Structure-based identification of sensor species for anticipating critical transitions

Andrea Aparicio, Jorge Velasco-Hernández, Claude H. Moog, Yang-Yu Liu, Marco Tulio Angulo

Structure-based identification of sensor species for anticipating critical transitions

Andrea Aparicio[a,b], Jorge X. Velasco-Hernández[a], Claude H. Moog[c], Yang-Yu Liu[b], and Marco Tulio Angulo[d,1]

[a]Institute of Mathematics, Universidad Nacional Autónoma de México, 76230 Juriquilla, México; [b]Channing Division of Network Medicine, Department of Medicine, Brigham and Women's Hospital and Harvard Medical School, Boston, MA 02115; [c]Laboratoire des Sciences du Numérique de Nantes, UMR 6004, CNRS, Nantes 44321, France; and [d]Consejo Nacional de Ciencia y Tecnología–Institute of Mathematics, Universidad Nacional Autónoma de México, 76230 Juriquilla, México

Edited by Marten Scheffer, Wageningen University and Research, Wageningen, The Netherlands; received March 10, 2021; approved October 15, 2021

Ecological systems can undergo sudden, catastrophic changes known as critical transitions. Anticipating these critical transitions remains challenging in systems with many species because the associated early warning signals can be weakly present or even absent in some species, depending on the system dynamics. Therefore, our limited knowledge of ecological dynamics may suggest that it is hard to identify those species in the system that display early warning signals. Here, we show that, in mutualistic ecological systems, it is possible to identify species that early anticipate critical transitions by knowing only the system structure—that is, the network topology of plant–animal interactions. Specifically, we leverage the mathematical theory of structural observability of dynamical systems to identify a minimum set of "sensor species," whose measurement guarantees that we can infer changes in the abundance of all other species. Importantly, such a minimum set of sensor species can be identified by using the system structure only. We analyzed the performance of such minimum sets of sensor species for detecting early warnings using a large dataset of empirical plant–pollinator and seed-dispersal networks. We found that species that are more likely to be sensors tend to anticipate earlier critical transitions than other species. Our results underscore how knowing the structure of multispecies systems can improve our ability to anticipate critical transitions.

critical transitions | observability | pollination | seed dispersal

The well-being of our species depends on maintaining the integrity of Earth's ecological systems (1, 2). Still, these systems can suffer abrupt, undesired, and possibly irreversible changes known as "critical transitions" (3). Critical transitions occur when the system state exhibits a fold bifurcation. Therefore, near the bifurcation (i.e., tipping) point, a gradual change in external conditions results in abrupt system responses. For example, in plant–animal mutualistic systems, declining visitation rates (4) or phenological mismatches (5) can decrease the mutualistic benefits until the system undergoes a critical transition, causing abrupt species extinctions (6–8). Fortunately, theoretical and experimental works show that it is possible to anticipate critical transitions because they can generate early warning signals before they occur (9–14). However, for systems with many species, where it is unfeasible to monitor all of them, anticipating critical transitions is more challenging because the early warning signals can be absent (15) or weakly present (16) in some species. Indeed, which and how early species display early warnings depend on the dynamics of the whole system (17–19). This fact suggests that our limited knowledge of ecological systems dynamics renders it hard, if not impossible, to identify species that early anticipate a critical transition.

Here, we focus on mutualistic ecological systems and characterize conditions for identifying species that early anticipate critical transitions by knowing only the system structure. The structure of mutualistic systems is captured by a bipartite network with nodes representing species (plants and animals) and bidirectional edges connecting a plant species with an animal species if they have a mutualistic interaction (20, 21). Mapping the network structure of a mutualistic system is more straightforward than inferring its dynamics (22). Using the system structure, we leveraged the structural observability theory (23–26) to identify a minimum set of species whose measurement allows deducing changes that occur in the abundance of all other species in the system. We call this set a "minimum set of sensors" (MSS). By construction, monitoring an MSS allows anticipating critical transitions, as long as at least one species in the system displays early warning signals.

In general, a system will have several different MSSs. Some species may belong to all MSSs, indicating that their monitoring is essential. Other species may not belong to any MSSs, indicating that we do not need to monitor them directly. By analyzing 51 empirical plant–pollinator and seed-dispersal networks from the Web of Life database (22), we show that species detecting earlier the critical transition tend to appear more often in the MSS. We analyze the implications of this result on our ability to identify species that early anticipate critical transitions, assuming that all we know about the system is its structure. Specifically, we study the minimum number of species that we need to measure to obtain the earliest warning of the transition and how early we can anticipate the transition if we measure a single species. Overall, our analysis shows that most species in these empirical systems exhibit late or no warning of a critical transition, emphasizing the necessity of a methodology to identify species that early anticipate critical transitions in multispecies systems.

Significance

Ecological systems can undergo abrupt and often catastrophic changes known as critical transitions. These critical transitions can be preceded by early warnings that are useful to anticipate them. But anticipating critical transitions in systems with many species is challenging because the early warnings can be absent in some species. Here, we leverage the mathematical formalism of "structural observability of dynamical systems" to address this fundamental issue. In mutualistic ecological systems, we show that "sensor species" that early anticipate a critical transition can be identified by knowing only the structure of plant–animal interactions.

Author contributions: M.T.A. conceived the project; J.X.V.-H., C.H.M., Y.-Y.L., and M.T.A. designed research; A.A. and M.T.A. performed research; A.A. and M.T.A. analyzed data; A.A. did the numerical analysis; M.T.A. did the theoretical analysis; M.T.A. wrote the manuscript; and Y.-Y.L. edited and revised the manuscript.

The authors declare no competing interest.

This article is a PNAS Direct Submission.

[1]To whom correspondence may be addressed. Email: mangulo@im.unam.mx.

This article contains supporting information online at https://www.pnas.org/lookup/suppl/doi:10.1073/pnas.2104732118/-/DCSupplemental.

Published December 15, 2021.

Structure-Based Identification of Sensor Species in Mutualistic Systems

To identify species that early anticipate a critical transition, we leveraged the notion of "observability " (*Materials and Methods* and *SI Appendix*, section 1A). Let $\Sigma = \{1, 2, \cdots, N\}$ denote the set of species in the system. The system is observable from a subset $S \subseteq \Sigma$ of species if it is possible to deduce the abundance of all species by measuring the abundance of species in S only. If the system is observable from S, we say that S is a set of "sensors." Therefore, by construction, monitoring a set of sensors allows for anticipating a critical transition, as long as an early warning is present in at least one species of the system.

Crucially, the structural observability theory shows that we can generically identify sensors by knowing only the system structure and the class of dynamics that the system can adopt (*Materials and Methods* and *SI Appendix*, section 1B). That is, knowledge of the parameters' values of the system dynamics is not necessary. Without loss of generality, we assume that the system structure is a connected bipartite network. We also assume that the early warning signals occur as small deviations of species' abundances from their stable equilibrium values (27–29), implying that the system adopts approximately linear dynamics (30). Under these assumptions, we prove that $S \subseteq \Sigma$ is (generically) a set of sensors if and only if there is a node-disjoint path cover of the network, where each path of length $\neq 1$ ends in a sensor (*Materials and Methods* and *SI Appendix*, Theorem 2 of section 2). Note that if S is a set of sensors, then any other larger set containing S is also a set of sensors. We thus ask for a set of sensors with a minimum number of species, which we call an MSS. Our characterization of sensors allows for identifying MSSs (*Materials and Methods* and *SI Appendix*, section 2E). A system can have several different MSSs. Species appearing in most MSSs are essential to measure according to the system structure, while those appearing only in a few MSSs are not.

Fig. 1*A* illustrates the identification of sensors in a toy system of $N = 3$ species. Our characterization shows that this system has two MSSs: $S = \{1\}$, corresponding to the path-cover $\{2 \to 3 \to 1\}$, and $S' = \{2\}$, corresponding to the path-cover $\{1 \to 3 \to 2\}$ (Fig. 1*B*). Species 3 alone cannot be an MSS because the only path-cover $\{1 \to 3, 2 \to 3\}$ ending in that species is not node-disjoint. To understand why species 3 alone cannot be a sensor, we checked if this species can detect changes in the abundance of the other two species. We thus perturbed species 1 and 2, and then we checked if species 3 can detect that such perturbation occurred. We simulated the effect of these perturbations on the species abundance using a stochastic differential equation model for plant–animal mutualism (31) compatible with the network structure (*SI Appendix*, section 4A). Fig. 1*C* shows a simultaneous perturbation to species 1 and 2 that remains undetected by species 3, confirming that species 3 alone is not a sensor. This situation occurs because the network topology is such that the perturbations "cancel out" when they meet in species 3. The fact that species 1 (or species 2) is an MSS guarantees that, generically, there is no change in species 2 and species 3 (or species 1 and 3) that remains undetected (Fig. 1*D*). *SI Appendix*, Figs. S1 and S2 illustrate our structure-based identification of sensors in bigger mutualistic systems.

Results

Anticipating Critical Transitions using Sensor Species. We evaluated the sensor's ability to detect early warning signals of critical transitions, starting with our toy example of Fig. 1*A*. To generate a critical transition, we consider the stochastic mutualistic dynamics and slowly decrease the "mutualistic strength" $\mu \geq 0$ of the plant–pollinator interactions until a "critical value" $\mu^* \approx 0.75$ where no species survives (*SI Appendix*, sections 3C and 4A). Consistent with the theory of early warning signals (32), the species abundance shows a higher variance as the system approaches the critical transition (Fig. 1*E*). But, although all species show early warning signals, the variance increase is stronger and occurs earlier in the species appearing in the MSS (Fig. 1*F*). That is, species that anticipate the transition earlier

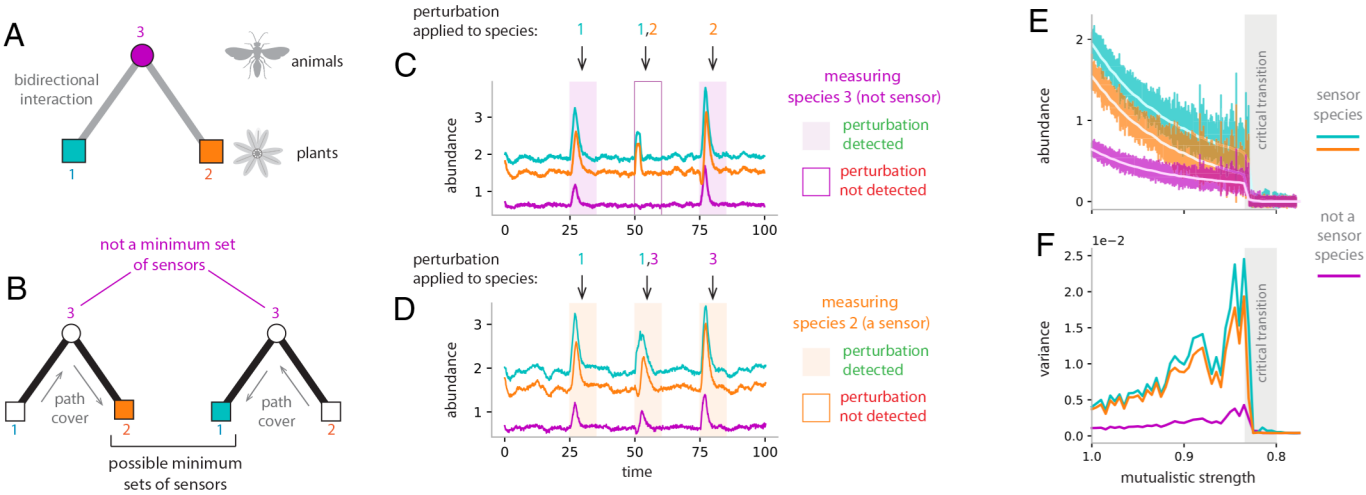

Fig. 1. Detecting critical transitions in mutualistic systems by identifying a minimum set of their sensor species. (*A*) The bipartite network characterizing the structure of a hypothetical mutualistic system of two plants {1, 2} and one pollinator {3} (i.e., $N = 3$ species). Edges represent bidirectional interactions. (*B*) We show that $S \subseteq \Sigma = \{1, 2, 3\}$ is a set of sensors if and only if there exists a node-disjoint path cover of the network, where each path of length $\neq 1$ ends in S. Shown are the two possible node-disjoint path covers $\{1 \to 3 \to 2\}$ and $\{2 \to 3 \to 1\}$, implying that either plant {1} or {2} can be a solo sensor. Species {3} cannot be a sensor by itself because the only path cover that ends in that species $\{1 \to 3, 2 \to 3\}$ is not node-disjoint. (*C* and *D*) We tested the ability of each species to detect perturbations in the system under stochastic mutualistic dynamics (see *SI Appendix*, section 4A for equations). The system is first allowed to reach its equilibrium. Then, a perturbation is applied by changing the abundance of the indicated species. In *C*, the simultaneous perturbation on species {1} and {2} remains undetected by species {3}, confirming that it cannot be a sensor. Since species {1} or {2} are sensor species, there is no change in the system that remains undetected by them. *D* illustrates this fact. (*E*) Using the mutualistic dynamics, we induced a critical transition in this toy system by slowly decreasing the mutualistic interaction strength between plants and the pollinator, producing the collapse of all species. (*F*) As expected from the theory of early warning signals of critical transitions, the abundance of species shows higher variance as the system approaches the critical transition. However, the variance increase is sharper and occurs earlier in the sensor species, making their early warning signals stronger than in the nonsensor species.

Table 1. Summary of variables

Name	Symbol	Definition
Sensor score	$s_i \in [0, 1]$	Probability that species i is in an MSS
Early warning score	$w_i \in [\mu^*, \infty)$	Maximum w such that $\mathrm{Var}_t[x_i(\cdot; \mu)] > \mathrm{Var}_t[x_i(\cdot; \mu')]$ for all $\mu < \mu' \leq w$
Earliest detection	$w_{\max}$	$w_{\max} = \max_{i\in\Sigma} w_i$, earliest detection of the critical transition
Early warning score ratio	$r(q) \in [0, 1]$	$r(q) = (\max_{i\in S_q} w_i)/w_{\max}$, how early a set $S_q \subseteq \Sigma$ of qN species, $q \in [0, 1]$, anticipates the critical transition compared to $w_{\max}$
Effective network structure	$\sigma \in \mathbb{R}$	$\sigma = \sigma(s_{\max}, \mathrm{Skew}(\{s_i\}))$, a function of $s_{\max} = \max_i s_i$ and the skewness $\mathrm{Skew}(\{s_i\})$ of the sensor scores
Deviations from equilibrium	$\rho \geq 0$	$\rho = \langle\, [\max_t x_i(t; \mu) - \min_t x_i(t; \mu)] \,/ \langle x_i(t; \mu)\rangle_t \,\rangle_{i,\mu}$, quantifies species' deviation from their equilibrium abundances

Here, N is the number of species, $\Sigma = \{1, \cdots, N\}$ is the set of species, μ^* is the mutualistic strength at which the critical transition occurs, $x_i(t; \mu)$ denotes the abundance of species i at time t in a system with mutualistic strength μ, and $\langle\cdot\rangle_p$ and $\mathrm{Var}_p[\cdot]$ denote average and variance with respect to parameters p, respectively.

are also more essential to measure according to the system structure.

To investigate the generality of the above findings, we studied critical transitions happening in three larger empirical seed-dispersal systems observed in Mount Missim ($N = 40$), the Genebra reserve ($N = 25$), and Calton ($N = 25$), obtained from the Web of Life database (22). We make our analysis quantitative by calculating, for each species $i \in \Sigma$, its "sensor score" s_i and its "early warning score" w_i, as defined in Table 1 (Fig. 2 *A–C*). On the one hand, species with a high sensor score appear more often in the MSS, implying that their measurement is essential (*Materials and Methods* and *SI Appendix*, section 2F). Note that the system structure is "poorly informative" when all sensor scores are small (i.e., no species is essential to measure). On the other hand, the early warning score w_i is the largest mutualistic strength at which species i detects the critical transition. Species with a higher w_i anticipate the critical transition earlier, and species with $w_i = \mu^*$ do not anticipate the critical transition before it occurs. We calculated the early warning score by simulating a critical transition using a stochastic differential equation model of obligate mutualism compatible with the system structure (*Materials and Methods* and *SI Appendix*, section 3). In this model, we randomly choose the species' mortality and competition parameters, producing different realizations of the system parameters (6). Across these realizations and realizations of the stochastic dynamics, the three systems undergo a remarkably similar critical transition, where all species collapse when the mutualistic strength decreases to $\mu^* \approx 0$ (Fig. 2 *D–F*). A substantial percentage of species do not anticipate such transition (median of 22.5%, 12%, and 12% across realizations for Missim, Genebra, and Calton, respectively). Note that the sensor score only depends on the system structure (i.e., it can be calculated without knowing the system dynamics and before the critical transition occurs). By contrast, the early warning score depends on the system dynamics.

We analyzed the performance of sensors for anticipating critical transitions by comparing the early warning and sensor scores in the above three systems (Fig. 2 *G–I*). Starting with Missim, species with a high early warning score tend to have a high sensor score, implying that knowing this system structure allows for identifying species that early anticipate the critical transition (Fig. 2*G*). Consistent with previous analysis (6), species with a high sensor score tend to be specialists (i.e., nodes with a low degree in the network). Therefore, we reasoned that measuring species one by one from high to low sensor score should soon allow measuring a species with the earliest detection $w_{\max}$ of the critical transition. To explore this hypothesis, for a fraction $q \in [0, 1]$ of the system species, we define the "early warning score ratio" $r(q)$ as in Table 1. Higher values of $r(q)$ imply that measuring qN species allows for anticipating earlier the critical transition, with $r(q) = 1$ if at least one measured species has the earliest detection. We focus our analysis on the minimum fraction q^* of species that we need to measure for obtaining the earliest detection and the ratio $r_1 = r(1/N)$ quantifying how early we can anticipate the transition by measuring a single species. A small q^* indicates that we need to measure fewer species to obtain the earliest detection. A higher r_1 indicates that measuring a single species detects the critical transition closer to $w_{\max}$.

Fig. 3*A* shows the early warning score ratio for Missim when choosing species according to their sensor score (from high to low), degree (from low to high), or randomly (*Materials and Methods*). In this system, choosing species by sensor score or degree is similar because species with a high sensor score are specialists (red in Fig. 2*G*). Compared to random, choosing species by sensor score or degree decreased by a third the fraction of species that we need to measure for obtaining the earliest detection (median $q^* \approx 0.125$, $q^* \approx 0.15$, and $q^* \approx 0.425$ when choosing species by sensor score, degree, and randomly, respectively). With probability 0.75, we obtained faster the earliest detection when choosing species by sensor score compared to random (Fig. 3 *B*, *Left*). Furthermore, choosing a single species by sensor score or degree yielded an early warning score ratio two times higher compared to choosing randomly (Fig. 3 *B*, *Right*). These results confirm that knowing the structure of this system improves our ability to identify species that early anticipate critical transitions.

We obtained similar results when analyzing the Genebra reserve (Fig. 2*B*). Species with a high early warning score again tended to have a large sensor score (Fig. 2*H*). Compared to random, choosing species by sensor score or degree reduced roughly by half the fraction of species that we need to measure for obtaining the earliest detection (Fig. 3*C*; median $q^* \approx 0.24$ and $q^* = 0.44$ when choosing species by sensor score or degree and randomly, respectively). With a probability of 0.66, we obtained the earliest detection faster by choosing species using sensor score or degree than when choosing randomly (Fig. 3 *D*, *Left*). Choosing a single species by sensor score yields an early warning score ratio two times higher compared to choosing randomly and 11% higher when compared to choosing by degree (Fig. 3 *D*, *Right*; median $r_1 = 0.56$, $r_1 = 0.45$, and $r_1 = 0.27$ when choosing by sensor score, degree, and randomly, respectively). Note that sensor score outperforms degree because, while species with a high sensor score tend to be specialists, not all specialists have a high sensor score. Indeed, from the five groups of specialists in this system, group 1 has higher sensor and early warning scores (Fig. 2*H*). This fact is also apparent in Fig. 2*B* because the nodes in group 1 are greener and bigger compared to the other specialist groups.

We finally turn to Calton, where neither specialists nor the species with the highest sensor score tend to be the first to anticipate the transition (Fig. 2 *C*, *F*, and *I*). Indeed, the first species

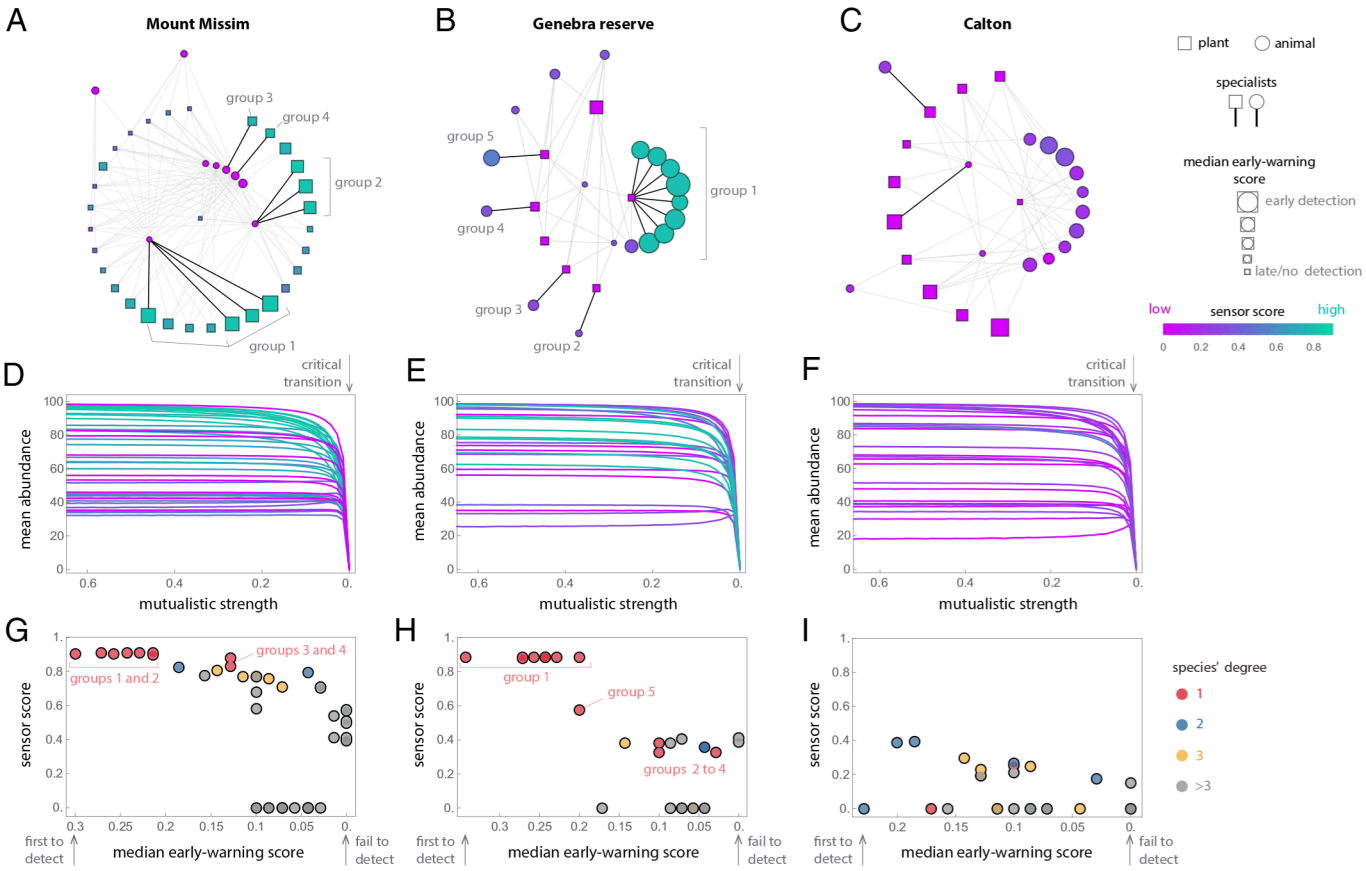

Fig. 2. Sensor species anticipate critical transitions in multispecies mutualistic systems. (*A–C*) Three empirical seed-dispersal networks, Mount Missim (Papua New Guinea), Genebra reserve (Brazil), and Calton (Great Britain), obtained from the Web of Life database. We calculated the sensor score of each species (i.e., the probability that such a species belongs to an MSS). Species with a higher sensor score are those whose measurement is essential (green color). Then, we simulated a critical transition in each system to calculate the species' early warning score (node size represents the median early warning score calculated from 20 realizations of the system parameters using the nominal values of *SI Appendix*, Table S3). Specialists are marked with black edges. Specialists belong to the same group if they interact with the same species. Species with a higher early warning score (green) anticipate earlier the critical transition (bigger nodes), showing how knowing the system structure allows identifying species that early anticipate critical transitions. (*D–F*) As the mutualistic interaction strength μ decreases, the three systems undergo a critical transition at $\mu^* \approx 0$, where all species go extinct. Lines represent different species colored by their sensor score. Panels represent one realization obtained by using the nominal parameter values. The critical transition is similar in the three systems, but the informativeness of their structure for identifying species that early anticipate critical transitions is different (colors). (*G–I*) Sensor score vs. median early warning score calculated for 20 realizations of the system parameters. Missim and Genebra contain species with a high sensor score, and those species have a high early warning score. In Calton, no species has a high sensor score, indicating that the network structure is not informative about which species is essential to measure. In this system, neither specialists nor the species with the highest sensor score are the first to anticipate the transition. The species with the second-highest median early warning score has the highest sensor score.

to anticipate the critical transition changes drastically across realizations of the system parameters (*SI Appendix*, Fig. S4). This is expected because the sensor-score distribution is closer to zero and has a lower maximum value than in Missim and Genebra, indicating that there are no species that are essential to measure according to the system structure. Consequently, the early warning score ratio is similar when choosing species by sensor score, degree, or randomly (Fig. 3 *E* and *F*).

The above results illustrate that having a large maximum sensor score indicates that the system structure is informative about which species are best to measure for anticipating critical transitions. And when the structure is informative, species with a high sensor score tend to anticipate earlier the critical transition. However, when is the system structure informative enough? We next address this question using a large dataset of empirical networks, helping us probe the performance of sensors across diverse systems.

Validation of Sensors in a Large Dataset of Empirical Networks. We analyzed a dataset of 51 empirical systems containing all plant–pollinator and seed-dispersal networks from the Web of Life database with $N \leq 80$ species (*SI Appendix*, section 3A). The MSS size is heterogeneous across these systems, requiring us to measure between 3% and 73% of their species depending on the system structure (median of 44%; *SI Appendix*, Fig. S6*A*). The MSS size increases in systems with a higher disparity between the number of plants and animals (Spearman's $\rho = 0.9668$, $P < 2.9415 \times 10^{-19}$; *SI Appendix*, Fig. S6*B*). This result can be explained as follows. Since empirical networks tend to be nested, increasing the disparity increases the number of specialists (33). Therefore, the MSS size increases because systems with more specialists require measuring more species. We also confirmed that species with a high sensor score tend to be specialists, but not all specialists have a high sensor score (*SI Appendix*, Fig. S7).

We simulated a critical transition in these systems as before to calculate the early warning score of each species (*SI Appendix*, section 3C). We found that most species in these systems do not show early warnings of the critical transition, and very few species provide early detection (Fig. 4*A*). Specifically, across networks and their realizations, the median proportion of species that provide the earliest detection $w_{\max}$ is only 3.8%.

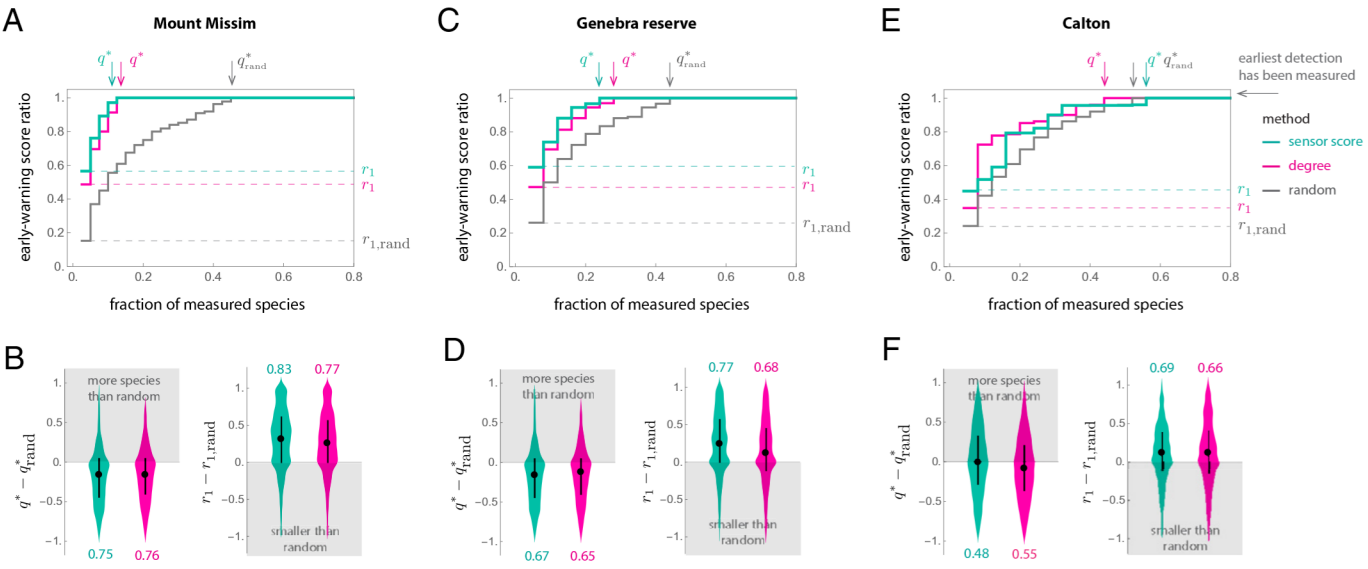

Fig. 3. Measuring species with a high sensor score improves our ability to anticipate critical transitions. Results are for the three empirical systems of Fig. 2. (*A*) Early warning score ratio $r(q)$ as a function of the fraction $q \in [0, 1]$ of species, quantifying how early we can anticipate the critical transition by measuring qN species compared to the earliest detection $w_{\max} = \max_{i \in \Sigma} w_i$. Recall that $r(q) = 1$ if the measured species contain the earliest detection. The panel shows the results for three methods of choosing species: by sensor score (from high to low; green), by degree (from low to high; pink), and at random (gray). Thick lines denote the median across realizations of the system parameters and choice of species with equal sensor score or degree (*Materials and Methods*). See *SI Appendix*, Fig. S5 for upper and lower quantiles. Choosing species by sensor score or degree significantly outperforms choosing species at random, indicating that the network structure contains useful information to anticipate critical transitions. *B*, *Left* shows the distribution for the fraction q^* of species needed to obtain the earliest detection of the critical transition (i.e., $r(q^*) = 1$), compared to random (lower values mean better performance). *B*, *Right* shows the distribution for the early warning score ratio $r_1 = r(1/N)$ of choosing a single species, compared to random (higher values mean better performance). Distributions were built by using 1,000 repetitions of choosing species over the 20 realizations obtained for different system parameters. Dots mark the median of the distribution and lines the lower and upper quantiles. Numbers below (respectively, above) each distribution are the probability that q^* is lower than random (respectively, r_1 higher than random). (*C* and *E*) Same as in *A*, but for Genebra and Calton, respectively. (*D* and *F*) Same as in *B*, but for Genebra and Calton, respectively.

A median of 21.42% of the species has a score $w_i \geq 0.5 w_{\max}$, implying that only one-fifth of the species provide "good" early warnings in the sense that they anticipate the transition before the mutualistic strength drops below half of the earliest warning. Moreover, a substantial fraction of species cannot detect the critical transition before it occurs (a median of 14.28% of the species have $w_i = \mu^*$, but for some systems and realizations, this percentage is as high as 87.5%). These findings underscore again the need for a methodology to identify species that early anticipate critical transitions in multispecies systems, indicating that finding such species is challenging because they represent a tiny fraction of all species.

Next, we compared the early warning and sensor scores of species across these empirical networks (1,855 species in total; Fig. 4*B*). We found again that species with a high early warning score tend to have a large sensor score, also confirming that those species represent a tiny fraction of all species (Fig. 4 *B*, *Left*). By contrast, species showing late or no detection of the critical transition are common, and they tend to have a low sensor score (Fig. 4 *B*, *Right*).

To quantify the implications of the above result, we calculated q^* and r_1 for each system, comparing their values when choosing species by sensor score, degree, and randomly (Fig. 4*C*). Compared to random, choosing species by sensor score typically decreases by half the fraction of species that we need to measure for obtaining the earliest detection, slightly improving the performance when choosing by degree (median $q^* = 0.2244$, $q^* = 0.2307$, and $q^* = 0.46$ when choosing species by sensor score, degree, and randomly, respectively). In 92.11% of the systems, choosing species by sensor score outperforms choosing species at random (white area in Fig. 4*C*), and in 60.78%, choosing species by degree. One species chosen by sensor score typically anticipates the critical transition twice as early compared to choosing it randomly and 20% earlier compared to choosing it by degree (median $r_1 = 0.48$, $r_1 = 0.39$, and $r_1 = 0.22$ when choosing species by sensor score, degree, and randomly, respectively). Choosing one species by sensor score outperforms choosing it randomly or by degree in 88.23% and 74.5% of the systems, respectively. These results show that the sensor score improves our ability to identify species that early anticipate critical transitions when all we know is the system structure, especially when we can only measure one or a few species.

Characterizing the Performance of Sensors. Here, we characterize the systems where sensors perform better than random in the sense that $q^* - q^*_{\text{rand}}$ is positive and large and $r_1 - r_{1,\text{rand}}$ is negative and large (white area in Fig. 4*C*). Our analysis shows that the performance of sensors can be explained by using only two covariates: the "effective network structure" σ and the magnitude $\rho \geq 0$ of species' "deviations from equilibrium" (see definitions in Table 1 and *SI Appendix*, section 4E for details). The effective network structure is a function of the maximum $s_{\max}$ and skewness of the sensor-score distribution (Fig. 4*D*). As its name indicates, σ summarizes how informative the system structure is for identifying species that early anticipate critical transitions. In contrast to σ, which only depends on the system structure, ρ also depends on the system dynamics. That is, as two systems with identical structures approach a critical transition, the early warning signals can produce deviations from the species' equilibrium with different magnitude depending on its dynamics. In this sense, ρ summarizes the effect of the system dynamics on sensors' ability to anticipate critical transitions. Note that ρ can be calculated from time-series data of species abundances without knowing the dynamics that generated it.

Using the above two covariates, two nonlinear regression models can predict the performance of sensors, indicating that sensors

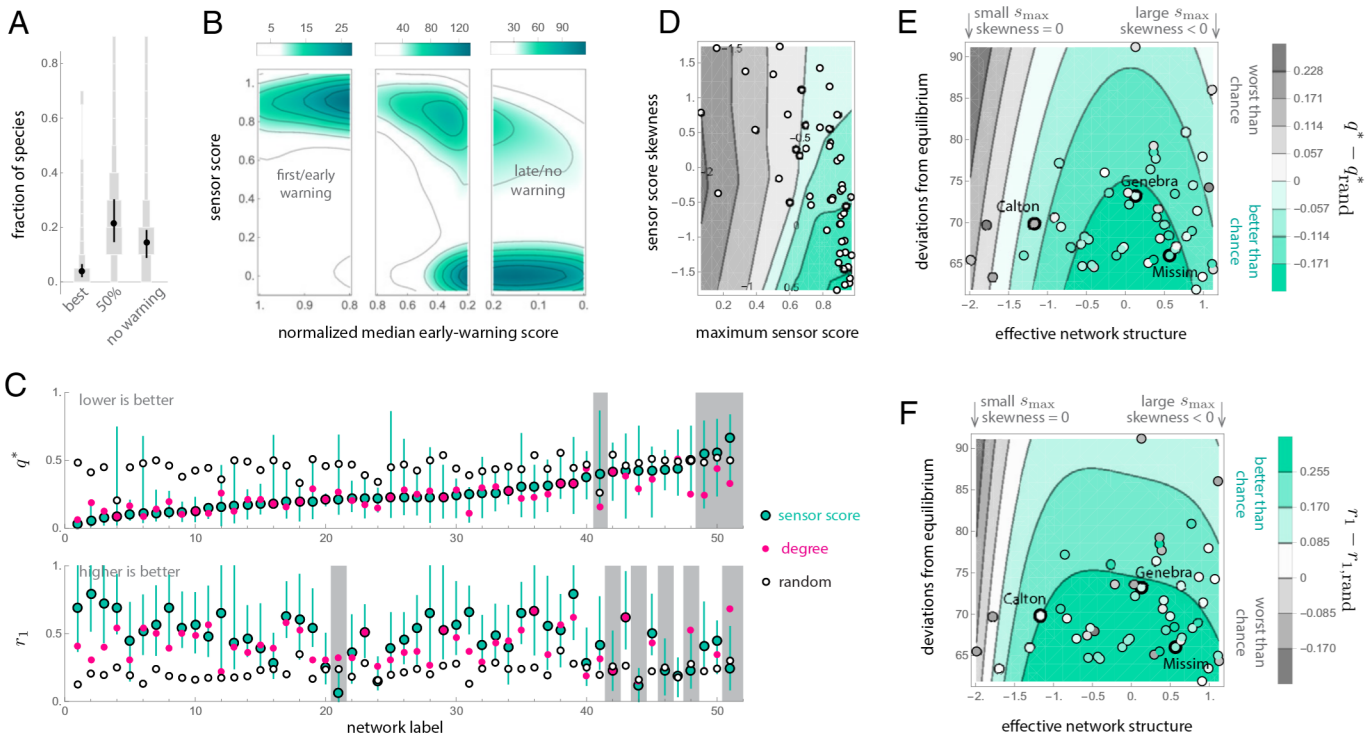

Fig. 4. Performance of sensors for anticipating critical transitions in a large dataset of empirical networks. Panels show the results of inducing a critical transition in our dataset of 51 empirical networks (26 pollinator and 25 seed-dispersal networks; *SI Appendix*, section 3A) with nominal parameters (*SI Appendix*, Table S3). (*A*) Distributions for the fraction of species with the best (i.e., $w_i = w_{max}$), up to 50% of the best (i.e., $w_i > 0.5w_{max}$), and no early warning (i.e., $w_i = \mu^*$). Black dots and lines indicate the median and upper/lower quantiles, respectively, across empirical networks and their realizations. Histograms show the distribution. Species early anticipating the critical transition represent a tiny fraction of all species. (*B*) Density distribution of normalized early warning score w_i/w_{max} vs. sensor score s_i for all the species in our dataset. Species are separated into three groups (panels) according to their normalized early warning score: the left group has 143 species, the center group has 1,104 species, and the right group has 608 species. Species with high early warning scores tend to have a large sensor score. (*C*) Fraction q^* of measured species needed to obtain the transition' earliest detection (*Upper*) and early warning score ratio r_1 when measuring a single species (*Lower*). Points show the median of these two quantities across the 51 empirical networks. Green lines show the upper and lower quantiles for choosing species by sensor score. Gray denotes systems in which choosing species by sensor score is worse than choosing at random. Networks are ordered by q^*, obtained when choosing species by sensor score. (*D*) We applied a nonlinear dimensional reduction to calculate the effective network structure σ, a one-dimensional variable capturing how informative is the system structure for identifying species that early anticipate critical transitions. This variable depends on the maximum s_{max} and the skewness of the sensor-score distribution. Shown are the level curves of σ with colors corresponding to values indicated by the numbers on top of each level curve. (*E* and *F*) Nonlinear regression models predicting the median sensor performance $q^* - q^*_{rand}$ (adjusted $R^2 = 0.76$, MSE $= 0.029$) and $r_1 - r_{1,rand}$ (adjusted $R^2 = 0.67$, MSE $= 0.046$), respectively, using the effective network structure σ and the magnitude ρ of species abundance deviations from their equilibrium. In both panels, background colors correspond to model predictions. Sensors perform better than chance in regions colored with stronger green. Both models indicate that sensors perform better as $\sigma \to 0$ and $\rho \to 0$. Circles correspond to the 51 empirical systems, colored according to the true sensor performance.

perform better as $\sigma \to 0$ and $\rho \to 0$ (Fig. 4 *E* and *F*). This result makes sense. On the one hand, smaller deviations ρ better satisfy our assumption of linear dynamics. On the other hand, a large positive σ occurs when s_{max} is high and the skewness is strongly negative, corresponding to a sensor-score distribution concentrated at high values with a large maximum. In this case, sensors perform as random because the structural observability theory predicts that it is essential to measure all species. A large negative σ occurs when s_{max} is small and the skewness is near zero, corresponding to a sensor-score distribution that is symmetric and has a small maximum. In this case, sensors perform as random because the structural observability theory predicts that no species is essential to measure.

Robustness Against Changes in the System Dynamics. We end by analyzing the performance of sensors to changes in the system parameters, including changes in the maximum interspecific competition strength within animals or plants, the mutualistic tradeoff, the baseline mutualistic strength, the handling time in the functional response of the mutualistic benefit, and the noise intensity (*SI Appendix*, sections 4F and 4G). Increasing the maximum competition strength decreases the deviation's magnitude of species from their equilibrium, improving the performance of sensors (*SI Appendix*, Fig. S8*A*). Sensors perform better with small values of the mutualistic tradeoff (*SI Appendix*, Fig. S8*B*). The handling time strongly affects the early warning score and the deviation's magnitude (*SI Appendix*, Fig. S8*D*). In particular, with a small handling time, sensors fail because no species in the system anticipates the transition.

A small baseline mutualistic strength makes it easier to anticipate critical transitions (*SI Appendix*, section 4G and Figs. S10–S12). More precisely, species have a high and very similar early warning score for small values of this parameter. Therefore, in this case, anticipating the critical transition is relatively easy because most species detect it very early. Consistent with previous results (7, 34), in this case, some species may go extinct before the community's complete collapse. Such species tend to have a high sensor score. On the other hand, increasing the baseline mutualistic strength decreases the species' early warning score until some fail to anticipate the critical transition. Here, reducing the mutualistic interaction strength produces a single critical transition where most species go extinct. In this case, species with a high sensor score tend to have a large early warning score, anticipating earlier the critical transition.

The performance of sensors is similar when using autocorrelation or variance to build the early warning signal (*SI Appendix*, section 4H). In addition, changes in the noise intensity have little effect on the sensors' performance (*SI Appendix*, Fig. S8*E*). Sensors also perform similarly when noise is multiplicative instead of additive, but, as found in previous works (6), the early warning signals must be constructed by using autocorrelation or the coefficient of variation (*SI Appendix*, section 4I).

Discussion

Our analysis supports previous studies indicating that critical transitions are challenging to anticipate in multispecies systems (15, 16): About 14% of the species in a mutualistic system may not show any warning, and an additional 65% of species may show "late" warnings of the critical transition (i.e., less than half of the earliest warning). We expect these percentages to be higher in more detailed ecological dynamics models and natural ecological systems.

Responding to the above challenge, we leveraged the structural observability theory to show that the structure of plant–animal interactions in mutualistic systems can provide information to identify sensor species that early anticipate critical transitions. In particular, our work justifies why specialists tend to early anticipate critical transitions in mutualistic systems (6). Namely, specialists tend to be essential from the structural observability viewpoint, meaning that they condense information to deduce changes happening in all other species of the system. Importantly, not all specialists are equally essential to measure. That is, from the set of all specialists, those who anticipate earlier a critical transition tend to have a high sensor score, as we illustrated in Genebra. Sensor species are different from "indicator species," which are sensitive to environmental parameters (35).

Some system structures are more informative for identifying sensors. This result is expected, as some system structures, such as a complete bipartite network, render all species equivalent. Our analysis shows that the information that a system structure has for identifying species that early anticipate critical transitions can be quantified by the effective network structure σ. A system structure is more informative as $\sigma \to 0$ ($\sigma \to -\infty$ for a complete bipartite network). Our analysis also shows that sensors perform better as the deviations ρ of species abundance from their equilibrium decreases. In particular, sensors perform better when the early warning dynamics are closer to linear. Small deviations are more likely to occur in mutualistic systems with a large handling time and stronger competition within plants and animals. In general, given a system structure and its corresponding σ value, the two regression models we built could be used to predict the range of deviations from equilibrium ρ, where we expect that sensors perform better than chance.

The predictions of the structural theory of dynamical systems have already been experimentally validated in other problems (36, 37). Similarly, the "structural stability" theory has shown how the structure of ecological systems can inform about their dynamical behavior (31, 38–40). Note that there is a so-called "duality" between structural observability and structural controllability for systems with linear dynamics (23, 25). A system is controllable from the species $A \subseteq \Sigma$ if we can steer the complete system to any desired state by directly influencing species in A only (41). Since the structure of mutualistic systems has bidirectional edges, the duality implies that a mutualistic system is structurally observable if and only if it is structurally controllable. Therefore, when the system structure is informative, we conjecture that directly influencing species with a high sensor score could help control critical transitions by delaying or producing them.

Our study has two significant limitations: It only considers mutualistic systems and assumes that their early warning dynamics are approximately linear. The generality of the structural theory of dynamical systems suggests that these two limitations could be overcome. For example, the general linear structural observability theory (23) could be applied to characterize sensors when species interactions are not necessarily mutualistic. If the early warnings dynamics is nonlinear, the structural theory could be used to analyze the observability of a general class of nonlinear dynamics (25, 42). For instance, in food chains with arbitrary nonlinear dynamics, the structural observability theory predicts that it is necessary to monitor all species in the lowest trophic level (25). This prediction agrees with experimental works (10). These extensions would be necessary for validating our approach with experimental data, underscoring how knowing the structure of multispecies systems allows better monitoring, especially when their dynamics remain uncertain and difficult to infer.

Materials and Methods

Observability and Sensor Species. Consider mutualistic systems with n plant and m animal species. Let $x(t) = (x_1(t), \cdots, x_n(t)) \in \mathbb{R}^n_{\geq 0}$ and $y(t) = (y_1(t), \cdots, y_m(t)) \in \mathbb{R}^m_{\geq 0}$ denote the abundance of plant and animal species at time $t \geq 0$, respectively. Here, x_i and y_j denote the abundance of the i-th plant and j-th animal species, respectively. Let $\Sigma = \{1, 2, \cdots, N\}$ denote the set of all species ($N = m + n$, where the first n species are plants). Consider that the system dynamics can be approximated by a Stochastic Differential Equation (SDE)

$$\begin{aligned} dx(t) &= f(x(t), y(t))dt + \omega\, dW(t), \\ dy(t) &= g(x(t), y(t))dt + \omega\, dW(t), \end{aligned} \quad [1]$$

with initial species abundances $(x(0), y(0)) \in \mathbb{R}^n_{\geq 0} \times \mathbb{R}^m_{\geq 0}$. Above, $f : \mathbb{R}^n \times \mathbb{R}^m \to \mathbb{R}^n$ and $g : \mathbb{R}^n \times \mathbb{R}^m \to \mathbb{R}^m$ are functions determining the system population dynamics. $W(t)$ is a vector of independent Wiener processes with adequate dimension, and $\omega \geq 0$ characterizes the noise intensity.

We assume that the early warning signals of a critical transition occur as small deviations of species' abundances from their (deterministic) stable equilibrium values $(x^*, y^*) \in \mathbb{R}^n_{\geq 0} \times \mathbb{R}^m_{\geq 0}$ (9, 27–29). This assumption implies that the dynamics of such deviations are approximately linear, conforming with the Hartman–Grobman Theorem (30, 43). Therefore, the early warning signal's dynamics is well-approximated by using Eq. **1** with

$$\begin{aligned} f(x, y) &= A\,(x - x^*) + B\,(y - y^*), \\ g(x, y) &= C\,(x - x^*) + D\,(y - y^*). \end{aligned} \quad [2]$$

The matrices $A \in \mathbb{R}^{n\times n}$, $B \in \mathbb{R}^{n\times m}$, $C \in \mathbb{R}^{m\times n}$, and $D \in \mathbb{R}^{m\times m}$ are calculated as $A = \partial f/\partial x$, $B = \partial f/\partial y$, $C = \partial g/\partial x$, and $D = \partial g/\partial y$, with partial derivatives evaluated at (x^*, y^*).

To identify species that anticipate critical transitions, we use the notion of observability, a cornerstone concept in control theory (see, e.g., refs. 44 and 45 and *SI Appendix*, section 1A). Consider a species subset $S \subseteq \Sigma$ and let

$$z = Fx + Hy, \quad [3]$$

denote the measured output containing the abundance of species in S, $z \in \mathbb{R}^P$, with $P = |S|$ being the number of species that S contains. Here, $F \in \mathbb{R}^{P\times n}$ and $H \in \mathbb{R}^{P\times m}$ characterize the measured species. Specifically, $f_{ij} \neq 0$ if the j-th plant species is in S, and $h_{jk} \neq 0$ if the k-th animal species is in S. The system of Eqs. **1**–**3** is observable from S if the abundance of all species $(x(t), y(t))$ at any time $t > 0$ can be deduced from knowing only the abundance of the measured species $z(t)$ up to time t (*SI Appendix*, Definition 1 in section 1A). If the system is observable from S, we say that S is a set of sensors. Given the matrices (A, B, C, D) and (F, H), Kalman's rank observability criterion gives necessary and sufficient conditions that a set of sensors needs to satisfy (*SI Appendix*, section 1A). If S is a set of sensors, and all matrices are known, we can optimally estimate the abundance of all species from the output using the Kalman–Bucy filter (ref. 46, p. 136). Therefore, by construction, a set of sensors can anticipate a critical transition happening in any part of the system if at least one species shows early warnings of this transition.

Structure-Based Identification of Sensors in Mutualistic Systems. Identifying sensor species is challenging because of our limited knowledge of the functions f and g in Eq. **1**. This limitation renders the matrices (A, B, C, D) of Eq. **2** uncertain, so Kalman's rank criterion cannot be used to probe observability.

To circumvent the above challenge, we take a structuralist approach, showing that it is possible to generically identify sensors by knowing only the

structure of plant–animal mutualistic interactions (*SI Appendix*, section 2). This structure is captured by a bipartite network $\mathcal{B}(\Sigma, \mathsf{E})$, where nodes Σ represent species, and edges $\mathsf{E} \subseteq \Sigma \times \Sigma$ represent plant–animal (bidirectional) interactions (31, 38, 39). We assume that this network is connected. Note that this bipartite network encodes the zero pattern of the matrices (B, C) in Eq. **2**. That is, the structure encodes the presence ($b_{ij} \neq 0$, $c_{ji} \neq 0$) or absence ($b_{ij} = 0$, $c_{ji} = 0$) of plant–animal interactions, but it does not provide information about their magnitude (i.e., the values of b_{ij} and c_{ji} are unknown). Importantly, knowing the structure of a mutualistic system gives no information about the matrices A and D (i.e., about the interactions within animals or plants). Above, the word "generically" means that sensors render the system observable for all values of the nonzero parameters, excepting a zero Lebesgue-measure set. See *SI Appendix*, Definition 6 in section 2E for the formal statement.

Given the system mutualistic structure captured by the bipartite network $\mathcal{B}(\Sigma, \mathsf{E})$, we prove that $S \subseteq \Sigma$ is generically a set of sensors for all possible values of (A, D) if and only if there is a node-disjoint path cover of the network where each path of length $\neq 1$ ends in S (*SI Appendix*, Theorem 2 of section 2). If the conditions of this theorem are satisfied, then the sensors guarantee their generic observability, regardless of the within-animals and within-plants interactions that may exist. Recall that a "path" is a sequence of nodes of the form $\{\mathsf{v}_1 \to \mathsf{v}_2 \to \cdots \to \mathsf{v}_\ell\}$ with $\mathsf{v}_i \in \Sigma$ and $(\mathsf{v}_{i-1}, \mathsf{v}_i) \in \mathsf{E}$ for $i = 1, 2, \ldots, \ell$. The node v_ℓ is the end node of this path. The length of a path is the number of edges it contains. An isolated node $\{\mathsf{v}_i\}$ is a trivial path of zero length. A "path cover" of a network is a collection of paths where each node belongs to at least one path. A path cover is "node-disjoint" if each node belongs to exactly one path. Our criterion for sensors follows from three facts: first, the classic characterization of structural observability for linear systems (ref. 23 and *SI Appendix*, section 1B). Second, the structure of mutualistic systems is described by bipartite networks. Finally, adding edges to a system (in this case, interactions within animals or within plants) cannot damage its structural observability (ref. 25 and *SI Appendix*, Lemma 1 in section 1B).

MSS and Sensor Score. Using *SI Appendix*, Theorem 2 of section 2, it is possible to identify one MSS by calculating a maximum matching on an augmented bipartite network (*SI Appendix*, section 2E).

In general, a system has more than one MSS. Given a bipartite network characterizing the system structure, the sensor score s_i of the i-th species is the probability that species i is in an MSS. Our characterization of sensors implies that $s_i < 1$ for all i (i.e., no species appear in all MSSs). Indeed, $s_i = 1$ if and only if species i had a single incoming edge (which is impossible because all edges in $\mathcal{B}$ are bidirectional). Note also that in a fully connected bipartite network, any species is a solo MSS. In this case, the sensor score is $s_i = 1/N$ for all species, and the system structure does not contain any information to decide which species is better to monitor.

In *SI Appendix*, section 2F, we introduce a random sampling method to estimate the sensor score. We validated the accuracy of this method by comparing the estimated sensor score with the exact sensor score obtained by calculating all maximum matchings (47). Calculating all maximum matchings is computationally infeasible in large networks. Our random sampling method estimates the sensor score for networks of about $N = 80$ species in 70 s with a relative error $< 0.1\%$ on a laptop computer.

Simulating Critical Transitions. We used the model introduced in ref. 31 to simulate the dynamics of mutualistic systems with n plants and m animals. The model reads as Eq. **1** with

$$\begin{aligned} f_i(x,y) &= x_i\left(\alpha_i^{(P)} - \sum_{k=1}^{n}\beta_{ik}^{(P)}\,x_k + \frac{\sum_{j=1}^{m}\mu\,\gamma_{ij}^{(P)}\,y_j}{1+h\sum_{j=1}^{m}\mu\,\gamma_{ij}^{(P)}\,y_j}\right), \\ g_j(x,y) &= y_j\left(\alpha_j^{(A)} - \sum_{k=1}^{m}\beta_{jk}^{(A)}\,y_k + \frac{\sum_{i=1}^{n}\mu\,\gamma_{ji}^{(A)}\,x_i}{1+h\sum_{i=1}^{n}\mu\,\gamma_{ji}^{(A)}\,x_i}\right). \end{aligned} \quad [4]$$

Above, $f_i(x,y)$ and $g_j(x,y)$ denote the i-th row and j-th row of the vectors $f(x,y)$ and $g(x,y)$, respectively. The superscripts (A) and (P) refer to animals and plants, respectively. The parameters α_i are the intrinsic growth rates. The parameters $\beta_{ik} \geq 0$ represent the intraspecific competition strengths within plants and within animals when $i = k$ and the interspecific competition strengths when $i \neq k$. Plants and animals experience mutualistic benefits modeled by using a saturating functional response with handling time $h > 0$. The parameter $\mu \geq 0$ scales the mutualistic interaction strength across all species, and $\gamma_{ij} \geq 0$ characterizes the strength of the mutualistic benefit between species j and species i. If species i and j do not have a mutualistic interaction, then $\gamma_{ij} = \gamma_{ji} = 0$. Otherwise, these parameters are given by

$$\gamma_{ij}^{(P)} = \gamma_0/\deg(\mathsf{x}_i)^{\delta}, \quad \gamma_{ji}^{(A)} = \gamma_0/\deg(\mathsf{y}_j)^{\delta}.$$

Here, $\deg(\mathsf{x}_i)$ and $\deg(\mathsf{y}_j)$ denote the degree (number of edges) of the i-th plant species and j-th animal species in the bipartite network $\mathcal{B}$, respectively. Above, $\gamma_0 > 0$ is the baseline (or basal) mutualistic benefit, and $\delta > 0$ is the mutualistic tradeoff (34).

We studied three sets of parameter values for the above model (*SI Appendix*, section 3C and Table S3). Each parametrization follows (6). The results presented in *Validation of Sensors in a Large Dataset of Empirical Networks* and *Characterizing the Performance of Sensors* consider the "nominal" parameter values. *SI Appendix*, section 4F shows the performance of sensors for the other two parameter sets, with "low" and "high" values. Additionally, *SI Appendix*, section 4G analyses the case when the mutualistic strength is reduced more significantly.

To induce a critical transition, we decreased μ from $\mu = 1$ to $\mu = 0$ using 70 equal steps. We ran our simulations for 5,000 time steps before every change, discarded the first 200 samples to avoid transients, and recorded every species' abundance for the remaining time steps. Following ref. 6, we used only feasible realizations of the randomly chosen parameters and solutions of the SDE where every variable in Eq. **1** has a positive equilibrium.

Early Warning Score. To calculate the early warning score w_i for species i in a particular system, we first computed the variance of its abundance at each of the 70 steps described above. Then, we used a rolling window of 30% of the data points to perform a linear regression (see *SI Appendix*, section 3D for details). We interpret the regression slope as the change rate of the variance over the corresponding window. A positive slope represents a variance increase at that particular value of μ, and a negative slope represents a variance decrease. We estimate the early warning score of species i as w_i if the variance increases monotonically for all $\mu \leq w_i$.

Early Warning Score Ratio when Choosing Species by Sensor Score, Degree, or Randomly. Let Σ_s and Σ_d denote the species set Σ ordered from high to low sensor score and from low to high degree, respectively. If two or more species have identical sensor scores or degrees, we order those species randomly, producing different realizations of the species order in such a case. Let Σ_{rand} be a random permutation of Σ. To calculate the early warning score ratio $r(q)$ for a given fraction $q \in [0, 1]$ of species, as defined in Table 1, when choosing species by sensor score (respectively, by degree or randomly), we build the set $S_q \subseteq \Sigma$ choosing the first $\lfloor qN \rfloor$ species in Σ_s (respectively, Σ_d or Σ_{rand}). Therefore, $r(q)$ is a random variable depending on the realization of the system dynamics and the realization of the species order (see details in *SI Appendix*, section 3E).

Data Availability. Data and code to reproduce our results and calculate the sensor score and effective network structure is archived in GitHub (https://github.com/SyntheticDynamics/SensorSpecies).

Previously published data were used for this work. This work has used the Web of Life dataset (https://www.web-of-life.es).

ACKNOWLEDGMENTS. M.T.A. was supported by Consejo Nacional de Ciencia y Tecnología Grant A1-S-13909 and Programa de Apoyo a Proyectos de Investigacíon e Innovación Tecnológica Grant 104915. Y.-Y.L. was supported by NIH Grants R01AI141529, R01HD093761, RF1AG067744, UH3OD023268, U19AI095219, and U01HL089856. We thank Serguei Saavedra, Mario Santana, Hugo Flores, Aaron Kelley, and Chuliang Song for comments. We thank the Laboratorio Nacional de Visualización Científica Avanzada for computational resources. This work used the Web of Life dataset.

1. G. A. De Leo, S. Levin, The multifaceted aspects of ecosystem integrity. *Conserv. Ecol.* **1**, 3 (1997).
2. E. O. Wilson, *Half-Earth: Our Planet's Fight for Life* (W. W. Norton & Company, New York, 2016).
3. M. Scheffer, *Critical Transitions in Nature and Society* (Princeton Studies in Complexity, Princeton University Press, Princeton, NJ, 2020), vol. 16.
4. L. A. Burkle, J. C. Marlin, T. M. Knight, Plant-pollinator interactions over 120 years: Loss of species, co-occurrence, and function. *Science* **339**, 1611–1615 (2013).
5. N. E. Rafferty, A. R. Ives, Effects of experimental shifts in flowering phenology on plant-pollinator interactions. *Ecol. Lett.* **14**, 69–74 (2011).
6. V. Dakos, J. Bascompte, Critical slowing down as early warning for the onset of collapse in mutualistic communities. *Proc. Natl. Acad. Sci. U.S.A.* **111**, 17546–17551 (2014).
7. J. J. Lever, E. H. van Nes, M. Scheffer, J. Bascompte, The sudden collapse of pollinator communities. *Ecol. Lett.* **17**, 350–359 (2014).

8. J. Jiang *et al.*, Predicting tipping points in mutualistic networks through dimension reduction. *Proc. Natl. Acad. Sci. U.S.A.* **115**, E639–E647 (2018).
9. C. Kuehn, A mathematical framework for critical transitions: Bifurcations, fast–slow systems and stochastic dynamics. *Physica D* **240**, 1020–1035 (2011).
10. S. R. Carpenter *et al.*, Early warnings of regime shifts: A whole-ecosystem experiment. *Science* **332**, 1079–1082 (2011).
11. L. Dai, D. Vorselen, K. S. Korolev, J. Gore, Generic indicators for loss of resilience before a tipping point leading to population collapse. *Science* **336**, 1175–1177 (2012).
12. M. Scheffer *et al.*, Anticipating critical transitions. *Science* **338**, 344–348 (2012).
13. C. Boettiger, N. Ross, A. Hastings, Early warning signals: The charted and uncharted territories. *Theor. Ecol.* **6**, 255–264 (2013).
14. M. L. Pace *et al.*, Reversal of a cyanobacterial bloom in response to early warnings. *Proc. Natl. Acad. Sci. U.S.A.* **114**, 352–357 (2017).
15. M. C. Boerlijst, T. Oudman, A. M. de Roos, Catastrophic collapse can occur without early warning: Examples of silent catastrophes in structured ecological models. *PLoS One* **8**, e62033 (2013).
16. V. Dakos, Identifying best-indicator species for abrupt transitions in multispecies communities. *Ecol. Indic.* **94**, 494–502 (2018).
17. C. Boettiger, A. Hastings, Quantifying limits to detection of early warning for critical transitions. *J. R. Soc. Interface* **9**, 2527–2539 (2012).
18. A. S. Gsell *et al.*, Evaluating early-warning indicators of critical transitions in natural aquatic ecosystems. *Proc. Natl. Acad. Sci. U.S.A.* **113**, E8089–E8095 (2016).
19. A. C. Patterson, A. G. Strang, K. C. Abbott, When and where we can expect to see early warning signals in multispecies systems approaching tipping points: Insights from theory. *Am. Nat.* **198**, E12–E26 (2021).
20. J. Bascompte, P. Jordano, Plant-animal mutualistic networks: The architecture of biodiversity. *Annu. Rev. Ecol. Evol. Syst.* **38**, 567–593 (2007).
21. J. Bascompte, P. Jordano, *Mutualistic Networks* (Monographs in Population Biography, Princeton University Press, Princeton, NJ, 2013), vol. 53.
22. Bascompte Lab, Web of life: Ecological networks database. https://www.web-of-life.es/. Accessed 19 December 2020.
23. J. M. Dion, C. Commault, J. Van Der Woude, Generic properties and control of linear structured systems: A survey. *Automatica* **39**, 1125–1144 (2003).
24. Y. Y. Liu, J. J. Slotine, A. L. Barabási, Observability of complex systems. *Proc. Natl. Acad. Sci. U.S.A.* **110**, 2460–2465 (2013).
25. M. T. Angulo, A. Aparicio, C. H. Moog, Structural accessibility and structural observability of nonlinear networked systems. *IEEE Trans. Netw. Sci. Eng.* **7**, 1656–1666 (2019).
26. G. Ramos, A. P. Aguiar, S. Pequito, Structural systems theory: An overview of the last 15 years. *arXiv* [Preprint] (2020). https://arxiv.org/abs/2008.11223 (Accessed 3 February 2021).
27. D. Piovani, J. Grujić, H. J. Jensen, Linear stability theory as an early warning sign for transitions in high dimensional complex systems. *J. Phys. A Math. Theor.* **49**, 295102 (2016).
28. S. Chen, E. B. O'Dea, J. M. Drake, B. I. Epureanu, Eigenvalues of the covariance matrix as early warning signals for critical transitions in ecological systems. *Sci. Rep.* **9**, 2572 (2019).
29. A. Ghadami, S. Chen, B. I. Epureanu, Data-driven identification of reliable sensor species to predict regime shifts in ecological networks. *R. Soc. Open Sci.* **7**, 200896 (2020).
30. P. Hartman, On the local linearization of differential equations. *Proc. Am. Math. Soc.* **14**, 568–573 (1963).
31. U. Bastolla *et al.*, The architecture of mutualistic networks minimizes competition and increases biodiversity. *Nature* **458**, 1018–1020 (2009).
32. M. Scheffer *et al.*, Early-warning signals for critical transitions. *Nature* **461**, 53–59 (2009).
33. J. Bascompte, P. Jordano, C. J. Melián, J. M. Olesen, The nested assembly of plant-animal mutualistic networks. *Proc. Natl. Acad. Sci. U.S.A.* **100**, 9383–9387 (2003).
34. S. Saavedra, R. P. Rohr, V. Dakos, J. Bascompte, Estimating the tolerance of species to the effects of global environmental change. *Nat. Commun.* **4**, 2350 (2013).
35. A. A. Siddig, A. M. Ellison, A. Ochs, C. Villar-Leeman, M. K. Lau, How do ecologists select and use indicator species to monitor ecological change? insights from 14 years of publication in ecological indicators. *Ecol. Indic.* **60**, 223–230 (2016).
36. G. Yan *et al.*, Network control principles predict neuron function in the *Caenorhabditis elegans* connectome. *Nature* **550**, 519–523 (2017).
37. A. Vinayagam *et al.*, Controllability analysis of the directed human protein interaction network identifies disease genes and drug targets. *Proc. Natl. Acad. Sci. U.S.A.* **113**, 4976–4981 (2016).
38. R. P. Rohr, S. Saavedra, J. Bascompte, Ecological networks. On the structural stability of mutualistic systems. *Science* **345**, 1253497 (2014).
39. S. Saavedra *et al.*, A structural approach for understanding multispecies coexistence. *Ecol. Monogr.* **87**, 470–486 (2017).
40. L. P. Medeiros, K. Boege, E. Del-Val, A. Zaldívar-Riverón, S. Saavedra, Observed ecological communities are formed by species combinations that are among the most likely to persist under changing environments. *Am. Nat.* **197**, E17–E29 (2021).
41. E. F. Cagua, K. L. Wootton, D. B. Stouffer, Keystoneness, centrality, and the structural controllability of ecological networks. *J. Ecol.* **107**, 1779–1790 (2019).
42. M. T. Angulo, C. H. Moog, Y. Y. Liu, A theoretical framework for controlling complex microbial communities. *Nat. Commun.* **10**, 1045 (2019).
43. E. A. Coayla-Teran, P. R. Ruffino, Stochastic versions of Hartman–Grobman theorems. *Stoch. Dyn.* **4**, 571–593 (2004).
44. R. E. Kalman, Mathematical description of linear dynamical systems. *J. Soc. Ind. Appl. Math. Ser. A Control* **1**, 152–192 (1963).
45. D. G. Luenberger, *Introduction to Dynamic Systems: Theory, Models, and Applications* (John Wiley & Sons, New York, NY, 1979).
46. M. Davis, *Linear Estimation and Stochastic Control: A Halsted Press Book* (Chapman and Hall, London, 1977).
47. T. Uno, "Algorithms for enumerating all perfect, maximum and maximal matchings in bipartite graphs" in *International Symposium on Algorithms and Computation*, H. W. Leong, H. Imai, S. Jain, Eds. (Lecture Notes in Computer Science, Springer, Berlin, 1997), vol. 1350, pp. 92–101.