PHILOSOPHICAL TRANSACTIONS B

royalsocietypublishing.org/journal/rstb

Research

Cite this article: Ghosh S, Cottingham KL, Reuman DC. 2021 Species relationships in the extremes and their influence on community stability. *Phil. Trans. R. Soc. B* **376**: 20200343. https://doi.org/10.1098/rstb.2020.0343

Accepted: 17 June 2021

One contribution of 17 to a theme issue 'Synchrony and rhythm interaction: from the brain to behavioural ecology'.

Subject Areas:
ecology, environmental science, plant science, theoretical biology

Keywords:
variance ratio, community stability, community variability, synchrony, tail association, compensatory dynamics

Author for correspondence:
Daniel C. Reuman
e-mail: reuman@ku.edu

†2101 Constant Avenue, Lawrence, KS, 66047.

Electronic supplementary material is available online at https://doi.org/10.6084/m9.figshare.c.5506726.

THE ROYAL SOCIETY PUBLISHING

Species relationships in the extremes and their influence on community stability

Shyamolina Ghosh[1], Kathryn L. Cottingham[2] and Daniel C. Reuman[1,3,†]

[1]Department of Ecology and Evolutionary Biology and Kansas Biological Survey, University of Kansas, Lawrence, KS 66045, USA
[2]Department of Biological Sciences, Dartmouth College, Hanover, NH 03755, USA
[3]Laboratory of Populations, Rockefeller University, 1230 York Avenue, New York, NY 10065, USA

DCR, 0000-0002-1407-8947

Synchrony among population fluctuations of multiple coexisting species has a major impact on community stability, i.e. on the relative temporal constancy of aggregate properties such as total community biomass. However, synchrony and its impacts are usually measured using covariance methods, which do not account for whether species abundances may be more correlated when species are relatively common than when they are scarce, or vice versa. Recent work showed that species commonly exhibit such 'asymmetric tail associations'. We here consider the influence of asymmetric tail associations on community stability. We develop a 'skewness ratio' which quantifies how much species relationships and tail associations modify stability. The skewness ratio complements the classic variance ratio and related metrics. Using multi-decadal grassland datasets, we show that accounting for tail associations gives new viewpoints on synchrony and stability; e.g. species associations can alter community stability differentially for community crashes or explosions to high values, a fact not previously detectable. Species associations can mitigate explosions of community abundance to high values, increasing one aspect of stability, while simultaneously exacerbating crashes to low values, decreasing another aspect of stability; or vice versa. Our work initiates a new, more flexible paradigm for exploring species relationships and community stability.

This article is part of the theme issue 'Synchrony and rhythm interaction: from the brain to behavioural ecology'.

1. Introduction

Understanding how the dynamics of individual species within an ecological community combine to determine the temporal stability of key aggregated properties of the system as a whole is a topic that has fascinated ecologists for decades [1–4]. An important early insight was that an aggregate community property such as the total biomass of all species can be relatively stable through time, even while the dynamics of individual species are highly variable, so long as different species exhibit offsetting, asynchronous or partially asynchronous fluctuations [5,6]. Such fluctuations are often referred to as compensatory dynamics, because decreases in some species' abundances are compensated for by simultaneous increases in other species [5–8]. However, when species instead exhibit dynamics that are positively correlated through time—synchrony—the aggregate community property tends to have increased variability [3,9–11]. Other fields, with which some readers may be more familiar, have very similar notions of synchrony. For instance, synchronous volume fluctuations of insect mating calls can produce large oscillations in the total volume of an acoustic signal [12]. Thus, species relationships, and specifically the degree of synchrony between the population dynamics of different species, are important contributors to the stability of an aggregate community property. Here, the terminology 'species relationships' encompasses direct species interactions as well as related responses to environmental and other drivers. In addition to total biomass, aggregated community properties can include total abundance, net primary productivity, decomposition rate, and various nutrient cycling rates.

Synchronous versus compensatory dynamics, and their influence on community stability, have been studied in a wide variety of ecological communities (e.g. [5,13,14]). Synchrony is often owing to common responses to environmental changes (e.g. [5,14–16]), whereas compensatory dynamics result from multiple mechanisms, including competitive interactions and differing responses to environmental changes (e.g. [4,14,15,17,18]). Severe environmental stressors that impact all populations are expected to synchronize dynamics, whereas stressors with differential impacts allow for compensatory dynamics [19]. Although past work on the general concept of 'community stability' has led to a myriad of precisely defined alternative measures of 'stability' [20], we here focus on the common practice of examining the variability, through time, of an aggregate community property such as the total biomass of all species, referring to this henceforth as 'community variability.'

Much past work on species relationships and their influence on community variability has examined covariances between population time series, to characterize species relationships, and the temporal variance of an aggregate community property such as total biomass, to characterize community variability [5,6,10,19,21,22]. For instance, if $x_i(t)$ are data representing the population abundance or biomass of species $i = 1, \ldots, N$ at the sampling times $t = 1, \ldots, T$, and if $\mu_i = \mathrm{mean}(x_i)$, $v_{ii} = \mathrm{var}(x_i)$ and $v_{ij} = \mathrm{cov}(x_i, x_j)$ (here $j = 1, \ldots, N$ is another species index), community variability has commonly been quantified using the squared coefficient of variation of $x_{\mathrm{tot}}(t) = \sum_i x_i(t)$, i.e. $V_{\mathrm{com}} = \mathrm{var}(x_{\mathrm{tot}})/(\mathrm{mean}(x_{\mathrm{tot}}))^2 = \left(\sum_{i,j} v_{ij}\right)/\left(\sum_i \mu_i\right)^2$. This quantity obviously connects to species relationships measured via the covariances, v_{ij}, which appear directly in the formula. Denoting by $V_{\mathrm{ind}} = \left(\sum_i v_{ii}\right)/\left(\sum_i \mu_i\right)^2$ the value that V_{com} would take if the dynamics of each species were independent of the dynamics of other species (so $v_{ij} = 0$ for all $i \neq j$), the classic *variance ratio* was defined [5] as $\phi_V = V_{\mathrm{com}}/V_{\mathrm{ind}} = \mathrm{var}(x_{\mathrm{tot}})/\left(\sum_i v_{ii}\right) = \left(\sum_{i,j} v_{ij}\right)/\left(\sum_i v_{ii}\right)$. Thus $V_{\mathrm{com}} = \phi_V V_{\mathrm{ind}}$. Because V_{ind} has been interpreted as what community variability would be, without species relationships, ϕ_V is the factor by which species relationships inflate or decrease community variability. The variance ratio has therefore been used [6,8] as an index of whether dynamics are synchronous or compensatory. If $\phi_V > 1$, the interpretation has been that dynamics are, on balance, synchronous because then $V_{\mathrm{com}} > V_{\mathrm{ind}}$, i.e. the aggregate property x_{tot} is more variable than it would be if species dynamics were independent. Conversely, if $\phi_V < 1$, the interpretation has been that dynamics are compensatory [10].

Extensions, improvements and alternatives to the original variance ratio have been proposed, notably the synchrony metric of Loreau & de Mazancourt [10], here denoted $\phi_{V,\mathrm{LdM}}$. Whereas newer metrics such as $\phi_{V,\mathrm{LdM}}$ have interpretive and other advantages over the original variance ratio [10,23], in this study, we show that the ways in which species relationships influence community variability are likely to be more nuanced than is revealed by either the variance ratio approach or by any existing extensions or alternatives. Whereas the variance ratio compares V_{com} to the baseline value V_{ind} that would be the community CV^2 under an assumption of species independence (by considering the ratio $\phi_V = V_{\mathrm{com}}/V_{\mathrm{ind}}$), $\phi_{V,\mathrm{LdM}}$ instead compares V_{com} to the baseline value V_{syn} that would be the community CV^2 under an assumption of perfect linear correlation between species dynamics (again by considering a ratio): $\phi_{V,\mathrm{LdM}} = V_{\mathrm{com}}/V_{\mathrm{syn}}$. It is straightforward to show [23] that $V_{\mathrm{syn}} = \left(\sum_i \sqrt{v_{ii}}\right)^2/\left(\sum_i \mu_i\right)^2$, i.e. that this expression is what community CV^2 would be under perfect linear correlation between species [23]; and therefore that $\phi_{V,\mathrm{LdM}} = \left(\sum_{i,j} v_{ij}\right)/\left(\sum_i \sqrt{v_{ii}}\right)^2$. Timescale-specific alternatives and extensions of the original variance ratio approach have also been developed [24–26]. Although each alternative metric improves upon the original variance ratio in at least some respects, we nevertheless construct the specific statistical tools of this paper as extensions of the original variance ratio. This is for mathematical simplicity, and also because the choice makes little difference: the concepts of this study can be applied as improvements to all the existing frameworks of which we are aware. We return to this topic in the Discussion.

We begin by giving an intuitive introduction to the main concepts of this study; we do so by considering two simulated community dynamics datasets (figure 1*a,b*; data generated as in the electronic supplementary material, S1; see also figure S1). Both communities consist of $N = 20$ species. Data were generated so that species marginal distributions (and therefore the means, μ_i, and variances, v_{ii}) were the same, up to sampling variation, in both scenarios. In other words, the distribution of abundances, through time, of any given species i was the same for both scenarios. Likewise, data were generated so that species covariances, v_{ij} for $i \neq j$, were essentially the same for figure 1*a,b*. Thus both V_{com} and V_{ind}, which depend only on the μ_i and v_{ij}, were the same in both scenarios, and so ϕ_V was the same. Because V_{syn} also depends only on the μ_i and v_{ii}, $\phi_{V,\mathrm{LdM}}$ was also the same in our two scenarios, up to sampling variation (figure 1*a,b*). Thus classical approaches do not distinguish between the scenarios, neither in the nature of species relationships nor in the effects of relationships on community aggregate dynamics $x_{\mathrm{tot}}(t)$. Any approach that considers only species marginal distributions through time and the covariances v_{ij} will not detect the differences between our two scenarios, by construction. The metric of [22] is another such.

Nevertheless, species relationships and aggregate dynamics differed in important and related ways in our scenarios. In both scenarios, the 20 species could be separated into two groups of 10 species (the red lines on figure 1*a,b* are one group and the black lines are the other), and the dynamics of species in the different groups were compensatory. However, in scenario 1 (figure 1*a*), species in the same group exhibited strongly synchronous dynamics when those species were scarce (below about 5.5 on the y-axis of figure 1*a*), and much less synchronous dynamics when those species were common (above about 5.5); whereas in scenario 2 (figure 1*b*) species in the same group exhibited strongly synchronous dynamics when common and much less synchronous dynamics when scarce. Dynamics like scenario 1 could occur ecologically if red species in figure 1*a* were sensitive to an environmental factor, F, but only when it goes below a threshold. So all the red species are controlled by the same factor, F, when it is below the threshold, and hence are synchronous when scarce. When F is above the threshold, the red species are each sensitive, instead, to other, distinct factors, rendering them asynchronous when common. Black species of figure 1*a* may also be sensitive to F, but in a reverse manner, being all adversely impacted by F when it is above the threshold, and unaffected by F

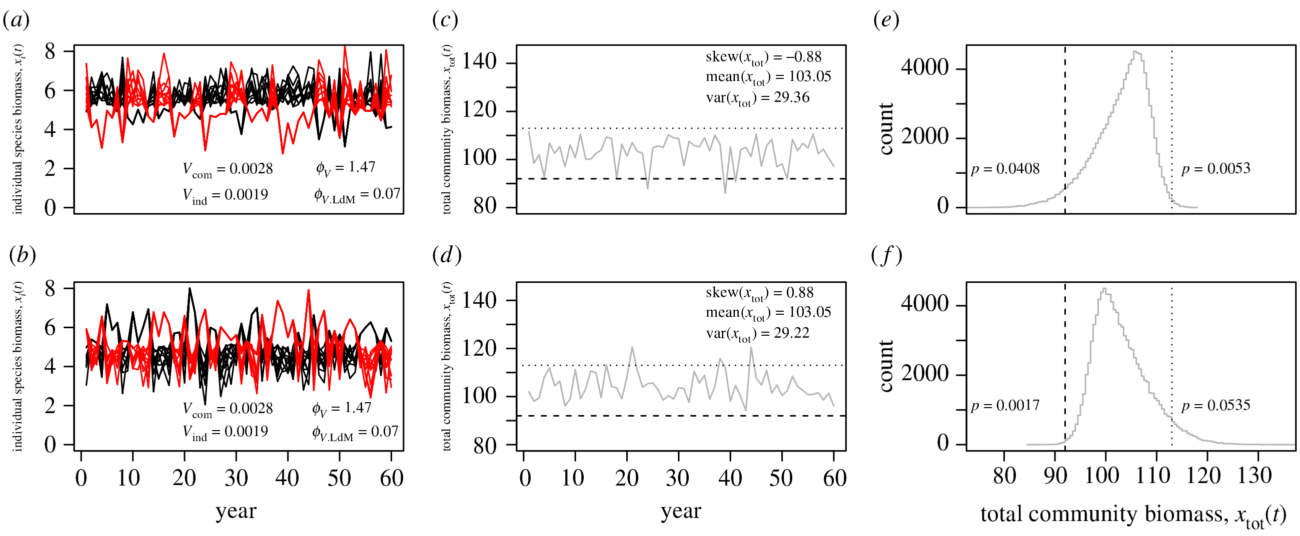

Figure 1. Pedagogical example showing two communities (*a,b*) of 20 species each that are not distinguishable via the classic variance ratio approach and related approaches, but that nonetheless differ markedly in potentially important ways. See Introduction for details and interpretation. See the electronic supplementary material, figure S1 for larger versions of (*a,b*). Dashed and dotted lines on (*c,f*) represent thresholds. If the lower threshold (dashed line) were a disastrous threshold for $x_{tot}(t)$ to cross, then the first scenario (*c,e*) would involve much more risk than the second scenario (see probabilities of crossing the threshold listed to the left of the dashed lines on (*e,f*)—the probability for scenario 1 is an order of magnitude higher). On the other hand, if the upper threshold (dotted line) were a disastrous threshold for $x_{tot}(t)$, the second scenario would involve more risk (see probabilities to the right of the dotted lines on (*e,f*)—the probability for scenario 2 is much higher). Time series were generated and statistics computed for 100 000 years, but are plotted on (*a,b*) for 60 years only, for clarity. (Online version in colour.)

below the threshold. Resulting community dynamics (figure 1*c–f*) differed between scenarios because of the distinct species relationships. The aggregate quantity, x_{tot}, was more likely to crash to low values in the first scenario (figure 1*c,e*) than in the second scenario (figure 1*d,f*; compare the probabilities on (*e,f*) listed to the left of the dashed lines), whereas $x_{tot}(t)$ was more likely to explode to high values in the second scenario than in the first (compare the probabilities on (*e,f*) to the right of the dotted lines). Thus total abundance was left-skewed, frequently crashing to low values, for scenario 1 (*c,e*); but was right skewed, frequently exploding to high values, in scenario 2 (*d,f*).

It is not difficult to imagine ecological ramifications or applied significance of the distinction illustrated here. For instance, grazers subsisting on a grass mixture will have different growth prospects if that mixture exhibits occasional crashes (which may harm the grazers) compared to if the mixture shows occasional explosions of abundance (unlikely to harm them). Thus species relationships and their effects on community aggregate dynamics may differ in potentially ecologically meaningful ways not detected by classical approaches. The reliance of earlier approaches on variances, covariances and related linear tools renders those approaches unable to register the differences illustrated here.

This study addresses three gaps in the research literature. First, commonly used past descriptions of species relationships have been oversimplified, based primarily on covariances or correlations of species time series. The example above suggests that more nuanced measures should be developed that identify whether species abundances are more related to each other when the species are common, or scarce (figure 1). Our first goal (G1) is to develop an approach to studying relationships between species that complements classic approaches relying on covariances. Second, past measures of community variability have typically focused on the variance or coefficient of variation of an aggregate property like x_{tot}. The example above suggests that more nuanced or complementary measures should be developed, taking skewness into account (figure 1). Our goal (G2) is to develop an approach to studying community variability that complements the coefficient of variation. Finally, the example above suggests that means by which species relationships influence community variability should be studied using the new measures of G1 and G2. Our goal (G3) is to do so, using data from long-term studies from mixed-grass prairie in Hays, Kansas and from the Konza Prairie Long-Term Ecological Research (LTER) site, also in Kansas.

The core ideas of this study relate to extreme events, and to how relationships between constituents of an aggregate quantity change during such events. There are strong connections to risk pricing in finance. A portfolio of investments should reduce risk if constituent investments are negatively correlated or uncorrelated. However, if constituents become positively correlated, risk can be accentuated. Some observers have blamed the 2008 financial crisis on analysts estimating correlations between housing default risks based on data from normal times, not taking into account that in a crisis, default risks become more correlated. In crisis, financial products based on aggregating mortgages therefore became riskier. A single modelling approach [27] was apparently widely used [28]. The relationship between two real quantities, be they investment risks or species populations, can probably never be completely captured by a single number such as a covariance [29], nor considered to be constant through time. Just as for mortgage defaults risks, the abundances of two species may be more or less correlated under different circumstances. Classical approaches that characterize relationships between species using covariances may therefore neglect something important. For instance, correlations between species may change as global climate change progresses, or when extreme climatic events occur [30]. Financial analysts have learned these lessons and have begun developing more appropriate models [29]. Ecologists can benefit from application of the same methods. Overall, this study initiates a new and more flexible paradigm for conceptualizing synchronous and compensatory dynamics in ecosystems, and their influence on community variability.

2. Theory

We now illustrate theoretically how relationships between species, beyond covariances, can influence variability of the total community. The theory here relates to, but goes substantially beyond, the example of figure 1. As in the Introduction, let $x_i(t)$ be an abundance measure for species $i = 1, \ldots, N$ at time $t = 1, \ldots, T$. We use $N = 20$ and $T = 100\,000$ in the theoretical examples here. This large value of T was used to ensure that sampling effects do not obscure our theoretical results. We assume the multivariate random variables $x(t) = (x_1(t), \ldots, x_N(t))$ are independent and identically distributed for distinct times, t. It should be straightforward to replace this assumption with an appropriate assumption of stationarity and ergodicity, with no real change to results, but we adopt the stronger assumption for simplicity. To describe species relationships, classic approaches use the covariances $v_{ij} = \mathrm{cov}(x_i, x_j)$, and related quantities. We will instead consider how x_i and x_j are related in their distribution tails. To describe community variability, classic approaches use the variance of $x_{\mathrm{tot}} = \sum_i x_i$, or its coefficient of variation. But these are only summaries of the distribution of x_{tot}. We will consider the whole distribution, including its skewness.

We present four scenarios (scenarios 1–2 below and 3–4 in the electronic supplementary material, S2 and figure S2). For all scenarios, the marginal distributions x_i and the covariances v_{ij} are the same, as are the quantities $\mathrm{var}(x_{\mathrm{tot}})$, V_{com}, V_{ind}, V_{syn}, ϕ_V and $\phi_{V,\mathrm{LdM}}$. Thus, as in the example of figure 1, our scenarios are indistinguishable to classic approaches but are nevertheless ecologically distinct: the total abundance x_{tot} achieves more extreme values under some scenarios than others. Hence our scenarios represent communities showing different characteristics of variability. Scenarios are intentionally simplified, with complexities such as autocorrelation excluded, for clarity, but the ideas also apply to real communities (see below). We additionally compare each scenario to a reference *comonotonic* scenario (called scenario C), for which the x_i are related via perfect positive relationships. This is a scenario of maximal covariance between species, and therefore maximal values of $\mathrm{var}(x_{\mathrm{tot}})$ [29] and of community variability, given fixed species marginal distributions, x_i.

In scenario 1, which is a baseline, $(x_1, \ldots, x_N)$ is a multivariate normal distribution with mean $(0, \ldots, 0)$ and covariance matrix having 1s along the diagonal and off-diagonal entries 0.6. A sample from (x_1, x_2) (figure 2*a*) and marginal histograms (side panels of figure 2*a*) help illustrate the distribution. The standard deviation of x_{tot} was 15.7, and its skewness was close to 0 (figure 2*b*). The probability of x_{tot} exceeding a high threshold is given in figure 2*c* for a range of high thresholds. Likewise the probability of x_{tot} falling below a low threshold is in figure 2*d* for a range of low thresholds. These probabilities represent the chances that x_{tot} will surpass or fall below a threshold considered disastrous or unacceptable from an applied or ecological viewpoint, and we refer to these as *disaster thresholds*. It is possible to imagine applied scenarios in which exceeding a high threshold is disastrous, and other scenarios in which disaster instead occurs when x_{tot} falls below a low threshold. Our theory can be considered in relation to either of these applied scenarios, so we keep them both in mind while developing the theory.

For the comonotonic scenario C, we assume $x_i = x_1$ for all i. This is a specific form of comonotonicity [29], with the relationship between x_1 and x_2 pictured in figure 2*e*. Because synchrony between the dynamics of individual species is perfect, $\mathrm{sd}(x_{\mathrm{tot}})$ is much larger than in scenario 1 (figure 2*f*; the dashed black lines on that panel and the panels below it are reproduced from figure 2*b*, to facilitate comparisons). Probabilities of x_{tot} exceeding a high disaster threshold or falling below a low one are likewise larger (red lines in figure 2*g*,*h*; dashed black lines on those panels and the panels below them are reproduced from panels (*c*,*d*), respectively).

Scenarios 2–4 illustrate the ideas, previously little recognized in ecology, of tail comonotonicity [29] and tail associations [30–32], and the consequences of tail associations for x_{tot} and its extreme values. In scenario 2, as in all our scenarios, each x_i is standard-normally distributed (figure 2*i* side panels), but $(x_1, \ldots, x_N)$ is not a multivariate normal distribution. Instead, $(x_1, \ldots, x_N)$ was engineered so that x_i and x_j are perfectly related when they exceed some threshold, but imperfectly related below the threshold (figure 2*i*; see the electronic supplementary material, S2 for details of each scenario). The strength of association below the threshold was chosen to make $\mathrm{cov}(x_i, x_j)$ equal to the same value as in scenario 1, so $\mathrm{sd}(x_{\mathrm{tot}})$ was also the same, up to sampling variation, as in scenario 1 (figure 2*j*). (Recall that $\mathrm{sd}(x_{\mathrm{tot}}) = \sqrt{\mathrm{var}(x_{\mathrm{tot}})}$ and $\mathrm{var}(x_{\mathrm{tot}}) = \sum_{i,j} \mathrm{cov}(x_i, x_j)$.) Comonotonicity in the upper tails produces right/positive skewness in x_{tot} (figure 2*j*). Skewness translates to elevated probabilities of x_{tot} exceeding high disaster thresholds, compared to scenario 1 (blue lines in figure 2*k*; the dashed red lines on panels (*k*,*l*) are copied from the solid red lines of panels (*g*,*h*), respectively, to facilitate comparisons). The probability of exceeding sufficiently high disaster thresholds is actually the same as scenario C (figure 2*k*). So although scenarios 1 and 2 are indistinguishable to classic approaches, the probability of disaster in scenario 2, for high disaster thresholds, is much elevated, and is as high as it can possibly be given the species marginal distributions used. Although it does not appear unstable to classic approaches, the scenario 2 community is maximally unstable with respect to high disaster thresholds.

Scenario 3 parallels scenario 2, but with comonotonicity in the left tails instead of the right, and concomitant elevated probabilities of x_{tot} falling below a low disaster threshold. Scenario 4 shows comonotonicity in both tails, and thus elevated probabilities both of x_{tot} exceeding high disaster thresholds and falling below low ones (electronic supplementary material, S2 and figure S2).

Although scenarios 1–4 are indistinguishable to classic approaches, the probabilities of disaster are elevated in scenarios 2–4, and are maximal, given fixed species marginal distributions, for some disaster types (high or low disaster thresholds), depending on the scenario. Perfect comontonicity in the tails of population distributions x_i simplifies our presentation of the theoretical ideas, but is not necessary to generate skewness in x_{tot} and elevated disaster probabilities. Similar results pertain for essentially any case where associations between species are asymmetric in the tails.

Our scenarios were constructed using mathematical results of [29,33]. We refer the reader to those papers for

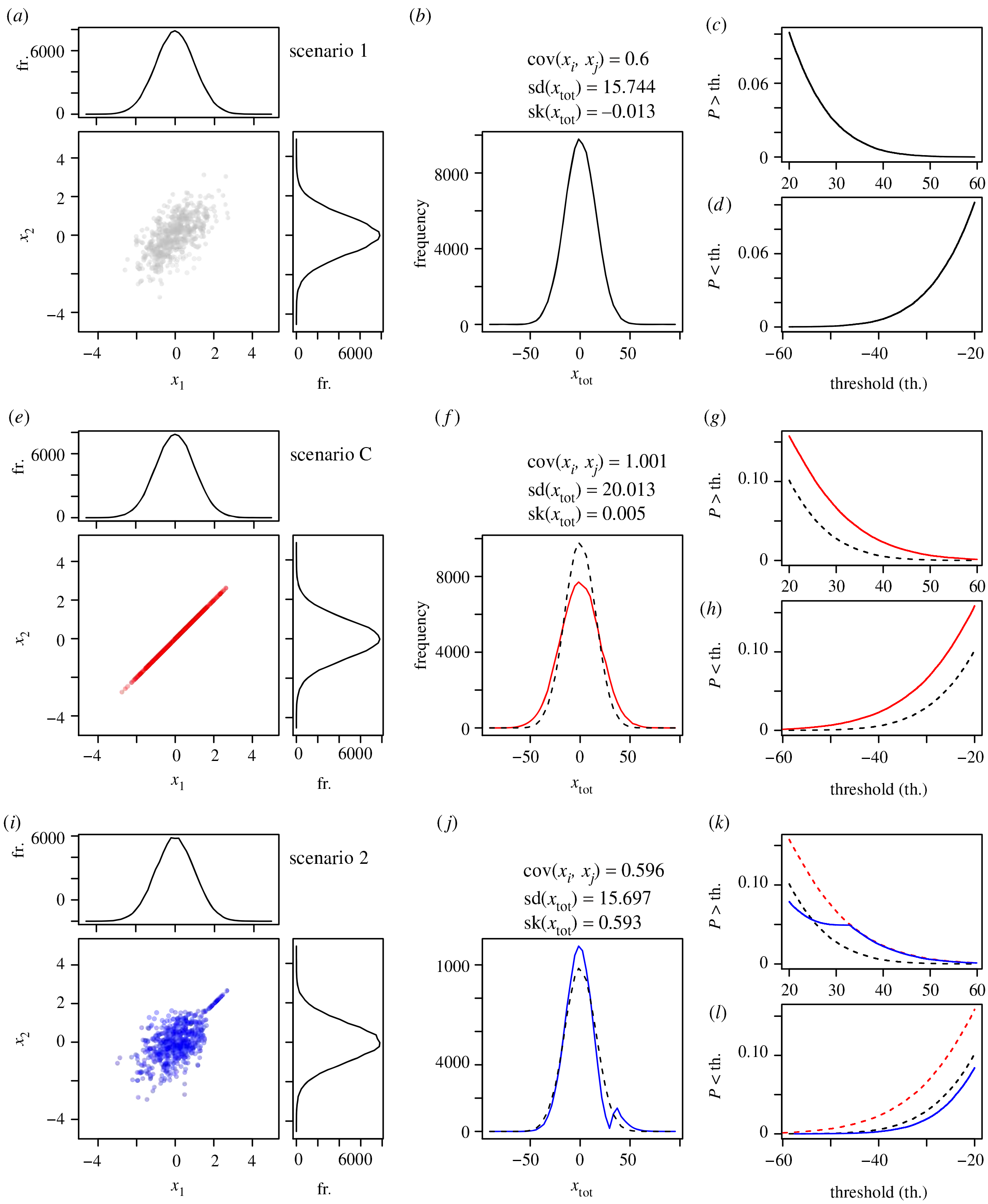

Figure 2. Theoretical scenarios demonstrating how species relationships can influence a community aggregate quantity in ways that are not detectable by classic approaches. The abbreviation 'sd' stands for 'standard deviation', 'sk' stands for 'skewness', 'th' stands for 'threshold' and 'fr' stands for 'frequency'. Scenario C is the reference, comonotonic scenario of maximal community variability. See the Theory section for other notation and interpretations, and the electronic supplementary material, section S2 and figure S2 for scenarios 3 and 4 and for mathematical details. (Online version in colour.)

mathematical specifics, while here summarizing aspects important for ecological applications. Given one-dimensional cumulative distribution functions $F_1, \ldots, F_N$, the *Fréchet space* $\mathcal{R}(F_1, \ldots, F_N)$ is the set of all N-dimensional random vectors $(x_1, \ldots, x_N)$ with the $F_1, \ldots, F_N$ as their marginal distribution functions. Thus the Fréchet space represents, in our modelling context, all possible interspecific relationships (different degrees and kinds of synchrony and compensatory dynamics between species) that can pertain, given fixed species marginal distributions. So the question of what forms the distribution of x_{tot} can take across the Fréchet space is a precisely formulated version of the classic question of how species interrelationships such as synchrony and compensatory dynamics can influence community variability. It is well known to statisticians that the sum x_{tot} exhibits maximal variance when the x_i are comonotonic [29]; this is the maximal-synchrony case. Cheung & Vanduffel [29] showed the converse, i.e. if $\mathrm{var}(x_{tot})$ is maximal in the Fréchet space, then $(x_1, \ldots, x_N)$ must be comonotonic (subject to some mild regularity assumptions about the F_i). The classical variance ratio and Loreau-de Mazancourt approaches essentially make sole use of $\mathrm{var}(x_{\mathrm{tot}})$ to quantify community variance (V_{com} is used, actually, but this is equivalent, since species marginals are fixed within the Fréchet space). However [29] showed that, even if we consider only the sub-Fréchet space consisting of random vectors $(x_1, \ldots, x_N)$ in $\mathcal{R}(F_1, \ldots, F_N)$ that additionally have $\mathrm{var}(x_{\mathrm{tot}}) = S$ for some fixed S, there are random vectors that achieve extreme values just as extreme as the comonotonic case (e.g. scenarios 2–4 in figure 2; electronic supplementary material, figure S2). Existing approaches in the ecological literature tell us nothing about true community variability, in the sense that if species relationships are tail associated, the community can exhibit the same extreme behaviour as the comonotonic case, while the classically important quantity $\mathrm{var}(x_{\mathrm{tot}})$ gives the impression of a much more stable community. We next transition to describing the data and the statistics used to assess to what extent the above theory applies to those data.

3. Methods

(a) Data

We used data from two grassland sites, both monitored for decades. After initial processing, data from the Hays, Kansas site consisted of annual estimates of basal cover, averaged over 36 quadrats from which livestock were excluded, for each species present, for the years 1932–1972 [34]. Data from Konza Prairie, after initial processing, consisted of annual canopy per cent cover data by species for the years 1983–2018, averaged over 20 annually burned plots on tully soils, ungrazed by livestock [35]. Additional details are in the electronic supplementary material, S3.

Community data often include numerous rare species. Rare species complicate analyses of interspecific relationships and their effect on community variability because the relationship of a rare species with another species is difficult to accurately assess. We therefore categorized species as ‘common’ (present in the system for at least 35 years), ‘rare’ (present for at most 2 years), or ‘intermediate’ (other species). Rare and intermediate species were combined into a single pseudo-species for analyses. Common species (electronic supplementary material, tables S1 and S2) made up the overwhelming majority of both the Hays and Konza communities (electronic supplementary material, figure S3), so ‘intermediate’ species were actually quite uncommon, and it is the variability of common species, and their relationships, that mainly determine community variability. Henceforth the terminology ‘species’ signifies both true species and the pseudo-species formed by combining rare and intermediate species.

(b) Statistical methods: quantifying tail associations

Figure 1 and the Theory section suggest that to accomplish goal G1 of the Introduction, we could apply measures of tail association between species, and asymmetries of tail association. We here define such measures, based on the *partial Spearman correlation* of [31]. Notation is summarized in the electronic supplementary material, table S3. Given $x_i(t)$ for $t = 1, \ldots, T$, we begin by defining the *normalized rank* $u_i(t)$, equal to the rank of $x_i(t)$ in the set $\{x_i(1), \ldots, x_i(T)\}$, divided by $T + 1$. The smallest element of this set is here considered to have rank 1. Plotting the normalized ranks $u_i(t)$ and $u_j(t)$ against each other for two positively associated species x_i and x_j provides a visual indication of tail association and asymmetries in tail association (e.g. the points of figure 3; see also the electronic supplementary material,

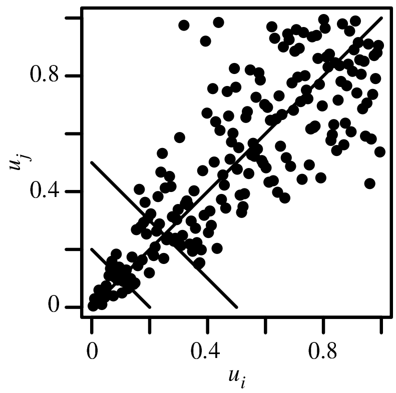

Figure 3. A normalized rank plot (see Methods) can visually help reveal that an association between variables is stronger in the lower tails of the variables' distributions than it is in the upper tails (as in this example), or *vice versa* (as in the electronic supplementary material, figure S4). In this example, points are clustered closer to the $u_i = u_j$ diagonal in the lower left than in the upper right, revealing stronger association between the variables in the lower tails of their distributions. The partial Spearman correlation, $\mathrm{cor}_{b_l,b_u}(x_i, x_j)$ (Methods), within a band defined by two bounds $0 \le b_l < b_u \le 1$ can be computed for any band to quantify the component of the overall Spearman correlation due to that band. Solid diagonal lines show the band associated with $b_l = 0.1$ and $b_u = 0.25$. Up to sampling variation, the (total) Spearman correlation is 0.8 on both this plot and the electronic supplementary material, figure S4, though patterns of tail association differ. This figure was slightly modified from fig. 4 of [30] and fig. 7 of [31].

figure S4), as described in [31]. Given two bounds $0 \le b_l < b_u \le 1$ (l stands for ‘lower’ and u for ‘upper’), we define the lines $u_i + u_j = 2b_l$ and $u_i + u_j = 2b_u$, which intersect the unit square $[0, 1] \times [0, 1]$ of the normalized rank plot (figure 3 shows these bounds for $b_l = 0.1$ and $b_u = 0.25$). The partial Spearman correlation associated with the band circumscribed by these bounds is

$$\mathrm{cor}_{b_l,b_u}(x_i, x_j) = \frac{\sum (u_i(t) - \mathrm{mean}(u_i))(u_j(t) - \mathrm{mean}(u_j))}{(T-1)\sqrt{\mathrm{var}(u_i)\mathrm{var}(u_j)}}, \tag{3.1}$$

where sample means and variances are computed over $t = 1, \ldots, T$ but the summation in the above formula is only computed over t such that $u_i(t) + u_j(t) > 2b_l$ and $u_i(t) + u_j(t) < 2b_u$ (e.g. the points in the band delineated by the two diagonal lines on figure 3). The partial Spearman correlation is the component of the Spearman correlation which can be attributed to the points in the band. For two positively associated variables, $\mathrm{cor}_{0,0.5}$ measures lower-tail association, and $\mathrm{cor}_{0.5,1}$ measures upper-tail association. The difference $\mathrm{cor}_{0,0.5} - \mathrm{cor}_{0.5,1}$ is a way to measure asymmetry of tail associations. Positive values of this statistic indicate stronger lower-tail association, and negative values indicate stronger upper-tail association. We henceforth use the shorthand cor_l for $\mathrm{cor}_{0,0.5}$ and cor_u for $\mathrm{cor}_{0.5,1}$. Genest & Favre [36] recommend making inferences about dependence structures such as tail associations using rank-based approaches such as the tools we introduced here. Work of Ghosh and colleagues [30–32] provides additional information on why rank-based approaches and the partial Spearman correlation are an appropriate statistical choice for purposes such as ours.

We computed $\mathrm{cor}_l - \mathrm{cor}_u$ for all positively associated species pairs (those with positive overall Spearman correlation); negatively associated pairs were ignored because theory suggested the importance of positively associated species pairs (but see the electronic supplementary material, S4 for thoughts on future work). Figure 1 and the Theory section suggest that communities with a preponderance of upper-tail association between positively associated species will have x_{tot} with distinct distributional properties from communities with more lower-tail

royalsocietypublishing.org/journal/rstb *Phil. Trans. R. Soc. B* **376**: 20200343

association. Community aggregated tail association between positively associated species was measured by counting the number, n_L, of values $\mathrm{cor}_l(x_i, x_j) - \mathrm{cor}_u(x_i, x_j)$, for $i \neq j$, that were positive, representing stronger lower- than upper-tail association between positively associated species; as well as the number, n_U, of values that were negative, representing stronger upper- than lower-tail association. We also computed the sum, A_{tot}, of $\mathrm{cor}_l(x_i, x_j) - \mathrm{cor}_u(x_i, x_j)$ across all positively associated species pairs, i, j. This was positive if lower-tail association was stronger, in aggregate, than upper-tail association, and negative if the reverse. We call A_{tot} the *total community tail association*. Note that the word 'positive' applies in three senses, here, all distinct: positive association of species pairs; positivity of $\mathrm{cor}_l(x_i, x_j) - \mathrm{cor}_u(x_i, x_j)$; and positivity of A_{tot}.

(c) Statistical methods: quantifying variability, and the skewness ratio

To accomplish goal G2 of the Introduction, we need a measure of the distribution x_{tot} that characterizes community variability in a way that complements $\mathrm{var}(x_{\mathrm{tot}})$ or V_{com}. We use the skewness, $\mathrm{sk}(x_{\mathrm{tot}})$ (see the electronic supplementary material, S4 for details). As illustrated in figure 1, skewness complements use of $\mathrm{var}(x_{\mathrm{tot}})$ or V_{com} to characterize community variability, because strongly negative values indicate that x_{tot} undergoes large downward departures from typical values (crashes), and large positive values indicate that x_{tot} undergoes upward departures (explosions); either of these dynamical behaviours can happen independently of the values of $\mathrm{var}(x_{\mathrm{tot}})$ and V_{com}. Because $\mathrm{sk}(x_{\mathrm{tot}})$ relates to upward or downward departures of x_{tot} from typical values, it also relates to probabilities that x_{tot} will surpass a high disaster threshold (see Theory) or fall below a low one.

Just as the variance ratio ϕ_V is the quotient $V_{\mathrm{com}}/V_{\mathrm{ind}}$, we analogously define a *skewness ratio*. Let $S_{\mathrm{com}} = \mathrm{sk}(x_{\mathrm{tot}})$, and consider the value that S_{com} would take if species dynamics were independent, i.e. $S_{\mathrm{ind}} = \frac{\sum_i m_3(x_i)}{\left(\sum_i v_{ii}\right)^{3/2}}$, where $m_3(x_i)$ is the third central moment of x_i (electronic supplementary material, S4). The *skewness ratio* is $\phi_S = S_{\mathrm{com}}/S_{\mathrm{ind}}$, so that $S_{\mathrm{com}} = \phi_S S_{\mathrm{ind}}$. It is also possible to define a skewness metric in a manner that attempts to parallel the choices made in constructing $\phi_{V,\mathrm{LdM}}$. This is elaborated in the Discussion.

Values of the skewness ratio provide a summary of how species relationships influence community variability that complements information provided by the variance ratio. The skewness ratio can be positive or negative because S_{com} and S_{ind} can be positive or negative, unlike V_{com} and V_{ind}. If $\phi_S > 1$, S_{com} has greater magnitude, but the same sign, as what its value would be in the absence of species relationships (S_{ind}). Assuming, for simplicity $\phi_V \approx 1$ (see the electronic supplementary material, S5 and table S4 for more on this assumption), if S_{ind} is positive, then $\phi_S > 1$ means $S_{\mathrm{com}} > S_{\mathrm{ind}} > 0$, i.e. species relationships magnify the tendency for the community aggregate quantity x_{tot} to explode to high values; and if S_{ind} is negative, then $\phi_S > 1$ means $S_{\mathrm{com}} < S_{\mathrm{ind}} < 0$, i.e. species relationships magnify the tendency for x_{tot} to crash to low values. This can be viewed as a new type of synchrony, in the sense that species relationships inflate community variability by accentuating the probability that x_{tot} will exceed large disaster thresholds, or fall below small ones.

If $0 < \phi_S < 1$, then S_{com} has lower magnitude, but the same sign, as S_{ind}. Again assuming, for simplicity, that $\phi_V \approx 1$ (see the electronic supplementary material, S5 for more on this assumption), if S_{ind} is positive, then $0 < \phi_S < 1$ means $0 < S_{\mathrm{com}} < S_{\mathrm{ind}}$, i.e. species relationships mitigate the tendency for x_{tot} to explode to high values; and if S_{ind} is negative, then $0 < \phi_S < 1$ means $S_{\mathrm{ind}} < S_{\mathrm{com}} < 0$, i.e. species relationships mitigate the tendency for x_{tot} to crash to low values. This can be viewed as a new type of compensatory dynamics, because species relationships reduce a kind of community variability.

If $\phi_S < 0$, what happens? If $S_{\mathrm{ind}} < 0$, so that without species relationships x_{tot} would have exhibited occasional crashes, we instead have $S_{\mathrm{com}} > 0$: with species relationships x_{tot} instead exhibits occasional explosions. If $S_{\mathrm{ind}} > 0$, so that without species relationships x_{tot} would have exhibited occasional explosions, we instead have $S_{\mathrm{com}} < 0$: with species relationships x_{tot} instead exhibits occasional crashes. Thus the tendency of x_{tot} to crash or explode is reversed by species relationships. We again assume, here, for simplicity, that $\phi_V \approx 1$; see the electronic supplementary material, S5.

Our skewness ratio does not supplant the variance ratio (or any of its alternatives, such as the Loreau-de Mazancourt metric), but rather complements it. Just as one understands more about a distribution by knowing both its variance and skewness, so one understands more about community dynamics by combining the variance and skewness ratio approaches. The values of ϕ_V and ϕ_S provide complementary information about how the distribution of x_{tot} is influenced by species relationships (see the electronic supplementary material, S5). See figure 4*a,b* for a summary of the two approaches.

(d) Statistical methods: the link from tail associations to variability

The skewness ratio $\phi_S = S_{\mathrm{com}}/S_{\mathrm{ind}}$ provides information about the influence of species relationships, of all kinds, on community skewness, $S_{\mathrm{com}} = \mathrm{sk}(x_{\mathrm{tot}})$, because S_{com} is influenced by species relationships of all kinds and S_{ind} is influenced by none. To understand the specific influence of species tail associations and asymmetries of tail associations on S_{com} (G3 of the Introduction), we developed a randomization procedure that rendered symmetric the tail associations between species, while keeping other statistical properties of community dynamics approximately fixed. Given the data $x_i(t)$ ($i = 1, \ldots, N$, $t = 1, \ldots, T$), the randomization procedure produced any desired number M of surrogate datasets $x_i^{(m)}(t)$, $m = 1, \ldots, M$, with properties detailed in the electronic supplementary material, S6 (see also the electronic supplementary material, figures S5–S7). Because these surrogate datasets had symmetric tail associations between species but were otherwise statistically similar to the original dataset, $x_i(t)$, $S_{\mathrm{com}} = \mathrm{sk}(x_{\mathrm{tot}})$ was computed for the original data, and a surrogate skewness value, $\mathrm{sk}\left(\sum_i x_i^{(m)}\right)$, was computed for each of the surrogate datasets. A test of whether asymmetry of tail associations between species significantly influenced S_{com} was obtained by examining whether S_{com} fell sufficiently in the tails of the distribution of surrogate values to meet a desired statistical confidence level. This approach produces a test of the null hypothesis that S_{com} was no more extreme than would have been expected by chance if species relationships were symmetric in their tail associations, against the alternative hypothesis that asymmetries of species relationships contribute meaningfully to community skewness. We used $M = 10\,000$ to ensure the p-values produced by this method were very precise.

The quantity S_{nta}, which represents what community skewness would be without asymmetries of tail association between species, was defined as the median of the values $\mathrm{sk}\left(\sum_i x_i^{(m)}\right)$; it is S_{nta}, and associated quantities which we will now define, that are used to accomplish goal G3 of the Introduction. Here and below, 'ta' stands for 'tail association' and 'nta' stands for 'no tail association'. We also defined the quantities $\phi_{S,\mathrm{ta}} = S_{\mathrm{com}}/S_{\mathrm{nta}}$ and $\phi_{S,\mathrm{cor}} = S_{\mathrm{nta}}/S_{\mathrm{ind}}$. The quantity $\phi_{S,\mathrm{cor}}$ satisfies $S_{\mathrm{nta}} = \phi_{S,\mathrm{cor}} S_{\mathrm{ind}}$, and therefore is the factor by which correlations between species, as distinct from asymmetries of tail association, influence community skewness. The subscript 'cor' is a reference to the effects of correlation. The quantity $\phi_{S,\mathrm{ta}}$ satisfies

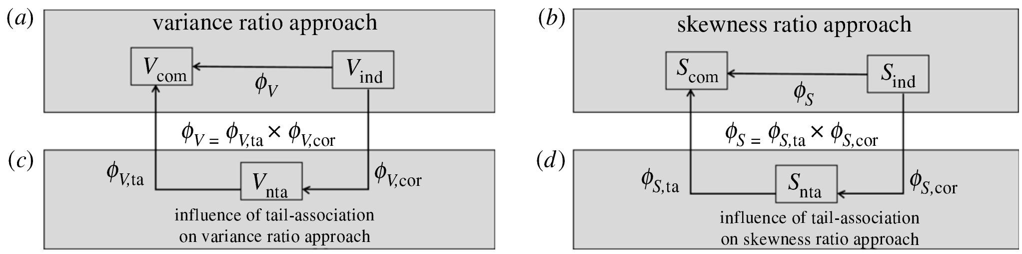

Figure 4. Summary of approach. Arrows in (*a*–*d*) imply an equation where the quantity in the box at the arrow tail, times the quantity labelled on the arrow itself, equals the quantity in the box at the arrow head. For instance, the arrow in (*a*) represents the equation $V_{\text{com}} = \phi_V V_{\text{ind}}$.

$S_{\text{com}} = \phi_{S,\text{ta}} S_{\text{nta}}$, and therefore is the factor by which asymmetric tail associations between species further influence community skewness. So a value of $\phi_{S,\text{ta}}$ which differs from 1 indicates that asymmetric tail associations have a role in determining S_{com}. Statistical significance, here, is judged as in the previous paragraph. It is straightforward to show that $\phi_S = \phi_{S,\text{ta}}\phi_{S,\text{cor}}$, so $S_{\text{com}} = \phi_{S,\text{ta}}\phi_{S,\text{cor}}S_{\text{ind}}$. Thus $\phi_{S,\text{ta}}$ and $\phi_{S,\text{cor}}$ separate the influence of species relationships on community skewness into factors owing to asymmetries of tail associations and correlations with symmetric tail associations. For comparison, V_{com} was also computed for the data and for surrogates, defining quantities V_{nta}, $\phi_{V,\text{cor}}$ and $\phi_{V,\text{ta}}$ which are related to each other in an analogous way. Figure 4*a*–*d* summarize the new quantities and their relationships. Because V_{com} depends only on species means and covariances, which are preserved by surrogates, we know, *a priori*, that $V_{\text{nta}} \approx V_{\text{com}}$ and $\phi_{V,\text{ta}} \approx 1$.

Having defined methods, we can now frame hypotheses suggested by figure 1 and the Theory section using the new methods; tests of the hypotheses will address goal G3 of the Introduction. Figures 1 and 2 and electronic supplementary material, figure S2 suggest that communities exhibiting a preponderance of lower-tail association should have total community dynamics that are more left-skewed, or less right-skewed, than would have been the case without these tail associations. This is the same as saying that for communities for which $n_L > n_U$ and $A_{\text{tot}} > 0$, we should have $S_{\text{com}} < S_{\text{nta}}$. Furthermore, figures 1 and 2 and electronic supplementary material, figure S2 suggest that communities exhibiting mostly upper-tail association should have total community dynamics that are more right-skewed, or less left-skewed, than would have been the case with symmetric tail associations. This is the same as saying that for communities for which $n_L < n_U$ and $A_{\text{tot}} < 0$, we should have $S_{\text{com}} > S_{\text{nta}}$. We test these hypothesis with the two grassland datasets in Results and discussion below.

4. Results and discussion

Addressing the goal G1 of the Introduction, of applying a new way of quantifying species relationships that complements covariance approaches, the quantities $\text{cor}_l - \text{cor}_u$ for all positively associated pairs of species within the Hays and Konza datasets were computed (electronic supplementary material, figure S8), as were the community aggregate quantities n_L and n_U and the total community tail association values A_{tot} (Methods). Results revealed marked differences between Hays and Konza in species' tail associations. Although both datasets had pairs of positively associated species with more lower-tail association and pairs with more upper-tail association, $n_L = 70$ was greater than $n_U = 40$ and $A_{\text{tot}} = 6.1$ was positive for Hays, and $n_L = 70$ was less than $n_U = 82$ and $A_{\text{tot}} = -1.3$ was negative for Konza. Thus lower-tail association was dominant at Hays and upper-tail association was dominant at Konza.

We earlier framed the hypotheses that communities exhibiting more lower- than upper-tail association should also have x_{tot} more left-skewed, or less right-skewed, than if tail associations were symmetric; and communities exhibiting more upper- than lower-tail association should have x_{tot} more right-skewed, or less left-skewed, than if tail associations were symmetric (Methods). We tested these hypotheses for Hays and Konza, thereby addressing goals G2 and G3 of the Introduction. The quantities of figure 4 are displayed for our data in figure 5. Because $n_L > n_U$ and $A_{\text{tot}} > 0$ for Hays (electronic supplementary material, figure S8A), S_{com} should be less than S_{nta}. This was confirmed (figure 5*b*). Both S_{com} and S_{nta} were positive and the component of the skewness ratio owing to tail association, $\phi_{S,\text{ta}} =$ 0.3, was less than 1. Becuse $n_L < n_U$ and $A_{\text{tot}} < 0$ for Konza (electronic supplementary material, figure S8B), S_{com} should be greater than S_{nta}. This was also confirmed (figure 5*d*). S_{com} and S_{nta} were again positive, so for Konza $\phi_{S,\text{ta}} = 2.8$ was greater than 1. These results were statistically significant: for Hays, S_{com} was less than a fraction 0.9973 of the analogous quantities computed using the Hays surrogates datasets (corresponding to $p = 1 - 0.9973 = 0.0027$ for a one-tailed test); whereas for Konza, S_{com} was greater than a fraction 0.9748 of the analogous surrogate values for Konza ($p = 0.0252$, one-tailed test; electronic supplementary material, figure S9). Thus asymmetric tail associations between species significantly modified the skewness of the total community abundance, x_{tot}, lowering it for Hays and raising it for Konza and thereby influencing community variability and probabilities of passing disaster thresholds. Figure 6 makes concrete some of the conclusions of this paragraph.

As anticipated (Methods), the variance ratio approach did not register the importance of tail associations for community variability and thus provided an incomplete picture of the influence of species relationships on community variability, i.e. $V_{\text{nta}} \approx V_{\text{com}}$ and $\phi_{V,\text{ta}} \approx 1$ for both Hays and Konza (figure 5*a*,*c*). This result was expected because V_{com} depends only on species means and covariances, which were preserved by the randomized, surrogate datasets from which V_{nta} is computed (Methods). V_{nta} and V_{com} were not statistically distinguishable for the grassland data: the distribution, across surrogates, of the squared coefficients of variation of the quantities $\sum_i x_i^{(m)}$ (recall that V_{nta} was defined to be the median of this distribution), was centred very close to the value of V_{com}, for both Hays and Konza (electronic supplementary material, figure S9A,C).

We now compare conclusions about Hays that would typically be drawn by using only the variance ratio approach, to conclusions drawn using both the variance and skewness ratios. Because $\phi_V < 1$ (figure 5), dynamics in Hays would classically be interpreted as compensatory, i.e. species

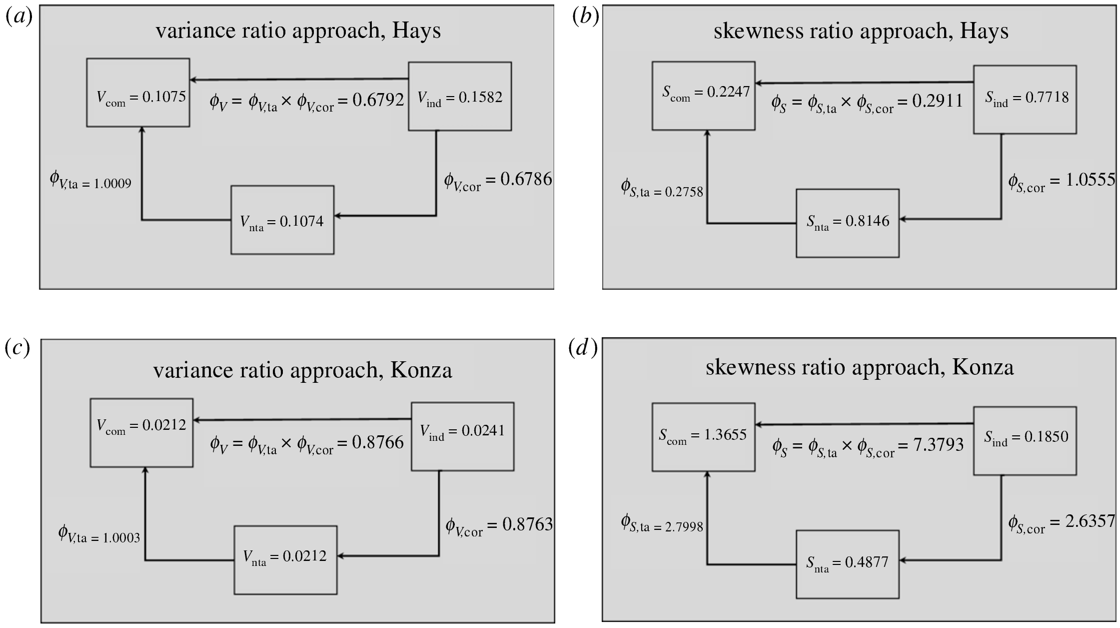

Figure 5. The quantities of figure 4 computed for the Hays (*a*,*b*) and Konza (*c*,*d*) datasets. See the electronic supplementary material, figure S9 for related results.

relationships buffer community variability. However, the combined approach yields more nuanced information. The skewness ratio, ϕ_S, was substantially less than 1, with $S_{com} = 0.2$ less than $S_{ind} = 0.8$ (figure 5), suggesting that the tendency for the total community abundance x_{tot} to explode to high values (right skewness) and surpass high disaster thresholds was mitigated by species relationships. This finding supports the variance-ratio-based conclusion that species relationships are generally stabilizing at Hays, but also extends the classic result by revealing the greater importance of community relationships for mitigating explosions of x_{tot} to high values, compared to their lesser role in mitigating crashes. Apparently, species relationships can mitigate or accentuate extreme values of x_{tot} differentially for high versus low extremes. See the electronic supplementary material, S7 for related results and discussion.

We next compare conclusions about Konza drawn by using only the variance ratio approach to conclusions drawn using both the variance and skewness ratios. Because ϕ_V was slightly less than 1 (figure 5), dynamics in Konza would classically be interpreted as slightly compensatory. However, using both the variance and skewness ratios together yields conclusions that differ substantially. The skewness ratio, ϕ_S, was much greater than 1, with $S_{com} = 1.4$ correspondingly much greater than $S_{ind} = 0.2$ (figure 5). So although species relationships slightly reduce the variance of x_{tot}, they also dramatically increase right skew of x_{tot}. This, in turn, leads to an overall greater propensity for explosions of x_{tot} to high values, and a correspondingly accentuated probability that x_{tot} will exceed high disaster thresholds. Crashes of x_{tot} to low values are reduced by species relationships. The variance ratio obscures the fact that species relationships modify extreme behaviour of x_{tot} differently in the two tails of its distribution. Although the variance ratio suggests that species dynamics at Konza are compensatory, reducing community variability, in fact this is true only for mitigating crashes of x_{tot} to low values; species relationships instead accentuate the tendency of x_{tot} to explode to high values. See the electronic supplementary material, S7 for related results and discussion.

Thus the skewness ratio newly reveals differences between Hays and Konza. For both communities, ϕ_V is slightly or moderately less than 1, suggesting compensatory, stabilizing species relationships. However, ϕ_S was less than 1 for Hays and was markedly greater than 1 for Konza owing partly to differing influences of interspecific tail associations on the distributions of x_{tot}. Hays can probably still be labelled as a community showing stabilizing, compensatory dynamics. By contrast, species relationships in Konza only reduce community variability in that they mitigate crashes of x_{tot}. Species relationships simultaneously accentuate explosions of x_{tot} to high values, a feature that should probably not be considered as stabilizing. For some applications, it may indeed be more important to avoid crashes than to avoid explosions of x_{tot} to high values, and for those applications it may still be reasonable to characterize species relationships at Konza as stabilizing; but this possibility depends on the applied context, and does not reduce the importance of our larger point that species relationships can be differently 'stabilizing' or 'destabilizing' (i.e. reducing versus increasing community variability) in the two tails of x_{tot}, a feature of empirical reality not captured by the variance ratio and other standard approaches.

These results highlight the previously unrecognized importance of tail associations for community dynamics. The influence of tail associations on the distribution of x_{tot} can be as great or greater than the influence of species correlations with symmetric tail association. For Konza, $\phi_{S,ta}$ and $\phi_{S,cor}$ were similar (figure 5), indicating that both tail associations and correlations were comparable in their effects on the skewness of x_{tot} (recall that $S_{com} = \phi_{S,ta}\phi_{S,cor}S_{ind}$). For Hays, however, $\phi_{S,cor}$ was close to 1 and $\phi_{S,ta}$ differed substantially from 1 (figure 5), indicating that essentially all of the influence of species relationships on $\mathrm{sk}(x_{tot})$ was owing to asymmetric tail associations between species. The influence of tail associations on the CV of x_{tot} was negligible, as indicated previously, but that reflects the inability of the CV to detect the impact of tail associations, rather than indicating that tail associations were unimportant.

Figure 6 makes more concrete some of the above conclusions through comparison of x_{tot} time series in Hays and

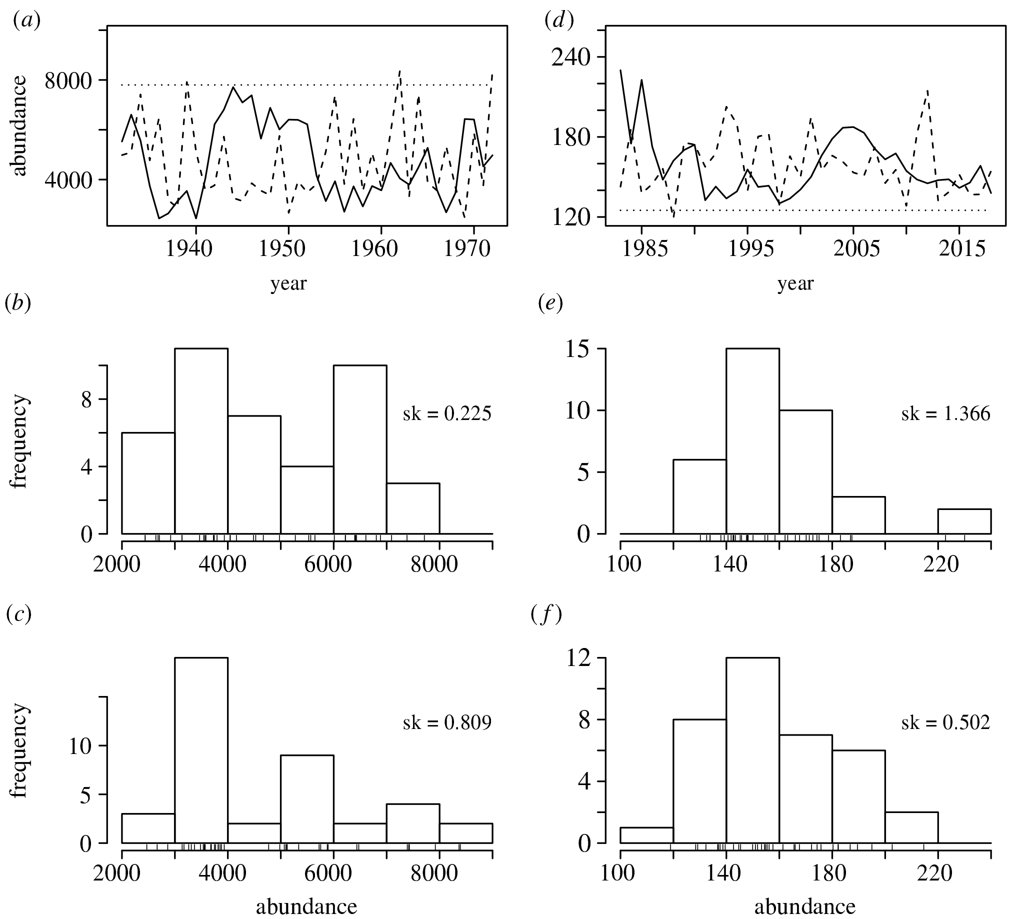

Figure 6. Example, using data, of the effect of tail association on dynamics of aggregate community abundance, $(x)_{\mathrm{tot}}$. Empirical $(x)_{\mathrm{tot}}$ time series for Hays (*a*) and Konza (*d*) are solid lines, with distributions of values across time shown in (*b*,*e*), respectively. Dashed lines on (*a*,*d*) show community abundance sums for 'typical' surrogate datasets produced via a randomization procedure that rendered tail associations symmetric, but left unchanged other statistical aspects of the data (Methods). Distributions across time for the dashed lines are in (*c*,*f*), respectively. Thus comparing dashed to solid lines, or comparing (*b*) to (*c*) (for Hays) or (*e*) to (*f*) (for Konza) reveals the effects of tail associations between species on community dynamics. Though the effect appears subtle, extreme high values are mitigated by tail associations in Hays (e.g. values on (*c*) exceed 8000, but values on (*b*) do not, and see also the dotted-line threshold on (*a*)); whereas extreme low values are mitigated in Konza (e.g. values on (*f*) go just below 120 but values on (*e*) do not, and see also the dotted-line threshold on (*d*)). If extreme values spell disaster in an applied scenario, these differences are important. This figure is an empirical parallel to aspects of figure 1. 'Abundance' was basal cover (cm^2) for Hays (*a*–*c*) and per cent cover for Konza (*d*–*f*). Vertical ticks between axes and histograms show the location of individual data points. Skewness, 'sk', is displayed on each histogram panel. Surrogates were selected for display which were median with respect to skewness of the total abundance, so that values of skewness on (*c*) and (*f*) approximate S_{nta}.

Konza to total abundance time series computed for randomized, surrogate datasets for which tail associations were rendered symmetric. The figure also helps illuminate future applied possibilities of the ideas of this study. For Hays, x_{tot} would have risen to more extreme high values were it not for the dominance of lower-tail associations between species in that system. For Konza, x_{tot} would have fallen to more extreme low values were it not for the dominance of upper-tail associations in that system. Because grazers seem more likely to be harmed by extreme low values of x_{tot} than by extreme high values, the differences between these systems illuminate some of the applied possibilities of our ideas. Whereas we emphasize that our data were from livestock-excluded plots, so applications to grazing are only suggestive, our overall results nevertheless show that details of species interactions can mitigate or accentuate extreme values of aggregate ecosystem functioning differently for high versus low extremes. Because human systems that depend on ecosystem functions (e.g. livestock grazing, though there are many examples) will react differently to high versus low extremes (and this will depend on the ecosystem function in question), the details of tail associations between species should be of clear applied interest in many circumstances. In essence, we showed that tail associations between species can have important effects on extreme behaviour of ecosystem functioning variables; so there is applied importance whenever extreme values would be a concern.

As described in the Introduction, the variance ratio ϕ_V compares V_{com} to the independence benchmark V_{ind} ($\phi_V = V_{\mathrm{com}}/V_{\mathrm{ind}}$), whereas the Loreau-de Mazancourt metric $\phi_{V,\mathrm{LdM}}$ compares V_{com} to the benchmark value V_{syn} that community CV^2 would take if species dynamics were perfectly linearly associated ($\phi_{V,\mathrm{LdM}} = V_{\mathrm{com}}/V_{\mathrm{syn}}$). This choice leads to substantial interpretive and other advantages of $\phi_{V,\mathrm{LdM}}$ over ϕ_V [10,23], and as a result $\phi_{V,\mathrm{LdM}}$ is increasingly widely used. An approach to community skewness using a paradigm similar to that of Loreau and de Mazancourt is available, and reveals the same main ideas we elaborated in this study using ϕ_S, but does not appear to have the same advantages over ϕ_S that $\phi_{V,\mathrm{LdM}}$ has over ϕ_V; nevertheless, such a metric may be worth considering in future work.

Because $\phi_{V,\mathrm{LdM}} = V_{\mathrm{com}}/V_{\mathrm{syn}}$, whereas $\phi_V = V_{\mathrm{com}}/V_{\mathrm{ind}}$, an alternative to $\phi_S = S_{\mathrm{com}}/S_{\mathrm{ind}}$ which follows the same paradigm may be $\phi_{S,\mathrm{LdM}} = S_{\mathrm{com}}/S_{\mathrm{syn}}$, where S_{syn} is the benchmark value that community skewness would take in the case of perfect synchrony between species dynamics. Whereas V_{syn} was community CV^2 computed under an assumption of perfect linear associations between the dynamics of different species [23], under our framework it may make more sense to consider S_{syn} to be community skewness under an assumption of species comonotonicity. Perfect linear associations between species cannot generally be realized without altering the marginal distributions of the species, and such alteration seems, to us, to be inconsistent with using a benchmark scenario in which species relationships are altered but other aspects of dynamics are kept constant. An alternative approach using $\phi_{S,\mathrm{LdM}} = S_{\mathrm{com}}/S_{\mathrm{syn}}$ would reveal essentially the same main ideas that we have revealed with the current approach of this study, because S_{nta} and V_{nta} would probably have to be defined in the same way. The main ideas of this study centre on the importance of tail associations for the skewness of x_{tot}, and therefore for extreme values of x_{tot}. Those ideas can be explored regardless of whether independence or perfect synchrony is used as a benchmark for comparison. One of the advantages of $\phi_{V,\mathrm{LdM}}$ over ϕ_V is that $\phi_{V,\mathrm{LdM}}$ always takes values between 0 and 1, facilitating comparisons across communities. But the same is not necessarily true for $\phi_{S,\mathrm{LdM}}$, as a careful examination of figures 2 and the electronic supplementary material, S2 will reveal (sk(x_{tot}) for scenario 2 divided by the same quantity for scenario C is much greater than 1, whereas computing the analogous quotient for scenario 3 gives a quantity much less than −1). Another advantage of $\phi_{V,\mathrm{LdM}}$ over ϕ_V is that $\sqrt{V_{\mathrm{syn}}}$ equals a weighted average of the CVs of individual species, and hence can be interpreted as population-level variability. Thus $\phi_{V,\mathrm{LdM}}$ can be elegantly interpreted as a scaling factor converting population- to community-level variability, whereas an analogous interpretation for ϕ_V is not available. However, such an interpretive advantage does not appear to extend to $\phi_{S,\mathrm{LdM}}$. Decades passed between the formulations of the original variance ratio and the Loreau-de Mazancourt and other improvements. Our goal with this study was to introduce the main ideas of tail association and its importance for community stability. We welcome the possibility that future research will reveal a metric that captures these main ideas, but with additional advantages.

Our work may also provide lessons for other fields. More detailed statistical descriptions of synchrony have become increasingly useful in recent years. Tail association metrics and related 'copula' statistics, though apparently seldom used so far in most biological research areas, may turn out to provide fruitful additional ways to make synchrony statistics more detailed and more useful across many fields. In many fields, techniques for quantifying the synchrony of signals have historically been crude, but have recently become more detailed, thereby enhancing progress. For instance, one paper in this special issue develops newly detailed frequency-specific descriptions of a kind of synchrony, and then uses them to make new inferences about musical and rhythmic behaviour in humans [37]. The examples of [38] on the interactive vocalizations of primates are also enhanced by a frequency-specific perspective: fig. 1C of [38] shows coordination of not only the timings but also the pitches of the calls of a pair of indris, a feature which would not be revealed without the additional statistical detail provided by the frequency-specific techniques of that paper (spectrograms). Sheppard, Zhao and colleagues [26,39,40] developed detailed frequency-specific measures of synchrony in population ecology which provided major inferential benefits. One common feature of all these studies is that progress was achieved through use of frequency-specific statistical descriptors of synchrony that were more detailed than earlier correlation and related basic tools. But spectral tools are not the only statistical approach by which additional information can be extracted from synchronous or asynchronous pairs of signals. Mathematically, the nature of the relationship between two quantities is an infinite-dimensional object, even when neither quantity exhibits spectral structure of any kind [31]. Most synchrony measures extract only a small portion of the available information. But the field of copula statistics [41] provides well-developed and general tools for assessing the complete information content. Tail association tools such as used in this study assess one aspect of copula structure. The potential of tail association and copula tools to reveal previously unsuspected useful information in patterns of synchrony across multiple fields of biology may be great, as demonstrated by our successes in this study. For instance, [42] indicates, in this special issue, that many childhood neurodevelopmental disorders are associated with deficits in rhythm, timing and synchrony. One wonders if diagnosis of disorders may be improved by applying copula statistics to data on synchrony tasks attempted by patients. Whereas we are not child psychologists and cannot critique work in that field or in many of the other disparate fields represented in this special issue, the power of copula statistics and the general scarcity of prior application of such tools to synchrony compel us to recommend these approaches to synchrony researchers in all fields.

5. Conclusion

It has been known for decades that species relationships can accentuate or mitigate the temporal variability of an aggregate community property such as the total abundance of all species, through synchronous or compensatory dynamics of species. However, our results show that consequences of species relationships for community stability are more complex than previously recognized. Synchronous relationships between the dynamics of different species are not necessarily well characterized by standard, commonly used covariance measures: species can exhibit contrasting patterns of tail association which are not detected by standard approaches. Moreover, the temporal stability of total abundance is also not necessarily well characterized by commonly used measures based on the coefficient of variation. Total abundance can exhibit patterns of skewness which correspond to increased probabilities of crashes, or explosions to high values, that are readily interpretable as instability. Thus the original variance ratio approach, whereby patterns of species' synchronous or compensatory dynamics are assessed using covariance tools and then related to community variability measured using a coefficient of variation, is probably inadequate, by itself. Species relationships can actually mitigate or accentuate community variability differentially for community crashes or explosions to high values, a fact which was apparently not previously recognized. In other words, species relationships can mitigate explosions of total abundance to high values, while simultaneously exacerbating the tendency to crash to low values; or vice versa. Our new skewness ratio approach complements the variance ratio to reveal a more multidimensional nature of species relationships and their effects on community variability. Overall, our work has begun a new and more flexible approach to synchrony, compensatory dynamics and their influence on community stability.

Data accessibility. All analyses were done in R. Complete computer code for this project is at https://github.com/sghosh89/Copula_spaceavg.

Authors' contributions. S.G. and D.C.R. designed the study, developed the theory and statistical tools, wrote the code, and analysed the data. All authors contributed to writing and revising. All authors gave final approval for publication and agree to be held accountable for the work performed therein.

Competing interests. We declare we have no competing interests.

Funding. The authors were partly supported by US National Science Foundation grant nos. 1714195 and 2023474, the James S. McDonnell Foundation, and the California Department of Fish and Wildlife Delta Science Program.

Acknowledgements. The authors thank Jim Bever, Bryan Foster, Lauren Hallett and Michel Loreau for helpful discussions and constructive criticism.

References

1. MacArthur R. 1955 Fluctuations of animal populations and a measure of community stability. *Ecology* **36**, 533–536. (doi:10.2307/1929601)
2. Pimm SL. 1984 The complexity and stability of ecosystems. *Nature* **307**, 321. (doi:10.1038/307321a0)
3. Micheli F *et al.* 1999 The dual nature of community variability. *Oikos* **85**, 161–169. (doi:10.2307/3546802)
4. Gonzalez A, Loreau M. 2009 The causes and consequences of compensatory dynamics in ecological communities. *Annu. Rev. Ecol. Evol. Syst.* **40**, 393–414. (doi:10.1146/annurev.ecolsys.39.110707.173349)
5. Peterson C. 1975 Stability of species and of community for the benthos of two lagoons. *Ecology* **56**, 958–965. (doi:10.2307/1936306)
6. Frost TM, Carpenter SR, Ives AR, Kratz TK. 1995 Species compensation and complementarity in ecosystem function. In *Linking species & ecosystems* (eds CG Jones, JH Lawton), pp. 224–239. New York, NY: Chapman & Hall.
7. Bai Y, Han X, Wu J, Chen Z, Li L. 2004 Ecosystem stability and compensatory effects in the Inner Mongolia grassland. *Nature* **431**, 181–184. (doi:10.1038/nature02850)
8. Hallett LM *et al.* 2014 Biotic mechanisms of community stability shift along a precipitation gradient. *Ecology* **95**, 1693–1700. (doi:10.1890/13-0895.1)
9. Keitt TH. 2008 Coherent ecological dynamics induced by large-scale disturbance. *Nature* **454**, 331–334. (doi:10.1038/nature06935)
10. Loreau M, de Mazancourt C. 2008 Species synchrony and its drivers: neutral and nonneutral community dynamics in fluctuating environments. *Am. Nat.* **172**, E48–E66. (doi:10.1086/589746)
11. Ma Z, Liu H, Mi Z, Zhang Z, Wang Y, Xu W, Jiang L, He JS. 2017 Climate warming reduces the temporal stability of plant community biomass production. *Nat. Commun.* **8**, 1–7. (doi:10.1038/s41467-016-0009-6)
12. Sheppard L, Mechtley B, Walter J, Reuman D. 2020 Self-organizing cicada choruses respond to the local sound and light environment. *Ecol. Evol.* **10**, 4471–4482. (doi:10.1002/ece3.6213)
13. Stange EE, Ayres MP, Bess JA. 2011 Concordant population dynamics of Lepidoptera herbivores in a forest ecosystem. *Ecography* **34**, 772–779. (doi:10.1111/j.1600-0587.2010.06940.x)
14. Houlahan J *et al.* 2007 Compensatory dynamics are rare in natural ecological communities. *Proc. Natl Acad. Sci. USA* **104**, 3273–3277. (doi:10.1073/pnas.0603798104)
15. Ives AR. 1995 Predicting the response of populations to environmental change. *Ecology* **76**, 926–941. (doi:10.2307/1939357)
16. Fischer JM, Frost TM, Ives AR. 2001 Compensatory dynamics in zooplankton community responses to acidification: measurement and mechanisms. *Ecol. Appl.* **11**, 1060–1072. (doi:10.1890/1051-0761(2001)011[1060:CDIZCR]2.0.CO;2)
17. Loreau M, de Mazancourt C. 2013 Biodiversity and ecosystem stability: a synthesis of underlying mechanisms. *Ecol. Lett.* **16**, 106–115. (doi:10.1111/ele.12073)
18. Mori AS, Furukawa T, Sasaki T. 2013 Response diversity determines the resilience of ecosystems to environmental change. *Biol. Rev.* **88**, 349–364. (doi:10.1111/brv.12004)
19. Frost TM, Carpenter SR, Kratz TK. 1992 Choosing ecological indicators: effects of taxonomic aggregation on sensitivity to stress and natural variability. In *Ecological indicators* (eds DH McKenzie, DE Hyatt, VJ McDonald), pp. 215–27. New York, NY: Elsevier Applied Science.
20. Ives A, Dennis B, Cottingham K, Carpenter S. 2003 Estimating community stability and ecological interactions from time-series data. *Ecol. Monogr.* **73**, 301–330. (doi:10.1890/0012-9615(2003)073[0301:ECSAEI]2.0.CO;2)
21. Schluter D. 1984 A variance test for detecting species associations, with some example applications. *Ecology* **65**, 998–1005. (doi:10.2307/1938071)
22. Gross K, Cardinale BJ, Fox JW, Gonzalez A, Loreau M, Wayne Polley H, Reich PB, van Ruijven J. 2014 Species richness and the temporal stability of biomass production: a new analysis of recent biodiversity experiments. *Am. Nat.* **183**, 1–12. (doi:10.1086/673915)
23. Thibaut L, Connolly S. 2013 Understanding diversity–stability relationships: towards a unified model of portfolio effects. *Ecol. Lett.* **16**, 140–50. (doi:10.1111/ele.12019)
24. Vasseur DA, Gaedke U. 2007 Spectral analysis unmasks synchronous and compensatory dynamics in plankton communities. *Ecology* **88**, 2058–2071. (doi:10.1890/06-1899.1)
25. Brown B, Downing A, Leibold M. 2016 Compensatory dynamics stabilize aggregate community properties in response to multiple types of perturbations. *Ecology* **97**, 2021–2033. (doi:10.1890/15-1951.1)
26. Zhao L *et al.* 2020 A new variance ratio metric to detect the timescale of compensatory dynamics. *Ecosphere* **11**, e03114. (doi:10.1002/ecs2.3114)
27. Li D. 2000 On default correlation: a copula function approach. Working paper number 99-07. 44 Wall Street, New York, NY: The RiskMetrics Group.
28. Salmon F. 2009 Recipe for disaster: the formula that killed Wall Street. Wired Magazine, 23 February 2009. See https://www.wired.com/2009/02/wp-quant/.
29. Cheung K, Vanduffel S. 2013 Bounds for sums of random variables when the marginal distributions and the variance of the sum are given. *Scand. Actuar. J.* **2013**, 103–118. (doi:10.1080/03461238.2011.558186)
30. Ghosh S, Sheppard L, Reuman D. 2020 Tail associations in ecological variables and their impact on extinction risk. *Ecosphere* **11**, e03132. (doi:10.1002/ecs2.3132)
31. Ghosh S, Sheppard LW, Holder MT, Loecke TD, Reid PC, Bever JD, Reuman DC. 2020 Copulas and their potential for ecology. *Adv. Ecol. Res.* **62**, 409–468. (doi:1101/650838)
32. Ghosh S, Sheppard L, Reid P, Reuman D. 2020 A new approach to interspecific synchrony in population ecology using tail association. *Ecol. Evol.* **10**, 12 764–12 776. (doi:10.1002/ece3.6732)
33. Cheung K. 2008 Characterization of comonotonicity using convex order. *Insur. Math. Econ.* **43**, 403–406. (doi:10.1016/j.insmatheco.2008.08.002)
34. Adler PB, Tyburczy WR, Lauenroth WK. 2007 Long-term mapped quadrats from Kansas prairie: demographic information for herbaceous plants. Ecological archives E088-161. *Ecology* **88**,

2673–2673. (doi:10.1890/0012-9658(2007)88[2673:LMQFKP]2.0.CO;2)

35. Hartnett D, Collins S. 2019 PVC02 Plant species composition on selected watersheds at Konza Prairie. See https://doi.org/10.6073/pasta/f7b9965ec6efa5247a97059bcdaf4a1b. Environmental Data Initiative.
36. Genest C, Favre A-C. 2007 Everything you always wanted to know about copula modeling but were afraid to ask. *J. Hydrol. Eng.* **12**, 347–368. (doi:10.1061/(ASCE)1084-0699(2007)12:4(347))
37. Dotov D, Trainor L. Cross-frequency coupling explains the preference for simple ratios in rhythmic behaviour and the relative stability across non-synchronous patterns. *Phil. Trans. R. Soc. B* **376**, 20200333. (doi:10.1098/rstb.2020.0333)
38. de Reus K, Soma M, Anichini M, Gamba M, de Heer Kloots M, Lense M, Bruno JH, Laurel T, Ravignani A. 2021 Rhythm in dyadic interactions. *Phil. Trans. R. Soc. B* **376**, 20200337. (doi:10.1098/rstb.2020.0337)
39. Sheppard LW, Bell JR, Harrington R, Reuman DC. 2016 Changes in large-scale climate alter spatial synchrony of aphid pests. *Nat. Clim. Change* **6**, 610. (doi:10.1038/nclimate2881)
40. Sheppard L, Defriez E, Reid P, Reuman D. 2019 Synchrony is more than its top-down and climatic parts: interacting Moran effects on phytoplankton in British seas. *PLoS Comput. Biol.* **15**, e1006744. (doi:10.1371/journal.pcbi.1006744)
41. Nelsen RB 2006 *An introduction to copulas*, 2nd edn. New York, NY: Springer.
42. Lense MD, Ledanyi E, Rabinowitz T-C, Trainor L, Gordon R. 2021 Rhythm and timing as vulnerabilities in neurodevelopmental disorders. *Phil. Trans. R. Soc. B* **376**, 20200327. (doi:10.1098/rstb.2020.0327)