\left(\frac{1}{\epsilon} - X_1 - X_2\right)$ $\beta X_0 + 1 > \delta$	32. $\frac{dX_1}{dt} = (1 - \delta)X_1\left(\frac{1}{\epsilon} \frac{\beta'X_2}{X_2 + 1/\beta} - X_1 - X_2\right)$ Violates assumption d) and unbounded growth	33. $\frac{dX_1}{dt} = (1 - \delta)X_1\left(\frac{1}{\epsilon} - X_1 - X_2\right) + \frac{\beta'X_2}{X_2 + 1/\beta}X_1$ Violates assumption d) and unbounded growth
34. $\frac{dX_1}{dt} = (1 - \delta\frac{1}{\beta X_2 + 1})X_1(1 - X_1 - X_2) - \epsilon X_1$ $\frac{1}{2}(\sqrt{\beta + 2} - \sqrt{\beta\epsilon})^2 > \delta$	35. $\frac{dX_1}{dt} = (1 - \delta)X_1\left(\frac{\beta'X_2}{X_2 + 1/\beta} - X_1 - X_2\right) - \epsilon X_1$ Violates assumption d)	36. $\frac{dX_1}{dt} = (1 - \delta)X_1(1 - X_1 - X_2) + \frac{1}{\epsilon} \frac{\beta'X_2}{X_2 + 1/\beta}X_1$ Unbounded growth
$\frac{dX_1}{dt} = X_1((1 - \delta_1) - X_1)$		
37. $\frac{dX_1}{dt} = \frac{1}{\epsilon} \frac{\beta'X_2}{X_2 + 1/\beta}X_1((1 - \delta) - X_1 - X_2)$ Violates assumption d)	38. $\frac{dX_1}{dt} = \frac{1}{\epsilon}X_1\left((1 - \delta\frac{1}{\beta X_2 + 1}) - X_1 - X_2\right)$ $\frac{(\beta + 2)^2}{4\beta} > \delta$	39. $\frac{dX_1}{dt} = \frac{1}{\epsilon}X_1(1 - \delta - X_1 - X_2) + \frac{\beta'X_2}{X_2 + 1/\beta}X_1$ Violates assumption d)
40. $\frac{dX_1}{dt} = \frac{\beta'X_2}{X_2 + 1/\beta}X_1\left((1 - \delta)\frac{1}{\epsilon} - X_1 - X_2\right)$ Violates assumption d)	41. $\frac{dX_1}{dt} = X_1\left((1 - \delta\frac{1}{\beta X_2 + 1})\frac{1}{\epsilon} - X_1 - X_2\right)$ Violates assumption d)	42. $\frac{dX_1}{dt} = X_1\left(\frac{1}{\epsilon} - \delta - X_1 - X_2\right) + \frac{\beta'X_2}{X_2 + 1/\beta}X_1$ $\frac{\epsilon(\sqrt{\beta\beta'} - \sqrt{2})^2 + \beta}{\beta\epsilon} > \delta$
43. $\frac{dX_1}{dt} = \frac{\beta'X_2}{X_2 + 1/\beta}X_1((1 - \delta) - X_1 - X_2) - \epsilon X_1$ Violates assumption d)	44. $\frac{dX_1}{dt} = X_1\left((1 - \delta\frac{1}{\beta X_2 + 1}) - X_1 - X_2\right) - \epsilon X_1$ $\frac{(\beta + 2 - \beta\epsilon)^2}{8\beta} > \delta$	45. $\frac{dX_1}{dt} = X_1(1 - \delta - X_1 - X_2) + \frac{1}{\epsilon} \frac{\beta'X_2}{X_2 + 1/\beta}X_1$ $\frac{(\sqrt{\beta\beta'} - \sqrt{2\epsilon})^2 + \beta\epsilon}{\beta\epsilon} > \delta$
$\frac{dX_1}{dt} = X_1(1 - X_1) - \delta_1 X_1$		
46. $\frac{dX_1}{dt} = \frac{1}{\epsilon} \frac{\beta'X_2}{X_2 + 1/\beta}X_1(1 - X_1 - X_2) - \delta X_1$ $\frac{4\beta' - 2\beta'\sqrt{2\beta + 4} + \beta\beta'}{\beta\epsilon} > \delta$	47. $\frac{dX_1}{dt} = \frac{1}{\epsilon}X_1\left(\frac{\beta'X_2}{X_2 + 1/\beta} - X_1 - X_2\right) - \delta X_1$ $\frac{(\sqrt{\beta\beta'} - \sqrt{2})^2}{\beta\epsilon} > \delta$	48. $\frac{dX_1}{dt} = \frac{1}{\epsilon}X_1(1 - X_1 - X_2) - \delta\frac{1}{\beta X_2 + 1}X_1$ $\frac{(\beta + 2)^2}{8\beta\epsilon} > \delta$
49. $\frac{dX_1}{dt} = \frac{\beta'X_2}{X_2 + 1/\beta}X_1\left(\frac{1}{\epsilon} - X_1 - X_2\right) - \delta X_1$ $\frac{4\beta'\epsilon - 2\beta'\sqrt{2\beta\epsilon + 4\epsilon^2} + \beta\beta'}{\beta\epsilon} > \delta$	50. $\frac{dX_1}{dt} = X_1\left(\frac{1}{\epsilon} \frac{\beta'X_2}{X_2 + 1/\beta} - X_1 - X_2\right) - \delta X_1$ $\frac{(\sqrt{\beta\beta'} - \sqrt{2\epsilon})^2}{\beta\epsilon} > \delta$	51. $\frac{dX_1}{dt} = X_1\left(\frac{1}{\epsilon} - X_1 - X_2\right) - \delta\frac{1}{\beta X_2 + 1}X_1$ $\frac{(\beta + 2\epsilon)^2}{8\beta\epsilon^2} > \delta$
52. $\frac{dX_1}{dt} = \frac{\beta'X_2}{X_2 + 1/\beta}X_1(1 - X_1 - X_2) - \delta\epsilon X_1$ $\frac{4\beta' - 2\beta'\sqrt{2\beta + 4} + \beta\beta'}{\beta\epsilon} > \delta$	53. $\frac{dX_1}{dt} = X_1\left(\frac{\beta'X_2}{\beta X_2 + 1} - X_1 - X_2\right) - \delta\epsilon X_1$ $\frac{(\sqrt{\beta\beta'} - \sqrt{2})^2}{\beta\epsilon} > \delta$	54. $\frac{dX_1}{dt} = X_1(1 - X_1 - X_2) - \delta\epsilon\frac{1}{\beta X_2 + 1}X_1$ $\frac{(\beta + 2)^2}{8\beta\epsilon} > \delta$

Effect of cost (ε) increases with partner density		
$\frac{dX_1}{dt} = (1 - \delta_1)X_1(1 - X_1)$		
$55. \frac{dX_1}{dt} = (1 - \delta(1 + \varepsilon X_2)) \frac{1}{\beta X_2 + 1} X_1(1 - X_1)$ $\frac{\beta X_0 + 1}{\varepsilon X_0 + 1} > \delta$	$56. \frac{dX_1}{dt} = (1 - \delta) \frac{1}{(1 + \varepsilon X_2)} X_1 \left(\frac{\beta' X_2}{X_2 + 1/\beta} - X_1 \right)$ <p>Violates assumption d) and unbounded growth</p>	$57. \frac{dX_1}{dt} = (1 - \delta(1 + \varepsilon X_2)) X_1(1 - X_1) + \frac{\beta' X_2}{X_2 + 1/\beta} X_1$ <p>Unbounded growth</p>
$58. \frac{dX_1}{dt} = (1 - \delta \frac{1}{\beta X_2 + 1}) X_1 \left(\frac{1}{(1 + \varepsilon X_2)} - X_1 \right)$ $\beta X_0 + 1 > \delta$	$59. \frac{dX_1}{dt} = (1 - \delta) X_1 \left(\frac{1}{(1 + \varepsilon X_2)} \frac{\beta' X_2}{X_2 + 1/\beta} - X_1 \right)$ <p>Violates assumption d) and unbounded growth</p>	$60. \frac{dX_1}{dt} = (1 - \delta) X_1 \left(\frac{1}{(1 + \varepsilon X_2)} - X_1 \right) + \frac{\beta' X_2}{X_2 + 1/\beta} X_1$ <p>Violates assumption d) and unbounded growth</p>
$61. \frac{dX_1}{dt} = (1 - \delta \frac{1}{\beta X_2 + 1}) X_1(1 - X_1) - (\varepsilon X_2) X_1$ $(\sqrt{\beta \varepsilon} - \sqrt{(\varepsilon + 1)(\beta + 1)})^2 > \delta$	$62. \frac{dX_1}{dt} = (1 - \delta) X_1 \left(\frac{\beta' X_2}{X_2 + 1/\beta} - X_1 \right) - (\varepsilon X_2) X_1$ <p>Violates assumption d)</p>	$63. \frac{dX_1}{dt} = (1 - \delta) X_1(1 - X_1) + \frac{1}{(1 + \varepsilon X_2)} \frac{\beta' X_2}{X_2 + 1/\beta} X_1$ <p>Unbounded growth</p>
$\frac{dX_1}{dt} = X_1((1 - \delta_1) - X_1)$		
$64. \frac{dX_1}{dt} = \frac{1}{(1 + \varepsilon X_2)} \frac{\beta' X_2}{X_2 + 1/\beta} X_1((1 - \delta) - X_1)$ <p>Violates assumption d)</p>	$65. \frac{dX_1}{dt} = \frac{1}{(1 + \varepsilon X_2)} X_1 \left((1 - \delta \frac{1}{\beta X_2 + 1}) - X_1 \right)$ $\frac{(\beta + 1)^2}{4\beta} > \delta$	$66. \frac{dX_1}{dt} = \frac{1}{(1 + \varepsilon X_2)} X_1(1 - \delta - X_1) + \frac{\beta' X_2}{X_2 + 1/\beta} X_1$ <p>Violates assumption d)</p>
$67. \frac{dX_1}{dt} = \frac{\beta' X_2}{X_2 + 1/\beta} X_1 \left((1 - \delta) \frac{1}{(1 + \varepsilon X_2)} - X_1 \right)$ <p>Violates assumption d)</p>	$68. \frac{dX_1}{dt} = X_1 \left((1 - \delta \frac{1}{\beta X_2 + 1}) \frac{1}{(1 + \varepsilon X_2)} - X_1 \right)$ <p>Violates assumption d)</p>	$69. \frac{dX_1}{dt} = X_1 \left(\frac{1}{1 + \varepsilon X_2} - \delta - X_1 \right) + \frac{\beta' X_2}{X_2 + 1/\beta} X_1$ <p>$f(\beta, \beta', \varepsilon) > \delta$</p>
$70. \frac{dX_1}{dt} = \frac{\beta' X_2}{X_2 + 1/\beta} X_1((1 - \delta) - X_1) - (\varepsilon X_2) X_1$ <p>Violates assumption d)</p>	$71. \frac{dX_1}{dt} = X_1 \left((1 - \delta \frac{1}{\beta X_2 + 1}) - X_1 \right) - (\varepsilon X_2) X_1$ $\frac{(\beta + \varepsilon + 1)^2}{4\beta(\varepsilon + 1)} > \delta$	$72. \frac{dX_1}{dt} = X_1(1 - \delta - X_1) + \frac{1}{(1 + \varepsilon X_2)} \frac{\beta' X_2}{X_2 + 1/\beta} X_1$ <p>$f(\beta, \beta', \varepsilon) > \delta$</p>
$\frac{dX_1}{dt} = X_1(1 - X_1) - \delta_1 X_1$		
$73. \frac{dX_1}{dt} = \frac{1}{(1 + \varepsilon X_2)} \frac{\beta' X_2}{X_2 + 1/\beta} X_1(1 - X_1) - \delta X_1$ $\frac{\beta \beta' (\sqrt{(\beta + 1)} - \sqrt{(\varepsilon + 1)})^2}{(\beta - \varepsilon)^2} > \delta$	$74. \frac{dX_1}{dt} = \frac{1}{(1 + \varepsilon X_2)} X_1 \left(\frac{\beta' X_2}{X_2 + 1/\beta} - X_1 \right) - \delta X_1$ $\frac{(\sqrt{\beta^2 \beta'} - \sqrt{\beta - \varepsilon + \beta \beta' \varepsilon})^2}{(\beta - \varepsilon)^2} > \delta$	$75. \frac{dX_1}{dt} = \frac{1}{(1 + \varepsilon X_2)} X_1(1 - X_1) - \delta \frac{1}{\beta X_2 + 1} X_1$ $\frac{(\sqrt{\beta} - \sqrt{(1 + \varepsilon)(\beta - \varepsilon)})^2}{\varepsilon^2} + 1 > \delta$
$76. \frac{dX_1}{dt} = \frac{\beta' X_2}{X_2 + 1/\beta} X_1 \left(\frac{1}{(1 + \varepsilon X_2)} - X_1 \right) - \delta X_1$ <p>$f(\beta, \beta', \varepsilon) > \delta$</p>	$77. \frac{dX_1}{dt} = X_1 \left(\frac{1}{(1 + \varepsilon X_2)} \frac{\beta' X_2}{X_2 + 1/\beta} - X_1 \right) - \delta X_1$ <p>$f(\beta, \beta', \varepsilon) > \delta$</p>	$78. \frac{dX_1}{dt} = X_1 \left(\frac{1}{(1 + \varepsilon X_2)} - X_1 \right) - \delta \frac{1}{\beta X_2 + 1} X_1$ <p>$f(\beta, \beta', \varepsilon) > \delta$</p>
$79. \frac{dX_1}{dt} = \frac{\beta' X_2}{X_2 + 1/\beta} X_1(1 - X_1) - \delta(1 + \varepsilon X_2) X_1$ $\frac{\beta \beta' (\sqrt{(\beta + 1)} - \sqrt{(\varepsilon + 1)})^2}{(\beta - \varepsilon)^2} > \delta$	$80. \frac{dX_1}{dt} = X_1 \left(\frac{\beta' X_2}{X_2 + 1/\beta} - X_1 \right) - \delta(1 + \varepsilon X_2) X_1$ $\frac{(\sqrt{\beta^2 \beta'} - \sqrt{\beta - \varepsilon + \beta \beta' \varepsilon})^2}{(\beta - \varepsilon)^2} > \delta$	$81. \frac{dX_1}{dt} = X_1(1 - X_1) - \delta(1 + \varepsilon X_2) \frac{1}{\beta X_2 + 1} X_1$ $\frac{(\sqrt{\beta} - \sqrt{(1 + \varepsilon)(\beta - \varepsilon)})^2}{\varepsilon^2} + 1 > \delta$

Table S1: Diverse coexistence criteria derived from a set of 81 mutualism models. The formulations are based on the locations of benefit (β, β'), cost (ϵ), and stress (δ) in logistic growth equations (see section II of SI for detailed method). Each 3×3 block represents different locations of the δ term. Each column and row represent a location of the benefit and cost term, respectively. Because we assume symmetry between the two populations, only the equation representing X_1 is shown for simplicity of presentation. Models in the first table assume cost is independent of partner density. In natural mutualism systems, mutualists can also compete for resources, so the third table contains models that describe the two partners sharing the same carrying capacity. In addition, cost can also scale with partner densities, so models in the second table add density-dependent cost as linear dependencies. Models highlighted in grey either violate model assumption (see Method section) or generate unbounded growth, or both. The rest of the 48 models satisfy the 4 model assumptions and have bounded growth. For criteria that have a long left-hand side are written as $f(\beta, \beta', \epsilon) > \delta$. The specific forms of $f(\beta, \beta', \epsilon) > \delta$ can be found in the supplemented MATLAB code. The MATLAB code also includes all 81 models in this table and the process of testing assumptions, calculating and verifying the above criteria.

Adding complexities	Inequality	Parameters
Base model criterion	21. $\frac{(\beta + 1)^2}{4\beta\varepsilon} > \delta$	β, ε
Include competition	82. $\frac{(\beta + a + 1)^2}{4\beta\varepsilon(a + 1)} > \delta$	β, ε, a
Include initial density	83. $\frac{1}{\varepsilon}(1 - x_0)x_0\beta + \frac{1}{\varepsilon}(1 - x_0) > \delta$	β, ε, x_0
Include asymmetry	84. $\frac{\rho(\beta_1 + 1/\theta_1 + 1)^2}{4\beta_1\varepsilon_1(\rho/\theta_1 + \theta_2)} > \delta_1$ $\theta_1 = \frac{\beta_2}{\beta_1}, \theta_2 = \frac{\varepsilon_2\delta_2}{\varepsilon_1\delta_1}$	$\beta, \varepsilon, \rho, \theta_1, \theta_2$
Include turnover rate	85. $\frac{(\beta(1 - \varepsilon\delta_0) + 2)^2}{4\beta\varepsilon} \left(\frac{\rho}{\rho + 1}\right) > \delta$	$\beta, \varepsilon, \rho, \delta_0$
N-mutualist system	86. $\frac{(\beta + \frac{n}{n-1})^2}{4\beta\varepsilon\frac{n}{n-1}} > \delta$	β, ε, n

Table S2: Increasing model complexity increases criterion complexity. The base model is identical to model 21 in Table S1. Number 82-86 and their corresponding criteria are obtained by relaxing assumptions in model 21. Including the 48 models in Table S1, we in total obtained coexistence criteria of 52 models. Note that number 21 and number 83 used the same model structure. The detailed models and criterion derivations are shown in section III of SI. Also see Supplementary Materials for the MATLAB code of above models and the process of calculating and verifying the corresponding criteria.

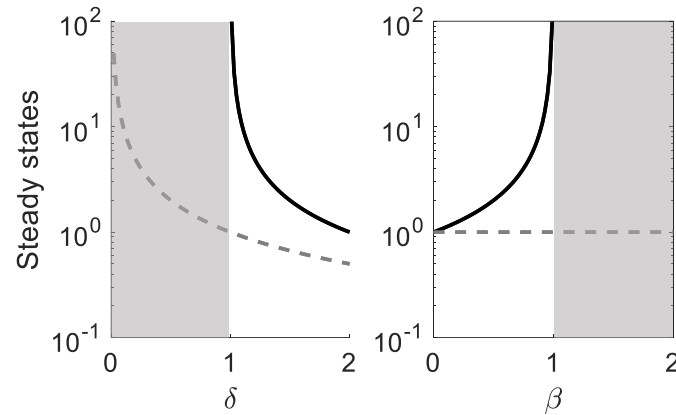


Figure S1: Typical bifurcation diagrams of models that create unbounded growth. Previously-developed general coexistence criteria describe the boundary between stable coexistence (white regions) and unbounded growth (grey regions). Solid lines represent stable steady state for coexistence. Grey dashed lines represent steady states of a mutualist in the absence of its partner. This diagram is generated using model:

$$\begin{aligned}\frac{dX_1}{d\tau} &= X_1(1 - \delta X_1) + \beta X_2 X_1 \\ \frac{dX_2}{d\tau} &= X_2(1 - \delta X_2) + \beta X_1 X_2\end{aligned}$$

The left panel is generated with $\beta = 1$ and the right panel is generated with $\delta = 1$.

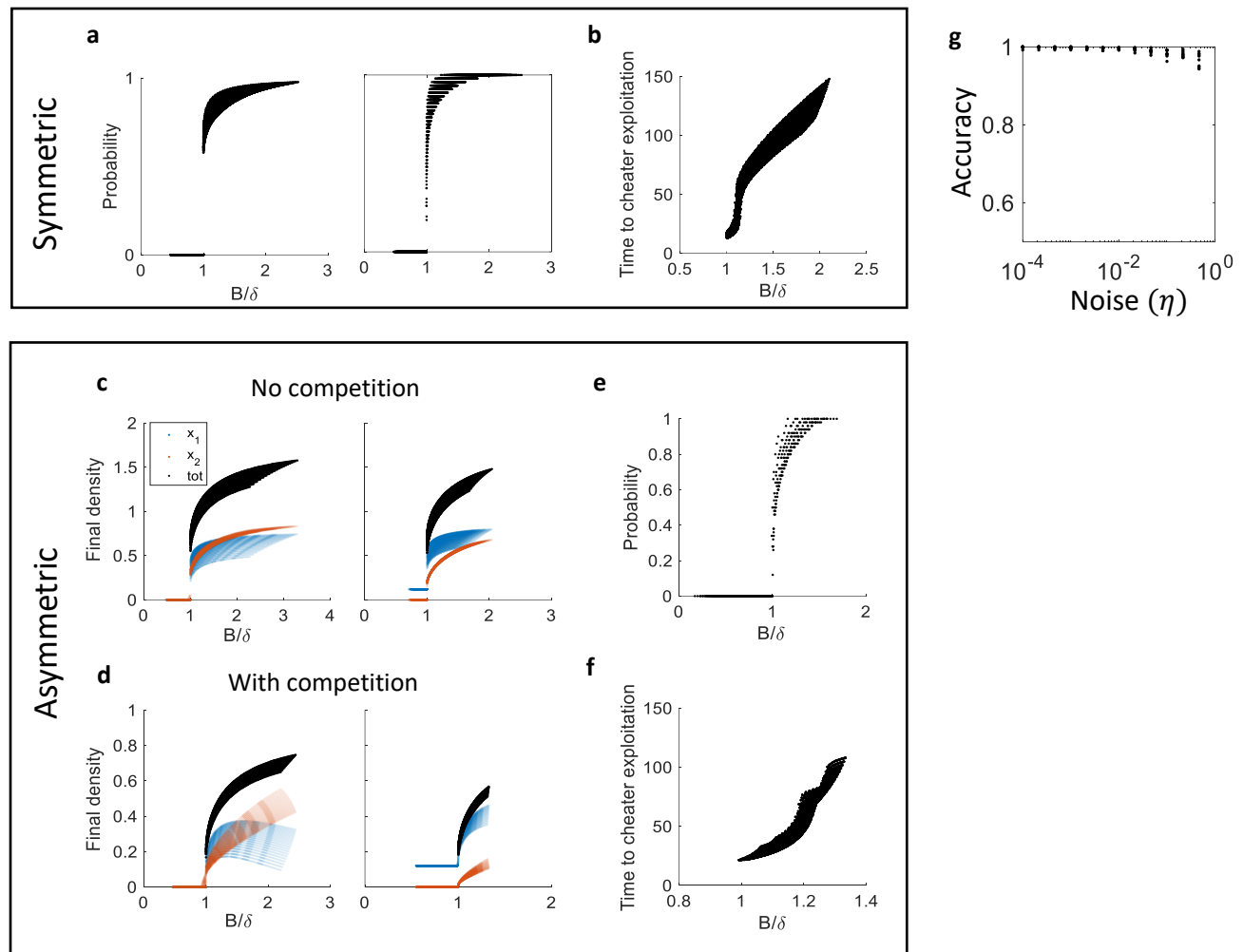


Figure S2: The predictor B/δ is a general metric for predicting performance of symmetric (panel a, b) and asymmetric mutualistic systems (panel c-f). All the traces are comprised of individual dots. Each dot represents result from one set of model parameter. **a.** Mutualistic systems have initial-density-dependent coexistence. In the left panel, we assume the same initial density of both populations that is a uniform random variable. In the right panel, we varied the ratio of the two initial densities and kept the total initial density constant. In both cases, B/δ is predictive of coexistence probability. **b.** When cheaters can arise in a system, B/δ is predictive of the time duration the system can persist before cheater exploitation. **c.** An asymmetric mutualistic system can be either obligatory (left panel), where both populations extinct when $B/\delta < 1$ or facultative (right panel), where one population can persist even when $B/\delta < 1$. The black dots represent total density and the two colors represent the total density of the two partners. The same representations are used in panel d. **d.** The predictive power of B/δ for total density also holds for asymmetric systems that share carrying capacity. **e.** B/δ is predictive of probability of coexistence for asymmetric systems. **f.** The predictive power of B/δ for resistance to cheater exploitation also holds for asymmetric systems. Please refer to SI for models and parameter values used in generating this figure. **g.** The accuracy of criterion is robustly maintained after addition of extrinsic Gaussian noise. X axis indicates the standard deviation of the Gaussian noise. Note that the noise is significant considering the maximum total density of the system is 1. Please see Supplementary Text section IV.B for the simulation details.

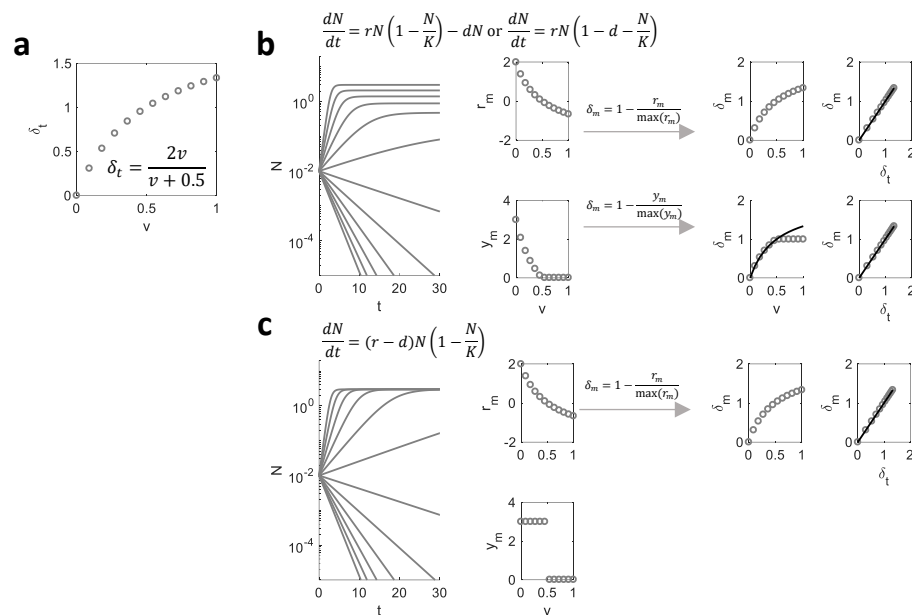


Figure S3: A general method for quantifying δ . Regardless of different model formulations, δ can always be measured using either growth rate or final population size of an individual population. **a.** We simulated the growth of a population that is modulated by a death term d . Stress δ is simply d normalized by growth rate r . In the simulations, the true δ (δ_t) depends on an independent variable v . **b.** One type of model structure has δ modulating both growth rate (r) and yield (y , final density). N is the population size and K is the carrying capacity. Based on the growth curve or final population size, measurements of growth rate (r_m) and yield (y_m) can be obtained when varying v . The δ measurement (δ_m) can be calculated by $1 - \frac{r_m}{\max(r_m)}$ or $1 - \frac{y_m}{\max(y_m)}$. Note that since y_m is always non-negative, $1 - \frac{y_m}{\max(y_m)}$ is always positive, but δ can be greater than 1 when the population size decreases. To address this discrepancy, extrapolation of δ_m is needed to capture death rate (the black trace in the δ_m versus v plot). **c.** In another type of model formulation, δ only modulates growth rate. Simulation shows that y_m either equals to carrying capacity or 0. Thus, in this case, only r_m can be used to calculate δ . The right-most panels show the comparison between the measured δ (δ_m) and true δ (δ_t). In all cases, δ_m is a good approximation of δ_t (the solid black line represents $\delta_m = \delta_t$).

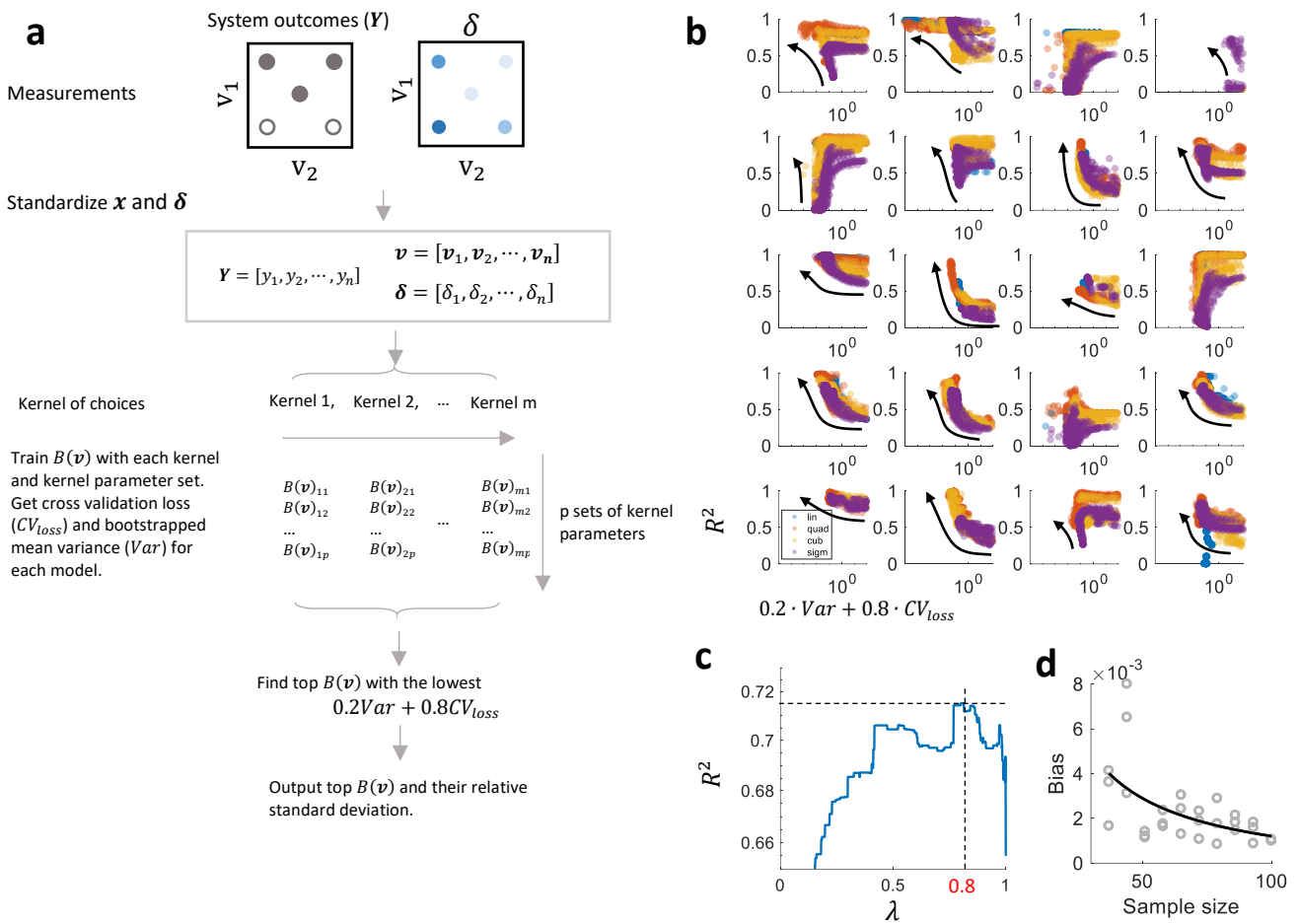


Figure S4: The calibration procedure using SVM. **a.** Measurements are first arranged in matrix forms and each standardized to have mean of 0 and standard deviation of 1. The standardized data are used as inputs to train SVM models with different kernels and kernel parameters. 10-fold cross validation accuracy loss (CV_{loss}) and variance (Var) are calculated for each trained $B(v)$. Top $B(v)$ are selected based on minimizing the metric $0.2 \cdot Var + 0.8 \cdot CV_{loss}$. We then output the top $B(v)$ (usually 5) and their corresponding relative standard deviation ($RSD = variance/mean$) obtained by bootstrapping to evaluate the confidence of the calibrated $B(v)$. **b.** To show that $0.2 \cdot Var + 0.8 \cdot CV_{loss}$ is a consistent indicator of high correlation (indicated by R^2 values) between $B(v)$ and the true B , we constructed 20 models that have known $B(v)$ (each subplot represents one model). Each dot represents the result of one set of kernel parameter and the colors represent the type of kernel used. These models show that minimizing $0.2 \cdot Var + 0.8 \cdot CV_{loss}$ leads to higher R^2 values (indicated by the black arrows). See SI for details of the specific model and model parameters used to generate data in this panel. **c.** 0.8 in the metric is picked by sweeping λ in $(1 - \lambda) \cdot Var + \lambda \cdot CV_{loss}$. Mean R^2 calculated using the top 3 $B(v)$ of each 20 model varies with λ values. The max mean R^2 occurs at $\lambda = 0.8$. **d.** Bias decreases exponentially with increasing sample size. This analysis is done using the model corresponding to the third subplot in the first column of panel b. A subset of the 100 data points is randomly sampled without replacement to train $B(v)$ and the bias of $B(v)$ is calculated. The black trace represents the data points fitted to the function $Bias = a \cdot (Sample\ size)^{-\alpha} + b$.

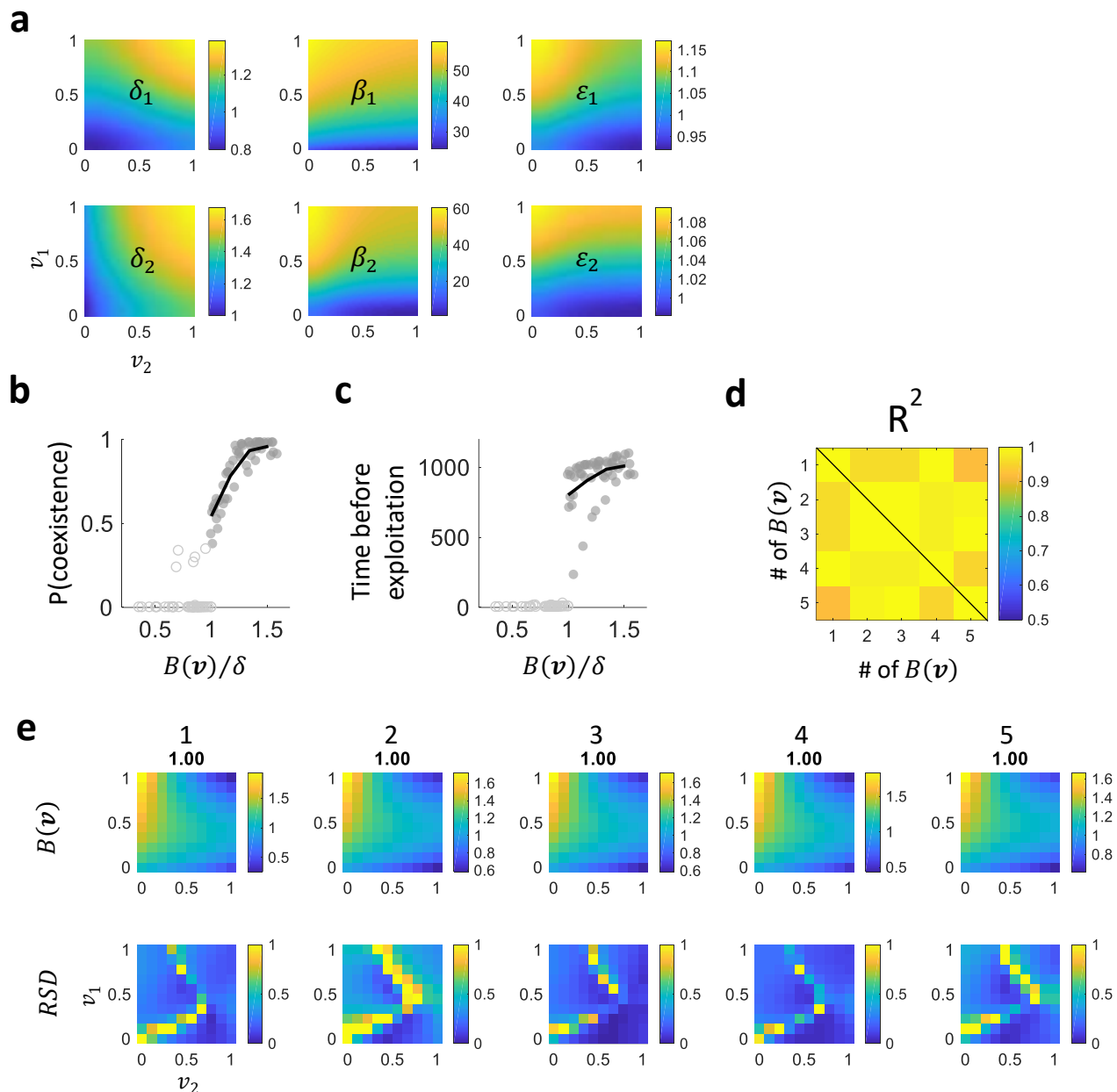


Figure S5: Application of the simple rule to a complex model with no explicit form of B . **a.** δ_1 , δ_2 , β_1 , β_2 , ϵ_1 , and ϵ_2 vary with v_1 and v_2 . **b.** The calibrated $B(v)$ along with δ is predictive of probability of coexistence. To get the probability of coexistence, at each (v_1, v_2) , we ran 100 simulations with 100 different ratios of initial densities while keeping the total initial density the same. **c.** $B(v)/\delta$ is also predictive of how well the system can resist cheater exploitation. The y axis indicates the time the system can persist before the exploitation of cheater populations. **d.** The high pairwise R^2 values of the top 5 $B(v)$ indicate high confidence of the calibration results. **e.** The top 5 $B(v)$ from the calibration procedure. The title for each $B(v)$ indicates its index corresponding to panel d and its corresponding training accuracy. Each $B(v)$ can also be evaluated based on their bootstrapped relative standard deviation (RSD).

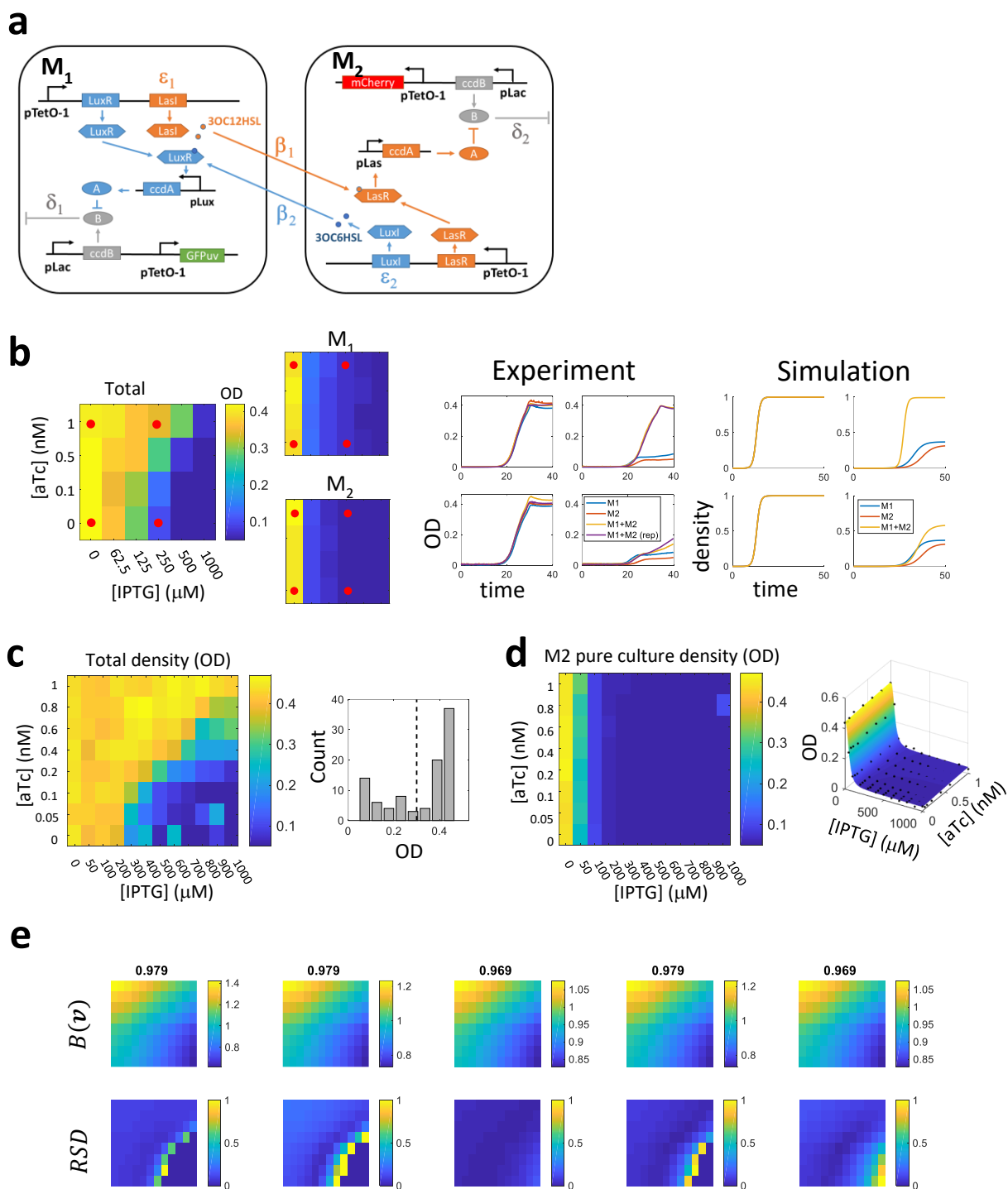


Figure S6: The QS-based mutualism system. **a.** The schematic of the synthetic gene circuit. δ_1 , δ_2 , β_1 , β_2 , ϵ_1 , and ϵ_2 are indicated in the diagram to show the molecular mechanisms each term is primarily associated with. **b.** Verification of the basic system dynamics. The final OD values of coculture and monocultures are recorded. Monocultures of M_1 and M_2 are significantly suppressed by IPTG, while the cocultures exhibit synergistic growth. The total density of the cocultures increases with increasing [aTc]. The comparison between experimental and simulated temporal dynamics indicate that our model captures

the basic mutualistic dynamics observed empirically. The four red dots in the heatmaps indicate the four experimental conditions of the simulated and experimental time courses. We used equations III.22 and 23 in the Supplementary Text for the simulation. The parameter values are: $\delta_1([IPTG] = 0) = \delta_2([IPTG] = 0) = 0$; $\delta_1([IPTG] = 1000\mu M) = 0.63$; $\delta_2([IPTG] = 1000\mu M) = 0.68$; $\beta_1([aTc] = 0\mu M) = \beta_2([aTc] = 0\mu M) = 2$; $\beta_1([aTc] = 1nM) = \beta_2([aTc] = 1nM) = 200$; $\varepsilon_1 = \varepsilon_2 = 1$; $\rho = 1$ **c.** Using higher resolution of [aTc] and [IPTG] gradients, we obtained the total density of cocultures using OD measurements. At 32 hours, a bimodal distribution of final densities emerges with a trough at OD=0.3. **d.** Quantification of δ based on OD of M_2 monoculture at 32 hours. The final OD is fitted using $OD([aTc], [IPTG]) = V_a \frac{[aTc]^{n_a}}{[aTc]^{n_a} + k^{n_a}} + V_I \frac{k^{n_I}}{[IPTG]^{n_I} + k^{n_I}} + b$ to reduce noise of the resulting δ measurements. δ is then calculated by $\delta = 1 - \frac{OD([aTc], [IPTG])}{\max(OD([aTc], [IPTG]))}$. **e.** The top 5 $B(v)$ and their corresponding relative standard deviation (*RSD*) (the axes are the same as heatmaps in panel c and d).

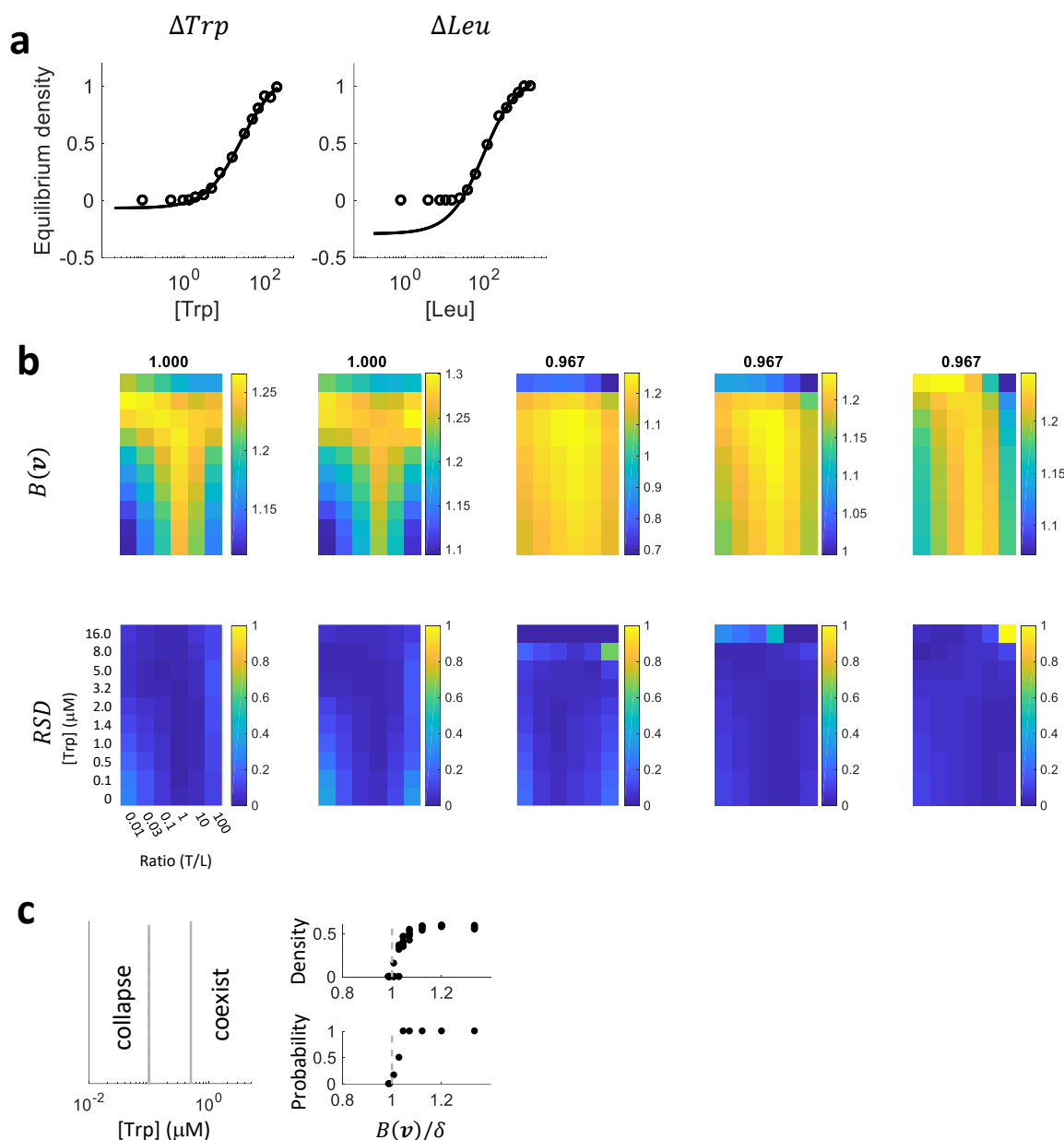


Figure S7: The yeast auxotroph system. **a.** The normalized equilibrium density of ΔTrp and ΔLeu were used to quantify δ for each strain. We only fitted density that is above the detection limit (10,000 cells/well). Since ΔLeu reaches below detection limit when $[Leu]$ is relatively low, extrapolation is done to estimate the death rate at lower $[Leu]$ (see Fig. S3a for the reasoning). As a result, δ of ΔLeu has a higher dynamic range than that of ΔTrp . **b.** 5 top $B(v)$ show both consistency and discrepancy. All $B(v)$ indicate an optimal initial density at an intermediate level. However, when $[Trp]$ increase to 16 μM , $B(v)$ can either increase or decrease. **c.** Our approach can also predict probability of coexistence. We excluded the ratio of initial density, leaving $[Trp]$ to construct a one dimensional v . The boundary between coexistence and collapse is between 0.1 μM tryptophan, 0.8 μM leucine and 0.5 μM tryptophan, 4.0 μM leucine. To calibrate a $B(v)$ for this system, we assume supplementing the two amino acids does not change the cooperation capability. This calibration demonstrates that our procedure can also predict probability of coexistence for experimental systems.

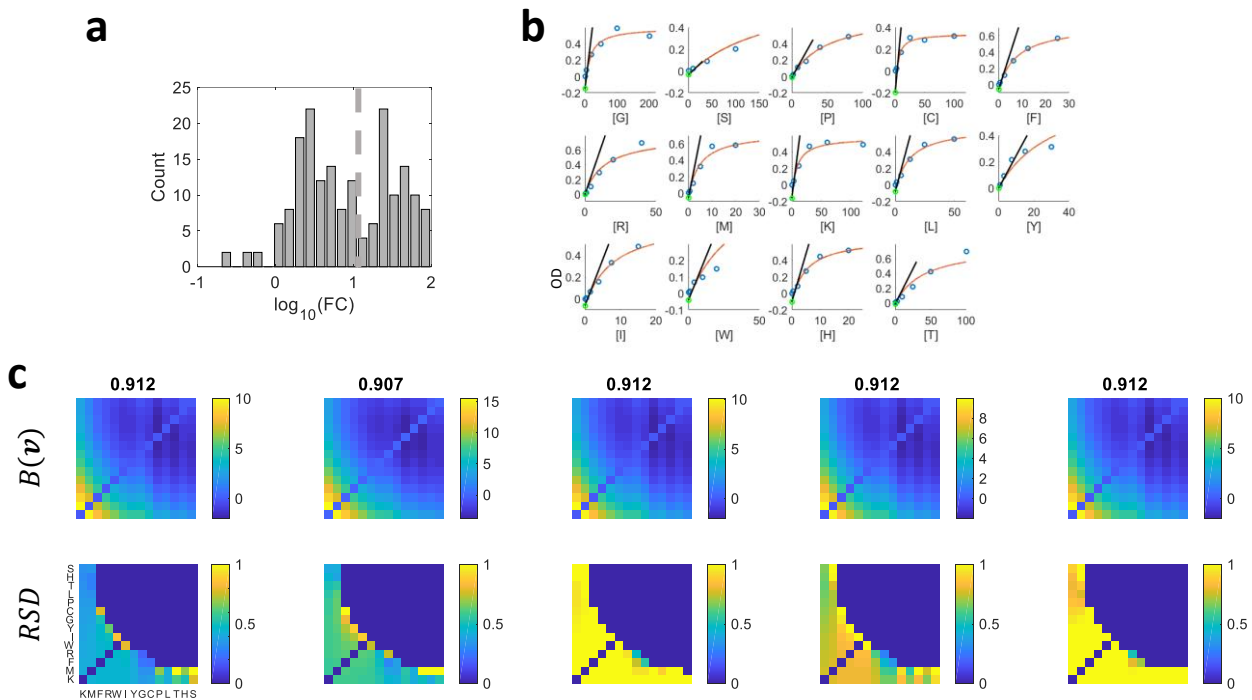


Figure S8: The 92 pairwise *E. Coli* mutualistic systems constructed using 14 auxotrophic strains. **a.** The logarithm of fold change (FC) of total final density for the 92 systems exhibit a bimodal distribution. The trough of the bimodal distribution corresponds to FC=10 (dashed grey line), which is used to classify coexistence and collapse. **b.** Quantification of δ for all 14 auxotrophs by fitting yield of the strain at different concentrations of their corresponding amino acid. Since none of the auxotrophs can grow without supplementary amino acid, $\delta \geq 1$ is expect for all auxotrophs at [AA]=0 (AA denotes the corresponding amino acid). Therefore, according to the reasoning demonstrated in Fig. S3, extrapolation of the data is required. We only fitted the data points that have OD>0.1, since OD<0.1 can be below the linear detection range of microplate readers. The data are fitted and extrapolated with Hill equations (red curves) or using a linear extrapolation of the fitted curves (the black trace). Green circles represent the values of OD of the fitted Hill equations at [AA]=0. δ is then calculated by $1 - \frac{OD([AA]=0)}{\max(OD([AA]))}$ for all 14 strains using OD([AA] = 0) obtained from either Hill equation or linear function. The two methods yield comparable results. We chose the δ calculated using the Hill equation for the calibration process. **c.** The top calibrated $B(v)$ and their variability are indicated by relative standard deviation (RSD).

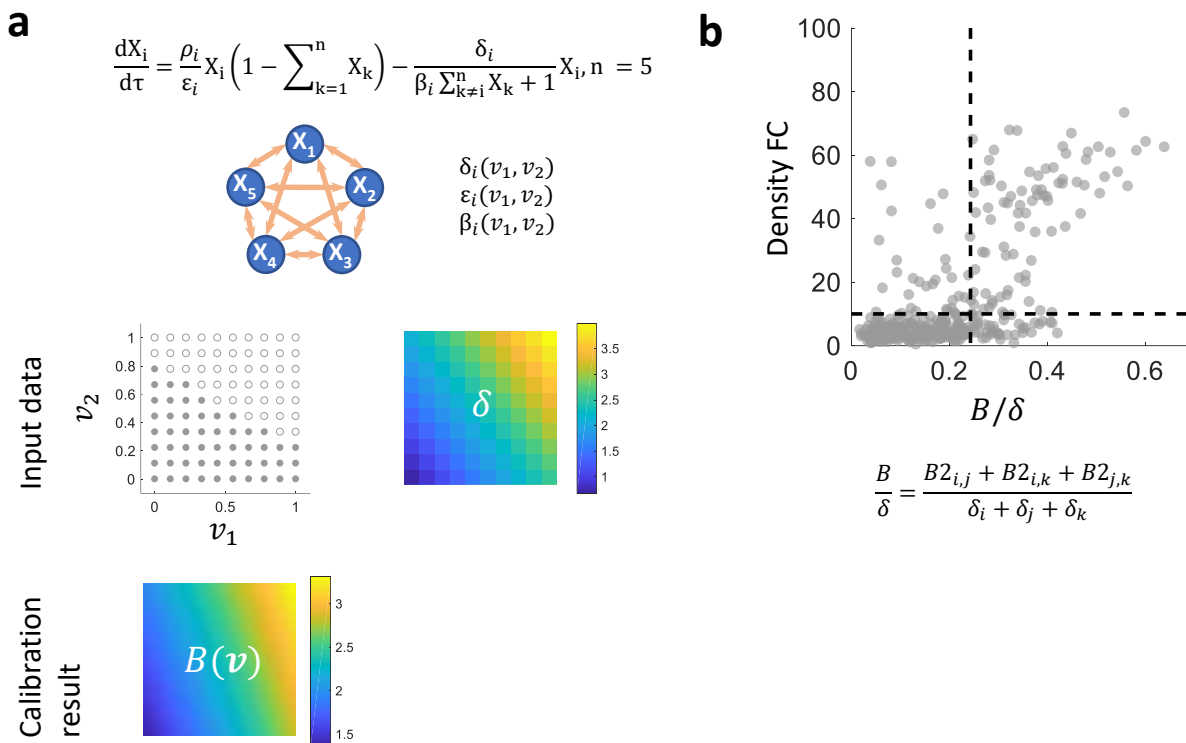


Figure S9: Mutualistic systems comprised of more than two partners. a. A mutualistic model with 5 partners. All the parameters in the model are linear functions of v_1 and v_2 and the parameters of the linear function are randomly generated. **b.** Constructing a predictive metric for 3-member systems from calibrated results of 2-member systems. We first normalized the range of calibrated B of 2-member systems ($B2$) to $[0, 1]$. The effective benefit of 3-member system is simply the average of $B2$ for all underlying 2-member systems and stress is the average of δ for all 3 members. The subscripts of $B2$ and δ represent the indices of the 3 members. We swept a threshold for B/δ between 0 and 1 to classify coexistence ($FC \geq 10$) versus collapse ($FC < 10$). The prediction accuracy reaches the maximum of 80.8% at a threshold equals to 0.24.

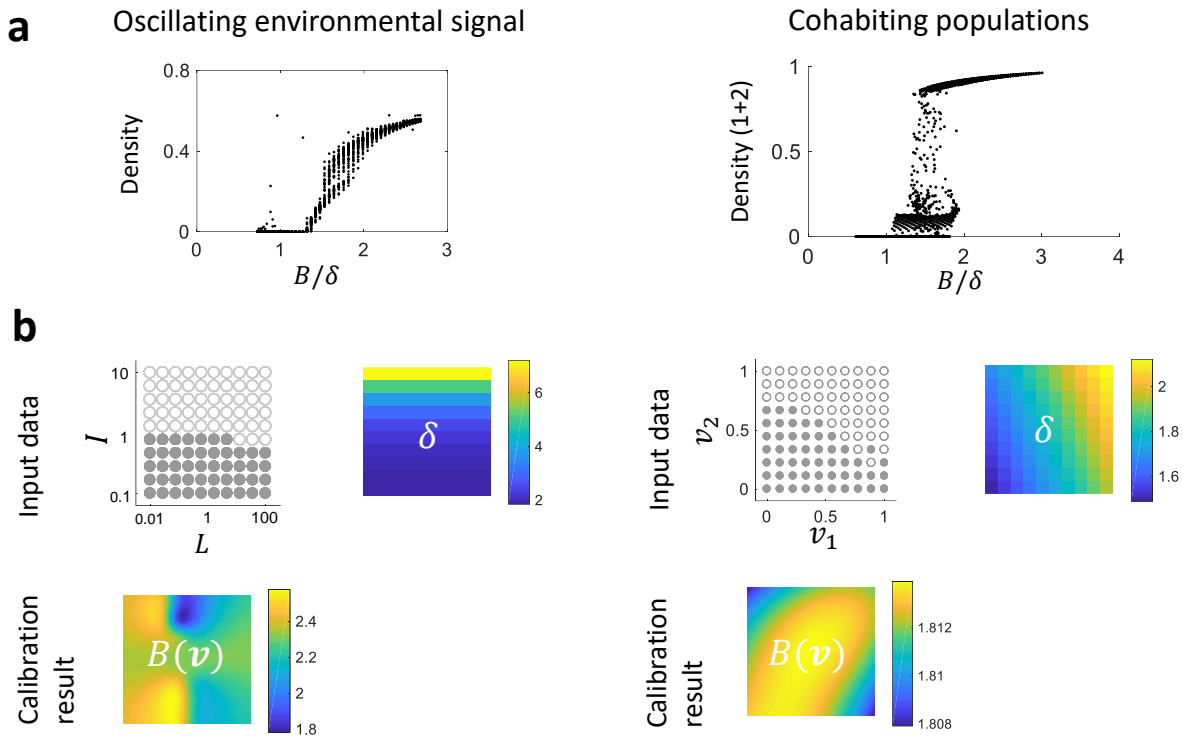


Figure S10: Mutualistic systems in dynamic environments. **a.** Two systems were analyzed: one system is a pairwise mutualism system that is modulated by oscillatory signals (the left panel) and the other system contains a pair of mutualists and 5 bystander populations that interact with the mutualists (the right panel). We verified that theoretically calculated B/δ roughly holds as a predictor for total density. The transition between coexistence and collapse, however, do not occur strictly at 1. Each dot represents a simulation result from one set of model parameters. **b.** The data used for the calibrations and the calibrated $B(v)$. The two $B(v)$'s are used for the prediction plots shown in Fig. 4c, d.

Movie S1. This video demonstrates the relative position of the boundary surface $F(\delta, \mathbf{v}) = 0$ and 5 observations in the schematic shown in Fig. 2b. Input data in Fig. 2b can be represented in a 3D space. The color of the dots and the z axis positions both indicate δ values. Coexistence is indicated by closed circles and collapse is indicated by open circles. The gray boundary surface separates coexistence and collapse. The surface is above observations that represent coexistence ($B > \delta$) and below observations that represent collapse ($B < \delta$). This surface can be directly interpreted as $B(\mathbf{v})$ since $F(\delta, \mathbf{v}) = 0 \Rightarrow F(B, \mathbf{v}) = 0 \Rightarrow B(\mathbf{v})$.

Movie S2 – S4. These 3 videos show the calibrated surface of $B(\mathbf{v})$ relative to input data in 3D space. The blue surfaces are the boundary $F(\delta, \mathbf{v}) = 0$ that separate coexistence and collapse. The surface is also equivalent to $B(\mathbf{v})$. The indices on x and y axes in video 4 correspond to the same strain orders in Fig. 3j.

MATLAB code and movie S1 – S4 can be found using the following Dropbox link:

<https://www.dropbox.com/sh/c65vqo0vw4sxb6k/AAC4QpRXs9KyiCZV1RjMaPfqa?dl=0>