Reconceptualizing beta diversity: a hypervolume geometric approach

Chuliang Song[1,2*,†], Muyang Lu[3*], Joseph R. Bennett[4], Benjamin Gilbert[2], Marie-Josée Fortin[2], Andrew Gonzalez[1]

[1] Department of Biology, Quebec Centre for Biodiversity Science, McGill University, Montreal, Canada

[2] Department of Ecology and Evolutionary Biology, University of Toronto, Toronto, Canada

[3] Department of Ecology and Evolutionary Biology, Yale University, New Haven, CT, USA

[4] Department of Biology, Carleton University, Ottawa, Canada

∗ These authors contributed equally

† Present address: Department of Ecology and Evolutionary Biology, Princeton University, Princeton, NJ, USA

Running title: Geometry of beta diversity

Number of words: abstract 189, main text 6691

Number of references: 90 references

Number of figures: 10 colored figures

Number of tables: 1 table

Keywords: beta diversity, community composition, hypervolume,

Author contributions: C.S. conceived of the idea of geometric beta diversity. C.S., M.L., M.-J.F. and A.G. designed and realized the study. C.S. and M.L. performed the theoretical analysis. C.S. and M.L. wrote a first draft of the manuscript. All authors contributed to substantial revisions. J.R.B. and B.G. compiled and provided data.

Contents

1 Introduction **4**

2 Geometry of beta diversity **6**

2.1 Illustration of the basic idea 7

2.2 Generalization to arbitrary metacommunity 8

3 A unified framework of beta diversity **10**

3.1 Duplications in presence/absence data as weighted embedding 11

3.2 Temporal changes of beta diversity as hypervolume overlap 12

3.3 Community/Species-specific contribution as hypervolume change 14

3.4 Species similarity and functional complementarity as transformed embedding . . . 16

3.5 Nestedness-turnover decomposition as filling-finding facets 18

4 Empirical applications **21**

4.1 Efficient estimation of beta diversity 21

4.2 Latitudinal pattern of beta diversity 21

4.3 How sampling efforts affect beta diversity 22

5 Discussion **24**

6 Conclusion **27**

Abstract

Beta diversity—the variation among community compositions in a region—is a fundamental indicator of biodiversity. Despite a diverse set of measures to quantify beta diversity, most measures have posited that beta diversity is maximized when each community has one distinct species. However, this postulate has ignored the importance of non-additivity of ecological systems (i.e., a community with two species is ecologically different from two communities with one species). Here, to account for this, we provide a geometric approach to measure beta diversity as the hypervolume of the geometric embedding of a metacommunity. We show that the hypervolume measure is closely linked to and naturally extends previous information- and variation-based measures. In addition, our hypervolume approach provides a unified geometric framework for widely adopted extensions on the basic measure of beta diversity: the contribution of duplications in presence/absence data, temporal changes, turnover-nestedness decomposition, species similarity and functional complementarity, and community/species-specific contributions. We apply our new geometric measures to empirical data and address two long-standing questions on beta diversity (latitudinal pattern and sampling efforts) and present novel ecological insights. In sum, our geometric approach reconceptualizes beta diversity, synthesizes previous measures and is immediately applicable to existing data.

1 Introduction

Beta diversity is one of the most important indicators of biodiversity (Anderson *et al.*, 2011; Mittelbach & McGill, 2019). In essence, beta diversity aims to measure the diversity of between-community components, or the number of effective communities. It serves as a bridge connecting ecological phenomena from local to regional scales. Unfortunately, it remains one of the most debated concepts in biodiversity research. Since the concept was conceived in the mid-20th Century (Whittaker, 1960, 1972), researchers have come up with a long list of measures (reviewed in Anderson *et al.* 2011 and Scheiner 2019). Some recent notable measures include Hill numbers (Jost, 2007; Ohlmann *et al.*, 2019), β-deviation (Kraft *et al.*, 2011; Xing & He, 2021a), turnover-nestedness decomposition (Baselga, 2012; Legendre, 2014), and variance of community composition matrix (Legendre & De Cáceres, 2013). Importantly, although these measures focus on different aspects of beta diversity, most of them obey the same set of mathematical axioms (Legendre & De Cáceres, 2013). However, some of these axioms may not align with the key ecological intuition on what beta diversity should be.

The key discrepancy rests on the condition by which beta diversity should be maximized. To explain the problem in a nutshell, let us consider two metacommunities (labeled as I and II; Figure 1), both with 3 local communities (labeled as 1-3) and 2 species (labeled as A and B). For simplicity, we use the Whittaker's multiplicative measure of beta diversity ($\beta = \gamma/\alpha$) to represent the traditional measures. In metacommunity I, community 1 and community 2 both only have species A, while community 3 only has species B. Metacommunity I's gamma diversity $\gamma = 2$, and its alpha diversity $\alpha = 1$, thus its beta diversity $\beta = \gamma/\alpha = 2$. Then, in metacommunity II, community 1 only has species A, community 2 only has species B, and community 3 has both species A and B. Metacommunity II's gamma diversity $\gamma = 2$, and its alpha diversity $\alpha = 4/3$, thus $\beta = \gamma/\alpha = 1.5$. Figure 1A illustrates these two metacommunities. Thus, the traditional measures argue that metacommunity I has a larger beta diversity than metacommunity II. Note that this is not a special property of the Whittaker's multiplicative measure but satisfied by almost all measures (reviewed in Legendre & De Cáceres 2013). However, metacommunity I only has two distinctive community compositions ($\{A\}$, $\{B\}$), while metacommunity II has three distinctive compositions ($\{A\}$, $\{B\}$, and $\{A, B\}$). Extensive theory and empirical validation have shown that the community with composition $\{A, B\}$ should behave differently, both dynamically (Levine *et al.*, 2017; Angulo *et al.*, 2021) and functionally (Maron *et al.*, 2018; van der Plas, 2019), from the communities with only $\{A\}$ or $\{B\}$; in other words, ecological communities are non-additive. If we accept this assumption of non-additivity, then maximum beta diversity should be achieved when all communities have different compositions, where the metacommunity has the highest *diversity*

of ecosystem dynamics. In sum, opposite to traditional measures, we should assign a higher beta diversity for metacommunity II than metacommunity I.

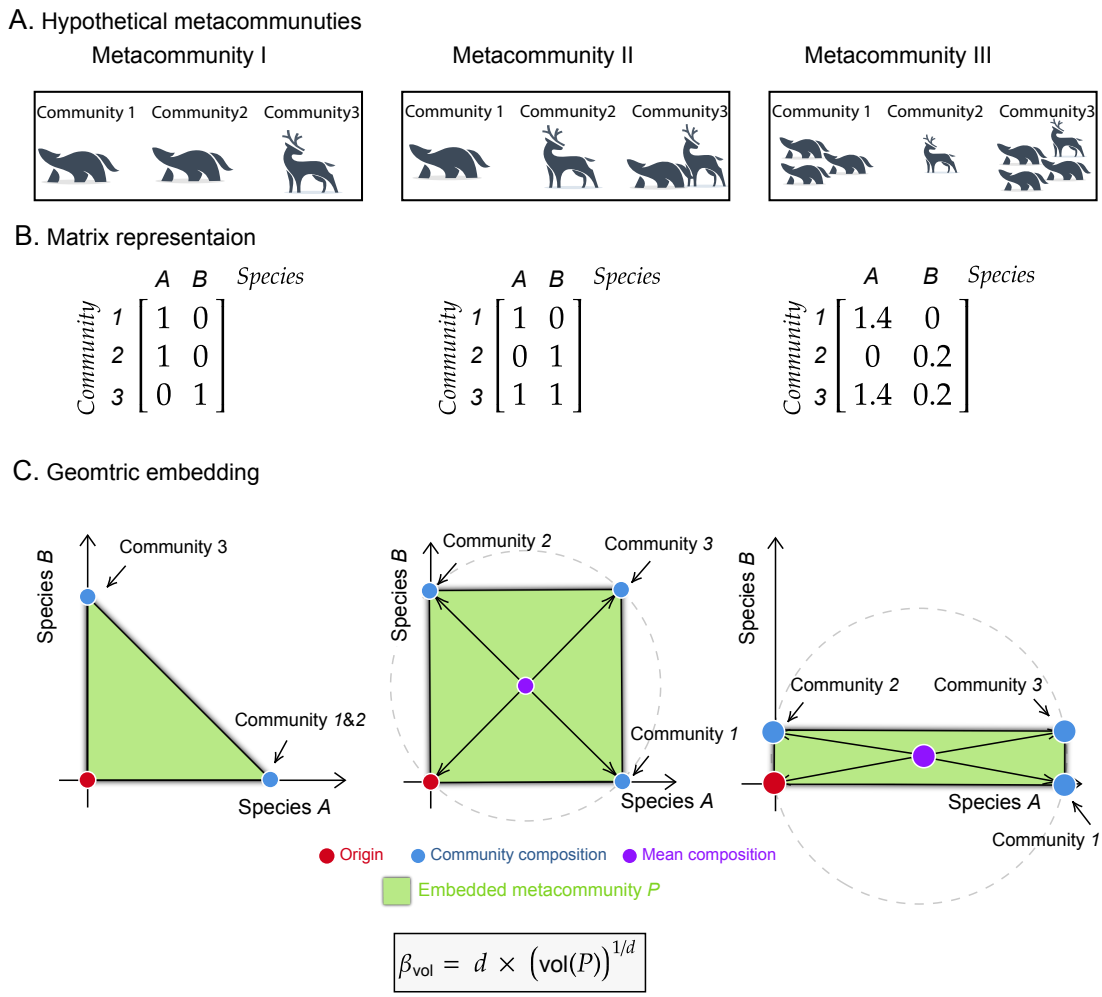

Figure 1: Illustration of the geometric approach to beta diversity. Panel (A) shows three hypothetical metacommunities (labeled as I-III) with 2 species (labeled as A and B) and 3 local communities (labeled as 1-3). The three communities of metacommunity I consist of species A only, species A only, and species B only, respectively. The three communities of metacommunity II consist of species A only, species B only, and species A, B together, respectively. The three communities of metacommunity III have the same species composition as metacommunity II, except that species A is of 7 times more abundant than species B. Panel (B) shows the equivalent matrix representations of the metacommunities in Panel (A). This matrix is known as community composition matrix, where the rows represent communities, the columns represent species, and elements represent species presence/absence. More generally, the elements can be any measure of species importance, such as abundance or biomass. Panel (C) provides a geometric embedding of the metacommunity. Here, as we have a smaller number of species than the number of communities, the number of species determines the dimension and axis of the embedded space (2-dimensional Euclidean space), while the communities determined the embedded points (blue points). Note that the origin (red point) is automatically embedded because adding an empty community should not affect beta diversity. The metacommunity is now realized as the spanned geometric object (green area; denoted as P) by all the embedded point and the origin. The geometric beta diversity is defined as the normalized hypervolume of the geometric object: $\beta_{\text{vol}} = d \times (\text{vol}(P))^{1/d}$, where d is the embedded dimension (2 here).

In addition to the non-additivity of ecological communities, the discrepancy is also manifest from the perspective of spatial associations among species. Despite the continued debates on the use of co-occurrence patterns, it is not doubted that co-occurrence patterns are partly driven by biological processes (Blanchet *et al.*, 2020). However, traditional beta diversity metrics, such as Whittaker's multiplicative metric (Whittaker, 1972) and variance-based metric (Legendre & De Cáceres, 2013), largely ignore species' association patterns. To see this, let us consider metacommunities II and III in Figure 1A with the variance-based metric. They have identical community compositions but different relative species abundances: species A and B are equally abundant in metacommunity II, while species A is 7 times abundant than species B in metacommunity III. Briefly, the variance-based metric measures how each community composition differs from the mean composition. As can be seen in Figure 1C, metacommunities II and III would have identical beta diversity as measured by the variance-based metric. But this has ignored how these differences in species compositions are interlinked to one another: the differences are orthogonal in metacommunity II, while they are not in metacommunity III. In other words, traditional measures mostly consider only the first order information (diagonal of the species covariance matrix) but not second order information (off-diagonal of the species covariance matrix).

To resolve these discrepancies on non-additivity and spatial association, we introduce a new measure of beta diversity using an intuitive and visual geometric approach. The key idea of this geometric approach is to view the metacommunity as a geometric object occupying hyperspace, and then quantify its beta diversity as the hypervolume of the geometric object. Firstly, we illustrate the key ideas with simple examples, and provide a generalization to metacommunities with arbitrary structure. Then, armed with this geometric perspective, we provide a unified treatment of common variants beyond basic beta diversity: duplications in presence/absence data, temporal changes, community/species-specific contribution, turnover-nestedness decomposition, and accounting for species similarity and functional complementarity. In contrast, traditional approaches require different formalisms to deal with these variants. We then show this geometric approach is linked to and naturally extends classic measures of beta diversity based on generalized covariance and information theory. Lastly, we apply our hypervolume measure of beta diversity to empirical datasets, including the trend of beta diversity along longitude and the sampling efforts.

2 Geometry of beta diversity

Most definitions of beta diversity stem from algebraic manipulations of metacommunity properties (Anderson *et al.*, 2011). Here, we provide an alternative geometric approach. This approach is

grounded on the idea of embedding an arbitrary metacommunity as a hyper-dimensional geometric object. We will show that this geometric shape of metacommunity provides a unifying bridge to various definitions of beta diversity.

2.1 Illustration of the basic idea

To illustrate the basic idea, let us consider again the hypothetical examples of metacommunities in Figure 1. Recall that in metacommunity II, community 1 only has species A, community 2 only has species B, and community 3 has both species A and B (middle panel in Fig 1A). We can represent the metacommunity in a matrix form (Figure 1B):

$$\begin{bmatrix} 1 & 0 \\ 0 & 1 \\ 1 & 1 \end{bmatrix}, \tag{1}$$

where the columns denote species (the first column for species A and the second column for species B), the rows denote the communities (the i-th row for the i-th community), and the elements denote whether the given species is present in the given community (1 for presence and 0 for absence). We call this matrix form the *metacommunity matrix*. Note this matrix form is also known as community matrix in the literature (Legendre & De Cáceres, 2013).

The crux of our new definition of beta diversity is to interpret this matrix as points in a hyper-dimensional space. In this example, the space is 2-dimensional (each species as an axis) and we have three points (rows in the matrix: $(1, 0)$, $(0, 1)$, and $(1, 1)$). The middle panel of Figure 1C illustrates the geometric embedding of the matrix. Beta diversity is related to the volume spanned by these points together with the origin. The ecological rationale to add the origin is known as double-zero asymmetry (appendix S3 of Legendre & De Cáceres 2013): beta diversity should not change when we "add" a ghost species that does not exist in all the communities (which is the origin in the space), because such a ghost species is not interpretable (Whittaker, 1972). Thus, the origin must be included for ecological consistency of beta diversity.

With this geometric embedding (Figure 1C), we can see that metacommunity II, which has three distinct community compositions, has a volume of 1. In comparison, metacommunity I, which has two distinct community compositions, has a volume of 0.5. Thus, this hypervolume approach naturally resolves the discrepancy regarding when beta diversity is maximized: more distinct compositions correspond to more unique points in the hyper-dimension, which leads to greater

hypervolumes. Note that the hypervolume would be minimized (= 0) if all communities have identical compositions, which align with the intuition of beta diversity with non-additivity.

Following Jost (2007), to make the hypervolume measure more ecologically meaningful, we define the beta diversity in this example as the rescaled volume:

$$\beta_{\mathsf{vol}} \coloneqq 2 \times (\mathrm{vol}(\mathcal{P}))^{1/2}, \tag{2}$$

representing an effective number of communities with 2 species. With this definition, the metacommunity I has a beta diversity of 1.4, while the metacommunity II has a beta diversity of 2 (the highest possible beta diversity).

2.2 Generalization to arbitrary metacommunity

We can generalize the simple cases above to complex metacommunities. For a general metacommunity with N local communities and γ species, we can represent it using a general metacommunity matrix:

$$\begin{bmatrix} z_{11} & \dots & z_{1\gamma} \\ \dots & \diagdown & \dots \\ z_{N1} & \dots & z_{N\gamma} \end{bmatrix}, \tag{3}$$

where columns denote species, rows denote community, and the elements (z_{ij}) denote the ecological measure of species importance. Note that z_{ij} can be any arbitrary measure (common measures include presence, abundance, and biomass). A caveat, though, is that z_{ij} need to be appropriately scaled to make it fully comparable across metacommunities (Legendre & De Cáceres, 2013).

We need to identify the constraint on beta diversity: the gamma diversity (γ), or the number of communities (N). Whichever is smaller is the constraint. For example, if there are more species (γ) than communities (N), then the number of communities (N) constrains beta diversity (i.e., maximally N effective communities). And vice versa. We take the constraint as the dimension of the embedded space, and the other quantity as the number of embedded points. The identification of the constraint assures that beta diversity is well-defined for all metacommunities.

Formally, the expanded convex hull $\mathcal{P}$ of the geometrically embedded points in the $d = \min(\gamma, N)$ -dimensional space is

$$\mathcal{P} \coloneqq \{ \sum_{i=1}^{\max(\gamma,N)} \lambda_i \mathbf{x}_i \mid \lambda_0, \lambda_1, \dots, \lambda_{\max(\gamma,N)} \geq 0, \sum_{i=0}^{\max(\gamma,N)} \lambda_i = 1 \}, \tag{4}$$

where λ_0 corresponds to the origin, and $\mathbf{x}_i$ corresponds to either the i-th column or row of the metacommunity matrix (depending on which is the constraint).

Following the definition above, our measure of beta diversity β_{vol} (the underscript highlights the use of hypervolume) is defined as the rescaled hypervolume of the convex hull $\mathcal{P}$:

$$\beta_{\mathsf{vol}} \coloneqq d \times (\mathrm{vol}(\mathcal{P}))^{1/d}, \tag{5}$$

where $d = \min(S, N)$ is the constraint and $(\mathrm{vol}(\mathcal{P}))^{1/d}$ is the normalized hypervolume. β_{vol} is interpreted as the number of effective communities, which ranges from 0 to d.

This rescaling of raw hypervolume in Eqn. 5 is fundamental for its interpretation as beta diversity. A heuristic argument is that, with γ species, the hypervolume beta diversity should range from 0 (achieved with only 1 unique community composition) to γ (achieved with $(2^\gamma - 1)$ distinct community compositions). This range of beta diversity is based on the argument that the effective number is mostly ecologically intuitive (Jost, 2007). To get rid of the effects of the exponential increase of distinct community compositions required, we need to take the $\gamma-$th root of the hypervolume. Of course, further rescaling of Eqn. 5 is possible depending on different ecological rationals (e.g., beta diversity should range from 0 to 1).

To validate the heuristic argument behind the rescaling of the hypervolume (Eqn. 5), we compute all possible beta diversities for metacommunities with three species (i.e., $\gamma = 3$). The maximum beta diversity ($\beta_{\mathsf{vol}} = 3$) is achieved with $(2^\gamma - 1) = 7$ distinct community compositions, while the minimum beta diversity is achieved with only 1 distinct community composition. Figure 2 shows the rescaled hypervolume has a close to linear association to the number of unique community compositions in an metacommunity. Appendix C shows the linear scaling persists for higher gamma diversity.

In addition to the linear trend, values of beta diversity have notable variations within the same number of unique community compositions. Thus, even though the number of unique community compositions is the key determinant of beta diversity, how distinct the unique community compositions are is also another determinant. For example, a community with species composition $\{A, B\}$ is more distinct from a community with composition $\{C\}$ than a community with composition $\{A\}$. In other words, β_{vol} does not treat all unique composition equally; instead, it puts higher weights to more distinct community compositions.

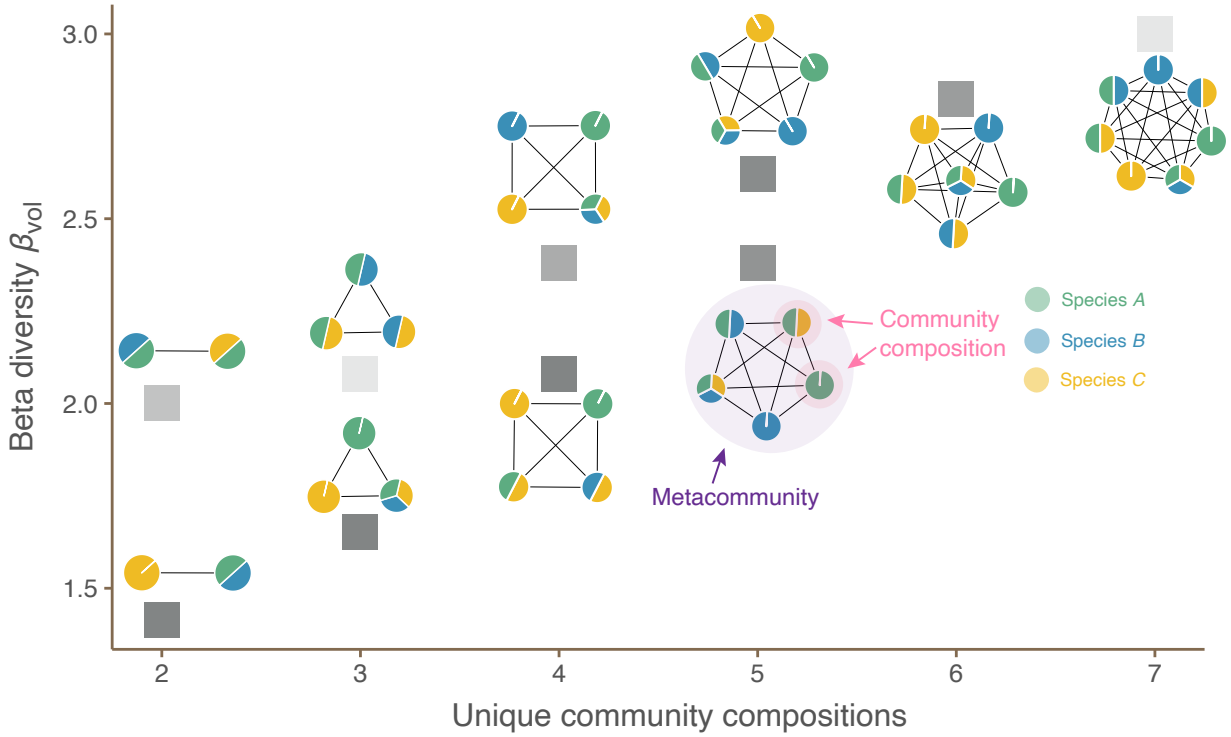

Figure 2: Possibilities of beta diversities for metacommunities with gamma diversity = 3. To explore the full possibilities, here we only consider metacommunities with information on species presence or absence. The horizontal axis shows the number of unique community compositions in the metacommunity, while the vertical axis shows the hypervolume beta diversity (β_{vol}) defined in Eqn. (5). For clarity of presentation, we only show metacommunities with nontrivial beta diversity (i.e., $\beta_{\text{vol}} > 0$). The transparency of the point denotes the number of distinct metacommunities that have identical beta diversity with the same number of unique community compositions, with more solid points indicating more metacommunities. For each point, we illustrate one example of metacommunity. In each example metacommunity, the communities are represented as nodes, and the colors of the nodes represent community compositions (green, blue, and yellow for species A, B, C, respectively). Beta diversity β_{vol} increases with the number of unique community compositions in a linear trend with notable variations. These variations are due to different levels of similarities in species compositions.

3 A unified framework of beta diversity

In the previous section, we have introduced a geometric approach to define beta diversity through a geometric embedding of a metacommunity (Eqn. 5). We have so far only focused on the most basic case of beta diversity. Many important extensions of beta diversity have been proposed through the study of beta diversity, such as temporal dimension (De Cáceres *et al.*, 2019) and accounting for species similarity (Leinster & Cobbold, 2012). Despite their importance, these extensions require different methodologies. With the flexibility empowered with geometry, here we provide a unified treatment to these extensions in beta diversity theory.

3.1 Duplications in presence/absence data as weighted embedding

Information on species presence or absence is often the only available data in empirical metacommunities. Mathematically, this means $z_i = 1$ or 0 in the metacommunity matrix (Eqn. 4). A common issue with these data is the duplication of identical community compositions. For example, metacommunity I in Figure 1A has two communities with only species A. However, the definition of β_{vol} (Eqn. 5) in the previous section does not take this into account because communities with duplicated compositions would be a single point.

To account for this, we provide a simple modification to account for duplicated community compositions through weighted embedding. We compress all communities with duplicated compositions into one community and then assign the frequency of identical communities as its weight. To illustrate the scheme, let us consider again metacommunity I in Figure 1A. Recall that it has two communities with species A and one community with species B. We can compress the three communities into two communities and then give double weights to the community with species A. This procedure is equivalent to transforming the original metacommunity matrix into a new one:

$$\begin{bmatrix} 1 & 0 \\ 1 & 0 \\ 0 & 1 \end{bmatrix} \rightarrow \begin{bmatrix} 2 \times \frac{2}{3} & 0 \\ 0 & 2 \times \frac{1}{3} \end{bmatrix}, \tag{6}$$

where 2 comes from the fact that there are 2 unique community compositions, and 1/3 and 2/3 correspond to the relative proportion of duplicated compositions. Figure 3 illustrates a more complicated example of metacommunity with 2 species and 6 communities.

Formally, suppose we have N local communities and S species in a metacommunity where the number of species is the constraint (i.e., $S < N$). The same modification can be applied when the number of communities is the constraint. The species composition of the i-th community is $\mathbf{x}_i \coloneqq \{z_{ij}\}$. Then suppose among the N communities, we have only m unique communities $\mathbf{y}_k$ $(k = 1, \ldots, m)$, where each unique community $\mathbf{y}_k$ appears n_k times. Then

$$\mathcal{P} \coloneqq \{\sum_{k=1}^{m} \lambda_k \frac{m \times n_k}{\sum_{l=1}^{m} n_l} \mathbf{y}_k \mid \lambda_0, \lambda_1, \ldots, \lambda_k \geq 0, \sum_{i=0}^{k} \lambda_i = 1\}, \tag{7}$$

where $\frac{m \times n_k}{\sum_{l=1}^{m} n_l}$ provides the weighted embedding of y_k. The weight would be 1 if all communities have distinct compositions (i.e., $n_k = 1, \forall k$). The computation of the beta diversity β_{vol} is identical to Eqn. (5). Note that it is straightforward to apply this modification on weighted embedding to other types of measures of species importance. Although in empirical data, the modification

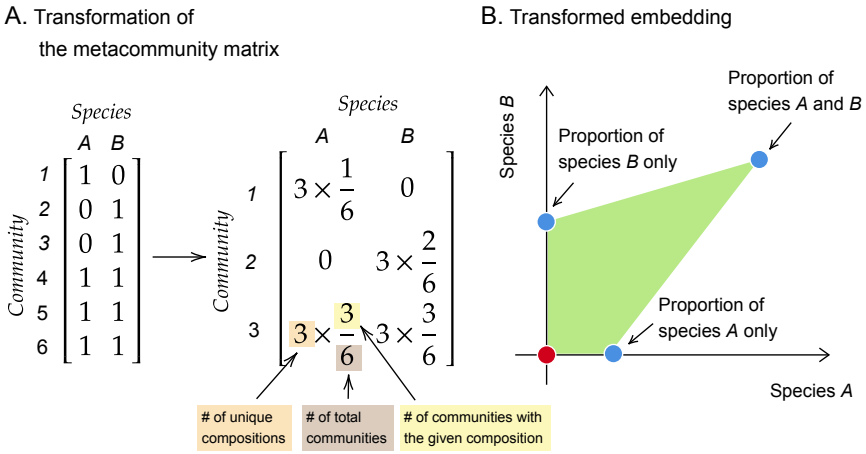

Figure 3: Weighted geometric embedding of metacommunity with presence/absence data. Panel (A) shows an example of metacommunity with 2 species and 6 communities. The left matrix is the original metacommunity matrix, while the right one is the transformed metacommunity matrix. This weighted transformation is given by (# of unique compositions) $\times \frac{\text{\# of communities with the given composition}}{\text{\# of total communities}}$. This embedding scheme ensures that metacommunities without duplications would be identical after the transformation. Panel (B) shows the transformed embedding of the metacommunity.

is unlikely to be needed. For example, it is unlikely that two communities would have identical abundances for all their constituent species.

As an application, we can ask the following question: for a metacommunity with 2 species (labeled as A and B), what is the proportion of communities with compositions $\{A\}$, $\{B\}$ and $\{A, B\}$ that maximize the beta diversity? We prove that the beta diversity is maximized when 1/4 communities have $\{A\}$, 1/4 communities have $\{B\}$, and the other 1/2 communities have $\{A, B\}$ (Appendix A).

3.2 Temporal changes of beta diversity as hypervolume overlap

Beta diversity *per se* is a measure on the spatial scale. To fully understand biodiversity changes, we need to study how beta diversity changes over time and over different temporal scales (Gonzalez *et al.*, 2020). One option is to simply compare the value of beta diversity at two times. Despite its simplicity, this option would ignore the information of whether and how community compositions are synchronized across the whole metacommunity. To capture this information, De Cáceres *et al.* (2019) has proposed a method based on trajectory distances. Here, we measure the temporal change using the overlap between two (oriented) geometric embeddings of metacommunities.

To illustrate the idea, let us consider a metacommunity with 2 communities and 2 species. At time (t), community 1 has only species B while community 2 has both species A and B (Figure 4A). Then at time ($t + 1$), community 1 still has species B while community 2 now only has species A (Figure 4B). To compute the hypervolume overlap, we need to assign the orientation

of the geometric embedding. This orientation specifies the direction of synchronization in the metacommunity. Note that the specific choice of orientation does not matter as long as it is fixed throughout time. Without loss of generality, we choose the orientation from origin to community 1 to community 2. Once we assign the orientation, the hypervolume would have signs, which means the hypervolume can be negative. From time (t) to $(t+1)$, the orientations of the geometric embeddings do not change (both are clockwise). The hypervolume overlap is simply the overlap between two positive hypervolumes, which equals to 1/4.

Then let us consider another example. Suppose at time $(t+2)$, community composition switches from time $(t+1)$ (i.e, community 1 only has species A, while community 2 only has species B; Figure 4C). In this case, the orientations of the geometric embeddings are opposite (clockwise versus anti-clockwise). The overlap now needs to consider the signed difference, which equals to $0.5-(-0.5)=1$ (despite the seemingly identical shape).

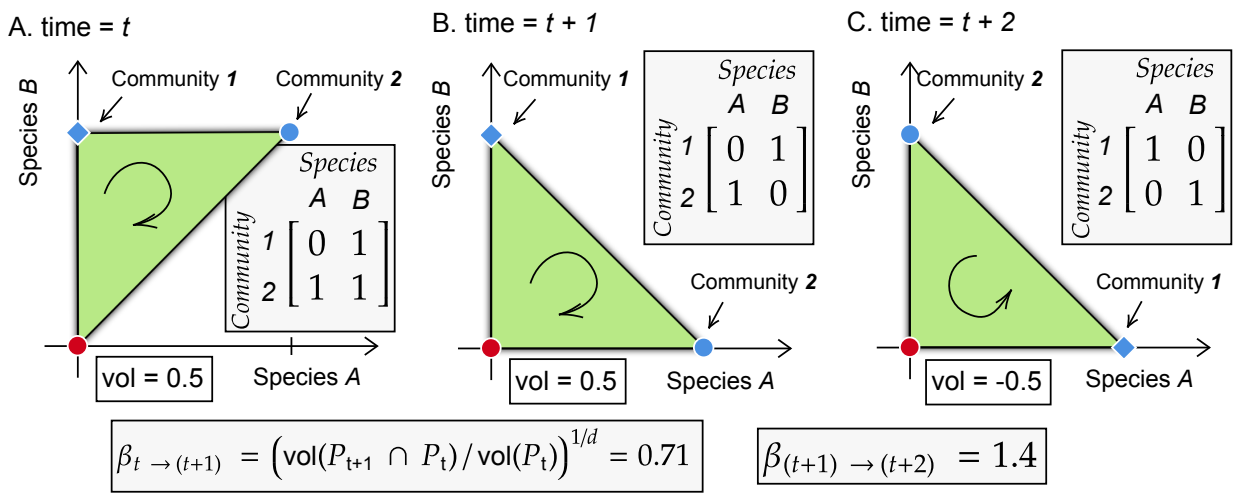

Figure 4: Measure temporal changes of beta diversity using (oriented) hypervolume overlap. We assign an orientation of hypervolume from origin to community 1 to community 2. The ecological interpretation of the orientation is the direction of synchronization in the metacommunity. Panels (A-C) represent metacommunity at time (t) to time $(t+2)$, respectively. From time (t) to time $(t+1)$, community 1 is unchanged while community 2 loses species B. The changes in community compositions are *asynchronized*, which are reflected in the identical orientations of their hypervolumes. In contrast, from time $(t+1)$ to time $(t+2)$, community 1 and community 2 switch their community compositions. The changes in community compositions are *synchronized*, which are reflected in the opposite orientations of their hypervolumes. With the definition of temporal change (Eqn. 8), beta diversity changes by 0.71 from time (t) to time $(t+1)$, while it changes by 1.4 from time $(t+1)$ to time $(t+2)$.

Formally, we can define the changes of beta diversity from time (t) to time $(t+1)$ as

$$\beta_{t\to(t+1)} = \left(\frac{\text{vol}(P_{t+1} \cap P_t)}{\text{vol}(P_t)}\right)^{1/d}, \tag{8}$$

which measures the extent of synchronous or asynchronous changes in community composition in the entire metacommunity. If $\left|\beta_{t\to(t+1)}\right| < 1$, then changes in community compositions are asynchronous or synchronous in the same direction. In contrast, if $\left|\beta_{t\to(t+1)}\right| > 1$, then changes in community compositions are synchronous in the opposite direction.

Applying this definition (Eqn. 8) to the examples above, the change in beta diversity equals to $(\frac{1}{4}/\frac{1}{2})^{(1/2)} = 0.71$ from time (t) to time $(t+1)$, while equals to $(1/\frac{1}{2})^{(1/2)} = 1.4$ from time $(t+1)$ to time $(t+2)$. This aligns with ecological intuition. From time (t) to time $(t+1)$, we see an asynchronous change in community compositions (community 1 remains fixed and only community 2 changes), which is reflected in a relatively smaller temporal change of beta diversity. In contrast, from time $(t{+}1)$ to time $(t{+}2)$, we see a synchronous change of community compositions (community 1 and community 2 switches composition), which is reflected in the relatively large temporal change of beta diversity.

3.3 Community/Species-specific contribution as hypervolume change

Communities do not contribute equally to biodiversity maintenance in a landscape. Thus, we need to disentangle the importance of community-specific contribution to beta diversity. Here, we measure the community-specific contribution using the relative change of hypervolumes.

From the perspective of our geometric approach, a community contributes to the overall beta diversity through its embedded points. Thus, to evaluate its relative contribution, we can compare the overlap between the hypervolumes with and without this community. To illustrate, we use the metacommunity example in Figure 3. Figure 5A shows the original metacommunity matrix and its embedded geometric object. Figure 5B-D shows geometric objects without site 1-3, respectively. A caveat observed here is that the sum of the hypervolumes in Figure 5B-D exceed the hypervolume of the original metacommunity. To account for this caveat, we introduce a normalization so that. Formally, the contribution of community i to beta diversity is

$$\mathscr{C}(\text{Community } i) = \frac{\mathsf{vol}(P_0)^{1/d} - \mathsf{vol}(P_i)^{1/d}}{\sum_j(\mathsf{vol}(P_0)^{1/d} - \mathsf{vol}(P_j)^{1/d})}, \tag{9}$$

where P_0 denotes the geometric object of the original metacommunity, P_i denotes the geometric object of the metacommunity without community i, and the summation index j runs through all communities. Applying Eqn. 9 to the above example, we found community 1 contributes 0.16, community 2 contributes 0.37, while community 3 contributes 0.47.

An important feature of our measure is that *all* communities have a non-negative contribution to beta diversity. This is because the hypervolume $\mathsf{vol}(P_0)$ of the original metacommunity is

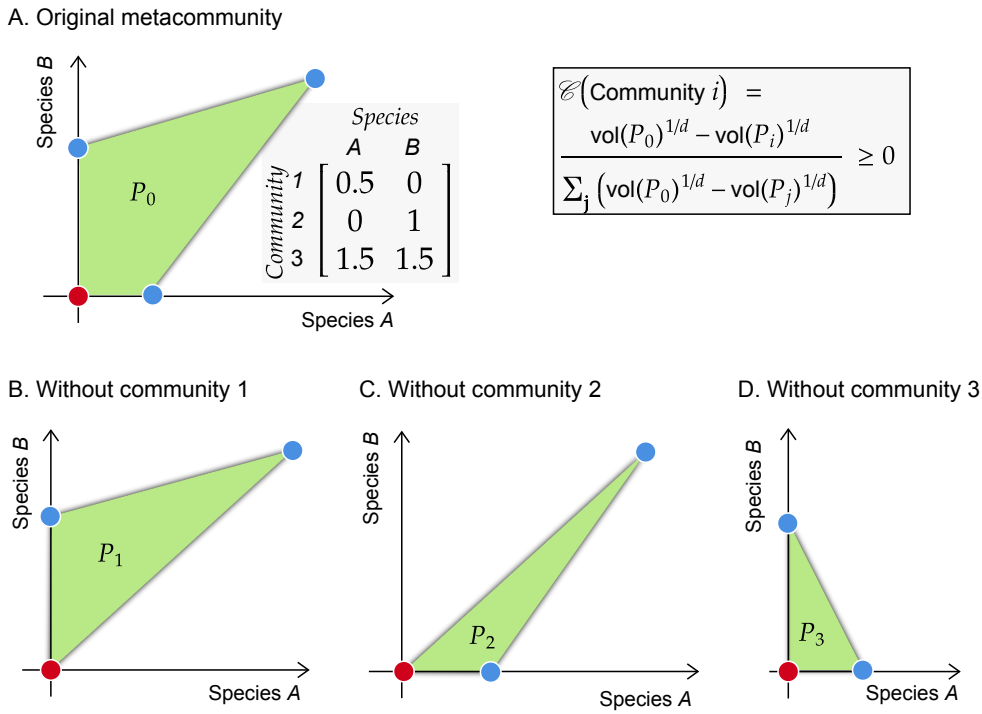

Figure 5: Disentangling site-specific contribution to beta diversity. Panel (A) shows an example metacommunity and its corresponding geometric embedding (the same as the example in Figure 3). Panels (B-D) show the metacommunities without community 1-3, respectively. The contribution of a community to the overall beta diversity is quantified as the normalized change in the hypervolumes. In this example, community 1 contributes 0.16, community 2 contributes 0.37, and community 3 contributes 0.47.

always greater than or equal to that $\text{vol}(P_i)$ of the metacommunity without a community (under the assumption of not using the modified schemes on duplications (Eqn. 7) and no changes in γ diversity). In contrast, in the traditional formalization of beta diversity, a community may have a negative contribution (i.e., its presence decreases the beta diversity). For example, in the metacommunity II in Figure 1B, community 3 would have a negative contribution with the traditional formalization (e.g., it decreases beta diversity by 25% with Whittaker's multiplicative measure). However, *ceteris paribus*, conservation management, in general, should *not* assign some community to be 'negative' for biodiversity (Hunter Jr & Gibbs, 2006). Thus, our framework is more appropriate to assess community contribution in conservation planning.

The above method parallelly applies to quantify species-specific contribution. A caveat is that we can only assess the contribution of either species or community (depending on which is the constraint on the embedded dimension), but not both.

3.4 Species similarity and functional complementarity as transformed embedding

Species are more similar to some species than others. To account for species similarity, we follow Leinster & Cobbold (2012) by introducing a S matrix where elements s_{ij} denote how similar species i is to species j. s_{ij} are scaled between 0 (totally dissimilar) to 1 (totally similar). For example, it can be a genetic or phenotypic (trait) similarity. Note that the S matrix is not required to be symmetric (i.e., $s_{ij} \neq s_{ji}$).

From our geometric perspective, the S matrix corresponds to a linear transformation of the embedded geometric object. For simplicity, let us consider 2 species. Originally, $(1, 0)$ denotes the presence of species A while $(0, 1)$ denotes the presence of species B. The two axes are orthogonal. With the introduction of the S matrix, the presence of species A is now indicated as $(\frac{1}{\sqrt{1+s_{AB}^2}}, \frac{s_{AB}}{\sqrt{1+s_{AB}^2}})$, while the presence of species A is now indicated as $(\frac{s_{BA}}{\sqrt{1+s_{BA}^2}}, \frac{1}{\sqrt{1+s_{BA}^2}})$. If all species are totally dissimilar, then the S matrix is an identity matrix. This corresponds to the same original axis (which is what we have been presenting so far; Figure 6A). For another example, if all species are similar, then the S matrix is a matrix with all 1s. This corresponds to all axes pointing to the exact same direction $(1, 1)$ (Figure 6B). In this case, the hypervolume would always be 0. This agrees with ecological intuition, because the system effectively only has 1 species and there is no beta diversity. For a non-trivial example, let us consider the S matrix $\begin{bmatrix} 1 & 0.5 \\ 0.5 & 1 \end{bmatrix}$. The hypervolume is now shrunk into a smaller region (Figure 6C).

Moving to the general case, we formalize the effect of the similarity matrix S as transforming the axes in the hyper-dimension space that the metacommunity is embedded into. To account for this, we simply need to compute the solid angle between all the axes. Mathematically, the solid angle $\Omega(S)$ formed by the similarity matrix S is given by (Ribando, 2006; Song *et al.*, 2018)

$$\Omega(S) \;=\; \frac{2^d}{(2\pi)^{S/2}\sqrt{|\,\mathsf{det}(S)|}} \int \cdots \int_{\boldsymbol{N}^* \geq 0} e^{-\frac{1}{2}\boldsymbol{N}^{*T} S^T S \boldsymbol{N}^*} d\,\boldsymbol{N}^*. \tag{10}$$

With the similarity matrix S, the hypervolume is transformed into $\Omega(S)\,\mathsf{vol}(\mathcal{P})$. Naturally, we can define diversity β_{vol} accounting for species similarity as

$$\beta_{\mathsf{vol}} \;= d \times (\Omega(S)\ \mathrm{vol}(\mathcal{P}))^{1/d}. \tag{11}$$

As the elements S_{ij} are always larger than 0, the transformed hypervolume and the associated beta diversity is always smaller accounting for species similarity. This aligns with ecological expectation because species being more similar would reduce the overall variation in the metacommunity.

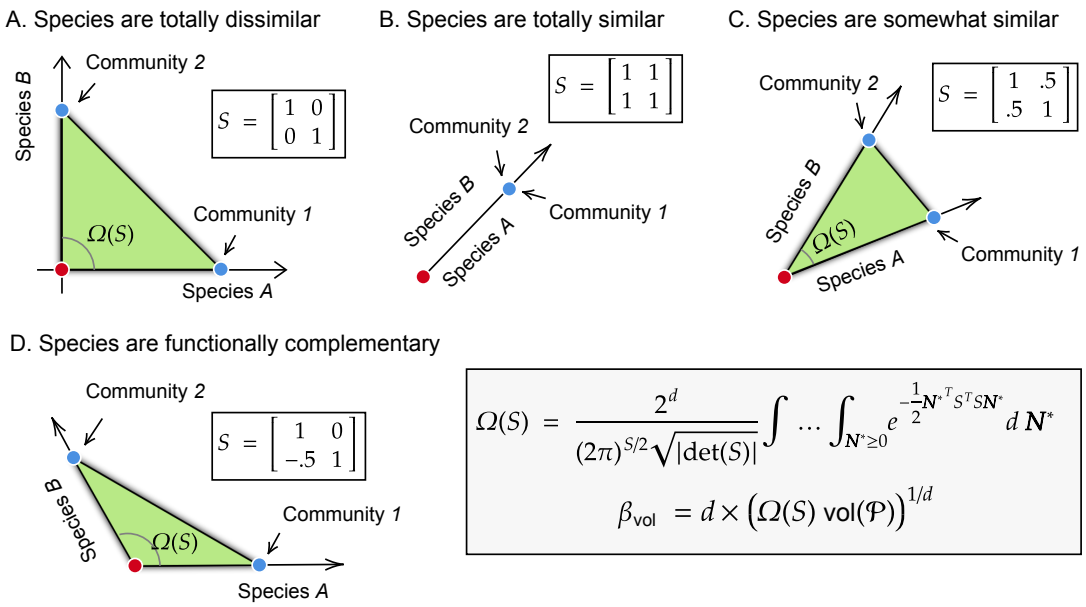

Figure 6: Accounting for species similarity and functional complementarity to quantify beta diversity. is equivalent to a coordinate transformation. All panels show the same metacommunity with different species' similarity or functional complementarity. The original metacommunity has two communities where one community only has species A and the other community only has species B (the same as metacommunity II in Figure 1A). Panel (A) shows the case where species are totally dissimilar. The hypervolume and the corresponding beta diversity remain the same. Panel (B) shows the case where species are totally similar. The hypervolume shrinks to 0 and there is no beta diversity. Panel (C) shows the case where species are a bit similar. The hypervolume is larger than 0 but shrinks compared to the case where the totally dissimilar case. Panel (D) shows the case where species are functionally complementary. This is reflected in $S_{21} < 0$. The hypervolume is expanded compared to the case where the totally dissimilar case.

In parallel to species similarity, we can also consider species functional complementarity. Functional complementarity means that two species provide additional ecological functioning than the addition of the functioning when both species are isolated (i.e., in monoculture) (Tilman *et al.*, 2014). Multiple methods are available to quantify functional complementarity from experiments (e.g., Loreau & Hector 2001; Alahuhta *et al.* 2017). We represent functional complementarity using the S matrix, where s_{ij} now denotes the level of functional complementarity species j provides to species i. Note that s_{ij} are negative, as they represent functional complementarity. Because of the negative s_{ij}, the hypervolume is now expanded (Figure 6D) compared to the case without any functional complementarity (Figure 6A). In general, the transformed hypervolume is always larger accounting for functional complementarity. This aligns with the ecological expectation because more variations in ecosystem functioning would increase the overall variation in the metacommunity.

While species similarity and functional complementarity are related concepts, they can have different focuses. For example, accounting for species similarity allows quantifying phylogenetic beta diversity, while accounting for functional complementarity allows quantifying functional beta diversity. As a side note, the formulas accounting for functional complementarity are identical to

those for species similarity (Eqns. 10 and 11). To our knowledge, despite the apparent symmetry between species similarity and functional complementarity, this extension of functional complementarity is not obvious to achieve using the traditional formalism using Hill's number (Leinster & Cobbold, 2012).

3.5 Nestedness-turnover decomposition as filling-finding facets

Decomposing beta diversity into turnover and nestedness components is a major advance in our understanding of beta diversity (Baselga, 2012; Legendre, 2014). Turnover (also known as replacement) means that species compositions tend to replace each other along spatial gradients. Nestedness (also known as richness difference) means that species composition in a community is a strict subset of the species composition in a richer community. Here, we provide a geometric interpretation of the nestedness-turnover decomposition.

For illustrative purposes, let us consider two metacommunities with one showing complete turnover and the other one being completely nested. The metacommunity matrix describing the metacommunity with complete turnover is (the corresponding geometric embedding illustrated in Figure 7A):

$$\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}, \tag{12}$$

and the metacommunity matrix of the nested metacommunity is (the corresponding geometric embedding illustrated in Figure 7B):

$$\begin{bmatrix} 1 & 1 & 1 \\ 1 & 1 & 0 \\ 1 & 0 & 0 \end{bmatrix}. \tag{13}$$

To compare the geometric embeddings of the two metacommunities, we observe that all the embedded points are located on *different* facets of the cube in the turnover metacommunity, while all the embedded points are located on the *same* facet of the cube in the nested community. In other words, the turnover process increases the beta diversity by finding new facets, while the nestedness process increases the beta diversity by filling a facet.

In contrast to previous sections, we did not provide an analytic measure to partition geometric beta diversity into nestedness and turnover parts. This is because our geometric approach suggests that this problem may be inherently ill-defined: nestedness is essentially a multidimensional property that cannot be reduced into a single scalar index. As a metacommunity with γ diversity has 2γ

facets, nestedness should be represented as a 2γ-dimensional vector, where each element denotes how much each facet is filled. To make it even more complicated, each element in the nestedness vector is interwinded with another element, as filling one facet can affect how another facet is filled. Thus, it is difficult, if not impossible, to summarize the nestedness vector into a 1-dimensional index without losing ecological information. Our observation complements the arguments that nestedness and turnover are interactive and thus cannot be partitioned (Šizling *et al.*, 2022).

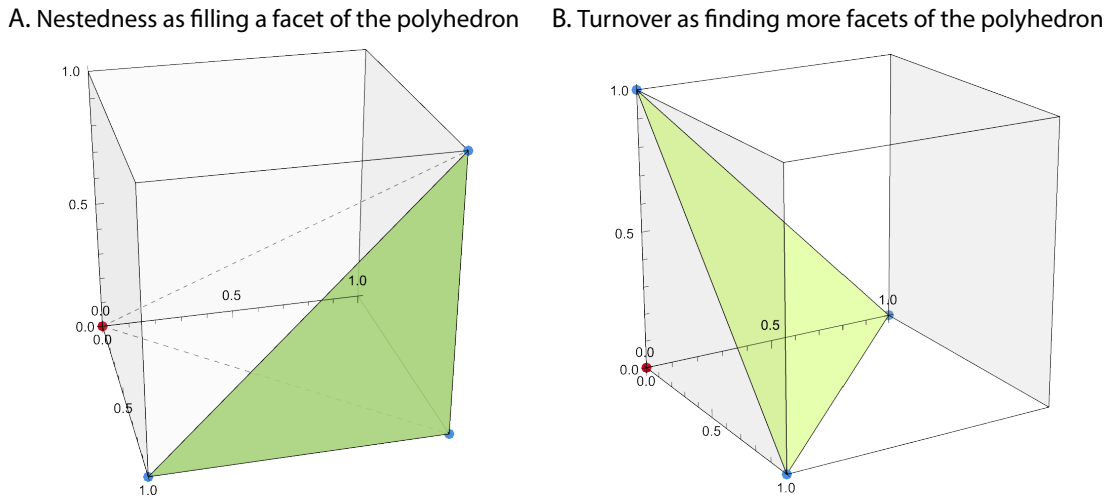

Figure 7: Geometric interpretation of nestedness and turnover decomposition. We consider here two metacommunities both with 3 communities and 3 species. Panel (A) shows the archetypical example of nestedness. All the points are located on the *same* facet of the 3-dimensional cube. The geometric interpretation of nestedness is to *fill* more facets of the cube. Panel (B) shows the archetypical example of turnover. All the points (each represents a community) are located on the *different* facets of the 3-dimensional cube. The geometric interpretation of turnover is to *find* more facets of the cube. These geometric interpretations generalize to higher dimension, where we replace cube as high-dimensional polyhedron.

Box 1 | Linking the geometric approach to traditional formalisms

Despite the differences of our geometric approach to traditional formalisms, our approach has a strong connection to them. The bridge across different formalisms exists by forming different geometric shapes from the same embedded points of the metacommunity. The main text has focused on forming a convex hull from the embedded points. However, there are other alternative choices (such as an ellipse). Different geometric shapes would result in different hypervolumes (consequently, different beta diversity). Importantly, different shapes (i.e., geometric beta diversity) emphasize different ecological properties. Here, in addition to the convex hull approach in the main text, this box introduces two other geometric approaches in forming shapes. Appendix E provides intuitions behind these two definitions, as well as mathematical derivations.

The first approach connects with the dominant formalism of beta diversity based on generalized variance (Legendre & De Cáceres, 2013). We define the hypervolume as determinant of the covariance matrix of the metacommunity matrix. The geometric interpretation of this hypervolume is the cor-

responding ellipse formed by the embedded metacommunity (Lu *et al.*, 2021). Similar to Eqn. (5), the geometric beta diversity β_{VAR} is defined as $d \times \mathsf{det}(\mathsf{VAR}(X))^{1/d}$ where X is the metacommunity matrix. Figure 8A-B illustrates two examples of metacommunity with β_{VAR}. This formalism naturally partitions composition variation into a traditional beta diversity measure and a spatial association component.

The second approach connects with the dominant formalism of beta diversity based on information theory (Chao *et al.*, 2014). We define the hypervolume as the joint entropy $H(X)$ of a multivariate Bernoulli distribution. The geometric interpretation of this hypervolume is the effective size of the support formed by the embedded metacommunity (Grendar, 2006). Similar to Eqn. (5), the geometric beta diversity β_{info} is defined as $d \times H(X)^{1/d}$ where X is the metacommunity matrix. Figure 8C-D illustrates two examples of metacommunity with β_{info}. This formalism allows interpreting beta diversity in the language of information, e.g., spatial association as mutual information. This formalism is also closely connected with zeta diversity (Hui & McGeoch, 2014).

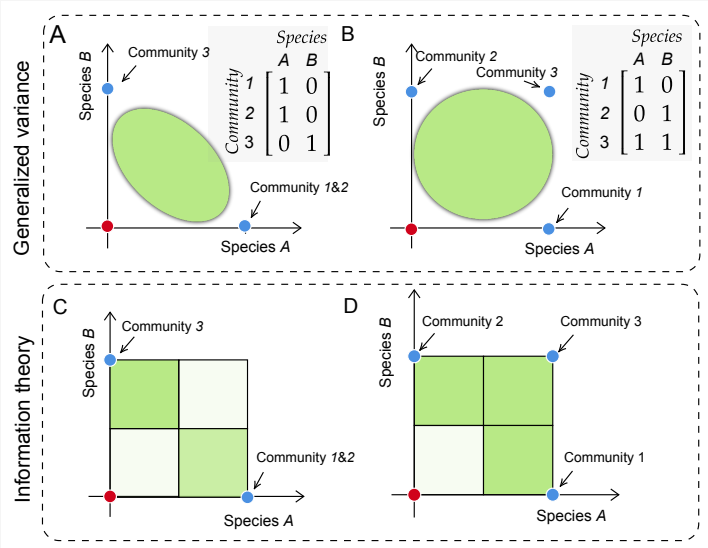

Figure 8: Connecting hypervolume beta diversity to traditional formalisms. We consider again metacommunity I and II from Figure 1A). Instead of forming a convex hull from the geometric embedding (Figure 1), we form either ellipse or effective support size from it. Panels (A) and (B) focus on the hypervolume of elliptic shapes formed by embedded metacommunity. This geometric shape closely connects with generalized variance. Specifically, the shape of the ellipse is determined by the structure of the covariance matrix. Panel (C) and (D) focus on the hypervolume of a multivariate Bernoulli random variable as entropy (a.k.a., the effective size of the support; Grendar 2006). Darker shading of the color represents a higher contribution of the state to the total entropy (given by the terms in Eqn. S10). The metacommunity with more evenly distributed community compositions has a higher effective size of the support and therefore higher beta diversity.

4 Empirical applications

4.1 Efficient estimation of beta diversity

To apply our measures of beta diversity to empirical data, we need to estimate the hypervolumes of the embedded metacommunity. The hypervolume of geometric shapes in high dimension is notoriously difficult to estimate. Fortunately, we do not need to compute the hypervolume of arbitrary geometric shapes (e.g., this is typically required for fundamental niches). Appendix B discusses how to compute hypervolume beta diversity (β_{vol}, β_{VAR} and β_{info}) in detail. We have provided an R package `betavolume` (`https://github.com/clsong/betavolume`) to assist with these calculations. In brief, although exact hypervolume is only computationally feasible for metacommunities with 15 or fewer communities or species, robust approximations can be made for metacommunities that are much larger (even for more than $10,000$ species or communities). A detailed discussion can be found in Appendix B. This package provides a user-friendly interface in R language to compute beta diversity β_{vol} and its various extensions (including duplications in presence/absence data, community/species-specific contribution, species similarity and functional complementarity).

4.2 Latitudinal pattern of beta diversity

Through the years, a high-profile debate has centered on latitudinal patterns of beta diversity (Currie *et al.*, 2004; Kraft *et al.*, 2011; Qian *et al.*, 2013; Xing & He, 2021a). The dataset used in the debate is forest transect data, which contains 198 locations along a latitudinal gradient (Gentry, 1988; Janni, 2003). Each location has a plot that can be considered as a metacommunity of 10 communities. Previous research using traditional measures of beta diversity has reached contrasting conclusions: beta diversity decreases along the absolute latitude gradient when using Whittaker's multiplicative measure (Currie *et al.*, 2004), while it shows a null pattern with absolute latitude when using an alternative measure known as beta deviation (Kraft *et al.*, 2011). Importantly, both patterns originate from the exponential decrease of gamma diversity along the absolute latitude gradient (Figure 9C). Specifically, the pattern with Whittaker's multiplicative measure is fully driven by gamma diversity, as an exponential decrease in gamma diversity completely masks the effects of alpha diversity. In contrast, the pattern with beta deviation is due to the ignorance of gamma diversity, as beta deviation removes the effect of changing gamma diversity (Bennett & Gilbert, 2016). However, both metrics largely ignore the role of the variance of alpha diversity in shaping the latitudinal pattern of beta diversity.

We applied our measure β_{vol} to this dataset (Gentry, 1988; Janni, 2003). In contrast to the previous consensus, we find a unimodal pattern of β_{vol}: it first increases and then decreases along the absolute latitude gradient (Figure 9B). This pattern emerges from the conflicting roles of alpha and gamma diversity. On the one hand, the decreasing gamma diversity has a negative effect on β_{vol} (Figure 9C). As a new species is equivalent to a new point in the embedded space, the hypervolume of a metacommunity would not decrease by adding a new species to the metacommunity. On the other hand, the increasing trend of the variance in alpha diversity has a positive force on β_{vol} (Figure 9A). This is because the increasing difference in alpha diversity across communities increases the nestedness of the metacommunity (Baselga, 2010). The high gamma diversity and low variance of alpha diversity in the lower latitude suggest that the metacommunities in the region are characterized by strong mutual exclusions among species, while the low gamma diversity and high variance of alpha diversity in the higher latitude suggest that the metacommunities are characterized by high nestedness. The unimodal pattern of β_{vol} therefore indicates that the-mid latitude metacommunities have the lowest spatial structure. The unimodal pattern is consistent with observations from another global dataset on plant diversity (Scheiner & Rey-Benayas, 1994).

Note that the unimodal pattern is not our key take-away. Given the spatial and temporal biases in global biodiversity datasets (Gonzalez *et al.*, 2016; Hughes *et al.*, 2021), there is plenty of room for disagreement on which is the true latitudinal pattern of beta diversity. Nonetheless, as both alpha and gamma diversity vary considerably in global syntheses of biodiversity data, a satisfactory measure of beta diversity should be able to extract the joint effects of both alpha and gamma diversity. Our measure β_{vol} is capable of doing this, while previous measures can mostly extract the information captured by gamma diversity.

4.3 How sampling efforts affect beta diversity

In empirical estimation of beta diversity, sampling efforts play a prominent role. That is, with traditional measures of multiple-site beta diversity, beta diversity always increases when more sites are sampled. This increase in beta diversity is mainly driven by the increase in gamma diversity (Bennett & Gilbert, 2016; Xing & He, 2021a). However, this begets two problems: first, more sampling may not pay off, as it provides exponentially diminishing returns; second, we cannot distinguish which metacommunity is more spatially heterogeneous. Our measure β_{vol} may solve these two problems, because it does not necessarily increase with higher gamma diversity. A higher gamma diversity would increase the dimension of the embedded space for the metacommunity, which could decrease the rescaled hypervolume.

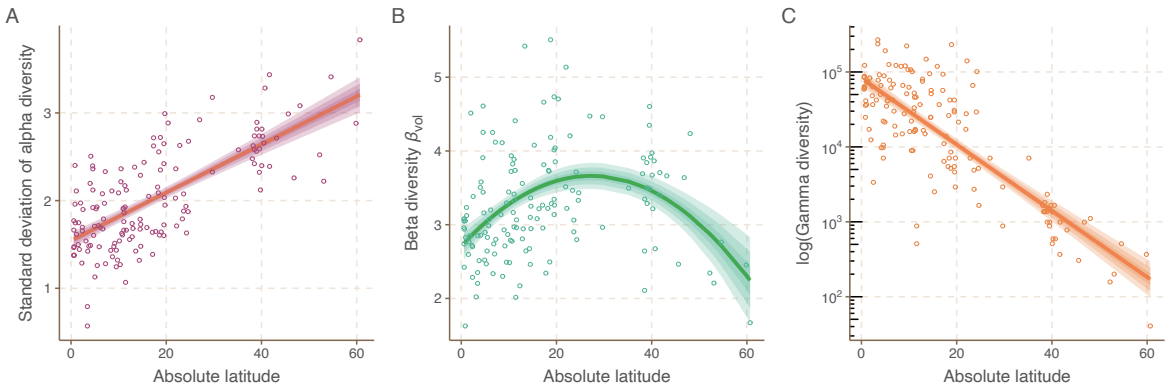

Figure 9: The pattern of beta diversity β_{vol} along latitudes and its origin. We show how (variance of) alpha diversity (Panel A), beta diversity (Panel B), and gamma diversity (Panel C) changes along the absolute latitude gradient. The horizontal axis shows the absolute latitude, while the vertical axis shows the measure of diversity. Each point represents a metacommunity. We depicted the generalized additive line with shaded confidence intervals. Panel (A) uncovers a monotonically increasing trend of the variance of alpha diversity (adjusted $R^2 = 0.64$). Panel (B) shows our measure β_{vol} has a unimodal pattern ($p = 0.99$ according to Hartigans' dip test; Hartigan & Hartigan 1985). This is in direct contrast to previous results, where beta diversity is either monotonically decreasing or does not change. Panel (C) shows gamma diversity exponentially decreases (adjusted $R^2 = 0.65$).

As a proof of concept, we focused on two datasets from Bennett & Gilbert (2016). One dataset contains 1-m^2 plots in early successional fields in the Koffler scientific reserve in Ontario, Canada. Another dataset contains 50-m^2 forest plots at Mont St. Hilaire near Montreal, Canada (Gilbert & Lechowicz, 2004). These two datasets were collected for different purposes. The data from Koffler Scientific Reserve were designed to sample a relatively homogenous area, while the data from the Mont St. Hilaire were acquired to capture environmental heterogeneity. Previous research has shown that traditional beta diversity in both datasets would increase with sampling effort with a power-law scaling (Bennett & Gilbert, 2016; Xing & He, 2021a). Thus, traditional measures fail to capture the ecological differences between the two datasets.

We apply our measure β_{vol} to these two datasets (Figure 10) using the random subsampling procedure of Bennett & Gilbert (2016). For the data from the Mont St. Hilaire, beta diversity always increases with the sampling effort. This increasing pattern is coherent with the fact that they were sampled from a relatively more heterogeneous area. In contrast, for the data from the Koffler Scientific Reserve, beta diversity initially increases with sampling effort but then decreases until it reaches a steady value. This pattern is also consistent with the fact that they were sampled from a relatively more homogeneous area. Thus, our measure β_{vol} has the potential to solve the long-standing issue on sampling efforts associated with beta diversity: more sampling is necessary to detect the ecological differences in spatial heterogeneity between the two datasets.

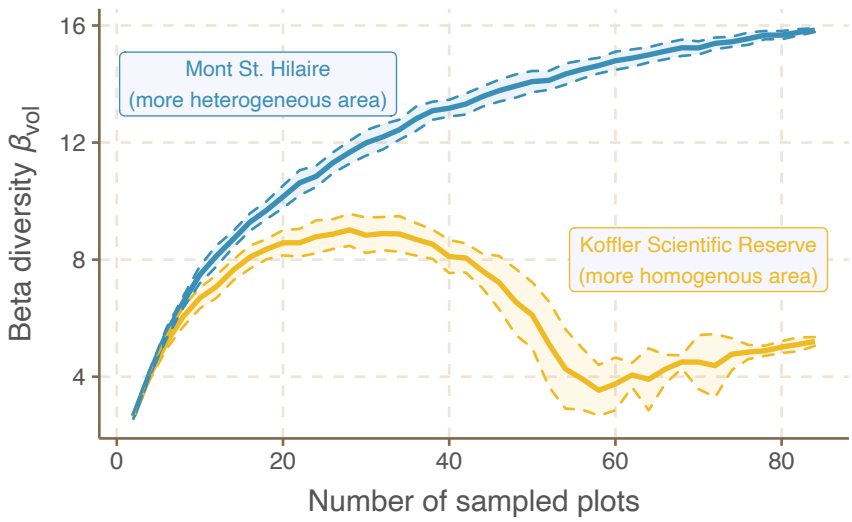

Figure 10: The pattern of beta diversity β_{vol} with sampling effort. The horizontal axis represents the sampling effort (as in the number of sampled plots). The vertical axis represents our measure of beta diversity β_{vol}. The thick lines show the average beta diversity with a given sampling effort for the data from Koffler Scientific Reserve (orange) and for the data from Mont St. Hilaire (blue). The dashed regions show one standard deviation of sampled beta diversity. While the trends of beta diversity are similar with a small sampling effort in both datasets, they quickly diverge. The beta diversity keeps smoothly increasing in the data from Mont St. Hilaire. In contrast, beta diversity decreases to a slowly increasing plateau in the data from Koffler Scientific Reserve.

5 Discussion

Beta diversity is a central concept in spatial ecology and conservation management (Mori *et al.*, 2018). Unlike the consensus on measures of alpha and gamma diversity (Jost, 2007; Chao *et al.*, 2014), there is a long list of beta diversity measures. One may argue that the pressing problem now should be classifying or reconciling these measures of beta diversity (Jurasinski *et al.*, 2009). If this is so, then why introduce a new measure at this point? We believe our new measure of beta diversity is much needed because (1) it reconceptualizes beta diversity, (2) its geometric nature makes it easily extendible and generalizable, (3) it synthesizes traditional measures, and (4) it provides novel ecological insights. We discuss these four advantages below.

First, our new measure behaves qualitatively differently from all traditional measures. Our measure is built upon a core observation that beta diversity should be maximized when we observe all possible community compositions in the region (Figure 1). In short, the more the merrier. In contrast, traditional wisdom posits beta diversity is maximized when each community only has a single distinct species (see review in Legendre & De Cáceres 2013). The traditional wisdom operates under an individualist perspective. That is, the 'ecology' of a species is the same with or without the presence of other species. Under this individualist perspective, each community with a single unique species would display the maximum possible variance of biodiversity in the

region. However, the individualist perspective is unlikely to be general in ecological systems where interactions abound. Thus, our measure is conceptually justified as long as species interactions in a local community affect species dynamics and functioning.

It is not a trivial problem to formalize this reconceptualization of beta diversity. To our knowledge, among the traditional measures, the only exception to traditional wisdom is the Shannon diversity of realized species combinations (Juhász-Nagy & Podani, 1983). This measure proposed to simply count the number of unique community compositions (Juhász-Nagy & Podani, 1983). However, this measure ignores quantitatively how community compositions are different. For example, a community with species A and B should be more distinct from a community with species C than a community with composition A. We have taken a hypervolume approach to solve this problem. Hypervolume is an old friend in ecology, and was used most famously by Hutchinson to frame the discussion of the niche (Blonder, 2018). The idea of hypervolume has been widely used in various areas of ecology research (Raup & Michelson, 1965; Violle & Jiang, 2009; Boucher *et al.*, 2013; Blonder *et al.*, 2014). Notably, researchers have measured functional beta diversity as the *overlap* between the functional trait spaces of two local communities (Mammola, 2019; Lu *et al.*, 2021). In contrast to these previous works, our measure is fundamentally different, as we directly interpret hypervolume of the metacommunity matrix as beta diversity. To do so, we have followed the idea of Hutchinson where he interpreted the fundamental niche as hyper-dimensional geometric shapes (Hutchinson, 1957). Our geometric measure provides a linear scaling between beta diversity and the number of unique community compositions, while it also quantifies the difference between unique community compositions (Figures 2 and S5).

Second, our approach provides a unifying framework for beta diversity. Given the importance of beta diversity, the basic quantification is far from enough for empirical study. We have extended our geometric measure to the following five cases with strong empirical importance: duplications in presence/absence data (Figure 3), temporal changes (Figure 4), community/species-specific contributions to beta diversity (Figure 5), species similarity and functional complementarity (Figure 6), and turnover-nestedness decomposition (Figure 7). While these extensions are possible with traditional measures of beta diversity, they require different theoretical formalisms. In part, this may result from the fact that most traditional measures are algebraic manipulations of metacommunity matrix without a simple geometric interpretation. In contrast, we present a geometric approach, which is fully visual in 2- or 3-dimensional space. This visual aspect of our geometric approach permits an intuitive and generalizable ecological interpretation. A psychological benefit with our approach is that humans are intrinsically more familiar with geometry than algebra

(Sablé-Meyer *et al.*, 2021). Thus, our geometric measure is, in general, easier to visualize, interpret, and generalize than traditional algebraic definitions.

Third, our measure provides a unifying approach to synthesize previous measures of beta diversity. We are not simply adding a yet another measure to the list of beta diversity measures. Instead, our measure considers new higher-order information that traditional measures have missed. Despite the variety of traditional measures, most of them can be classified into two schools of thought: variance-based or information-based. The variance-based measure considers the diagonal of the covariance matrix (Legendre & De Cáceres, 2013), while we have in addition considered the off-diagonal. These off-diagonal components represent ecologically the spatial associations of species (Figures 1 and 8). The most commonly used information-based measure considers the pooled marginal entropy of a joint distribution (Jost, 2007), while the joint entropy takes the mutual information into account (Juhász-Nagy & Podani, 1983). In other words, previous measures of beta diversity have a geometric basis, and our approach reveals their hidden geometric nature.

Fourth, our measure provides novel ecological insights into the patterns in empirical data. We have focused on two important empirical issues: global syntheses of biodiversity data and the sampling efforts. Focusing on global syntheses, traditional measures are masked by the exponentially changing gamma diversity; thus, the latitudinal pattern is mostly driven by gamma diversity. In contrast, our measure can reveal the joint effects of alpha and gamma diversity in shaping the patterns of beta diversity (Figure 9). Focusing on the sampling efforts, traditional measures fail to reveal additional information with increasing sampling effort. This is because traditional measures are again masked by increasing gamma diversity with increasing sampling effort. In contrast, we show that increasing sampling effort is necessary to detect hidden spatial heterogeneity, and our measure can help quantify this heterogeneity (Figure 10). Besides the demonstrated examples, we also expect that our metric should be particular useful in determining the relationship between species composition and ecosystem functioning (Grman *et al.*, 2018; Mori *et al.*, 2018) and stability (McGranahan *et al.*, 2018) because it explicitly takes species-association into account. When applied in the temporal context, the hypervolume-based beta diversity is also a measure of community change predictability (Song *et al.*, 2021; De Cáceres *et al.*, 2019): for example, in time-lag analysis, higher beta diversity indicates more random community composition changes over time while lower beta diversity indicates more directional changes (Jones *et al.*, 2017).

Like other beta diversity metrics, our method is not without limitations. One major issue is that hypervolume beta diversity is sensitive to normalization of the elements in the metacommunity matrix (Legendre & De Cáceres, 2013). For example, beta diversity is likely to be different when we consider the absolute versus relative species abundance. However, we consider it to be a feature

rather than a bug. For example, under some ecological rationales, we can argue that a metacommunity with more individuals, *ceteris paribus*, is more "diverse" than another metacommunity with fewer individuals (Legendre & De Cáceres, 2013). We suggest that every normalization method requires careful ecological interpretation. As long as we apply the same normalization method across metacommunities of interests, we can safely compare which metacommunity has a higher beta diversity (in accordance with the ecological rationale behind the normalization).

We believe that our proposed measure is readily applicable to existing data. To further expand its applicability, we envision the following extensions of our geometric framework. One future direction is to further explore geometric features of the embedded metacommunity. For example, we have not yet considered its geometric asymmetry. For example, let us consider two metacommunities are both embedded as triangles with identical volume, but one is equilateral while the other one is not. The ecological differences between them is that the equilateral metacommunity has more balanced species distributions across local communities. To quantify the association between geometric asymmetry and species balance, it might be useful to adapt tools from studies on the geometric asymmetry in different ecological contexts (Grilli *et al.*, 2017; Medeiros *et al.*, 2021). Another future direction is to develop analytic models of null models. Null models are widely used in beta diversity analysis to disentangle confounding factors. Analytic null models are available for many traditional measures of beta diversity (Xing & He, 2021b; Lu *et al.*, 2019; Lu, 2021; Deane *et al.*, 2022). An analytic expression for β_{vol} is challenging because the complexity in quantifying hypervolume(Appendix B). However, β_{VAR} and β_{info} are tightly linked to high-dimensional normal distributions, thus it is possible to obtain analytic expressions. Furthermore, a promising future direction is to extend our geometric approach to measure functional diversity (reviewed in Scheiner *et al.* 2017).

6 Conclusion

We have reconceptualized the measurement of beta diversity as a geometric hypervolume. We have shown the connections of our new measure to existing variance- and information-based measures. Our geometric approach provides a unified way to measure beta diversity that can deal with duplications in presence/absence data, temporal change, turnover, nestedness, species and functional complementarity. We demonstrated its application to two datasets and the novel insights it offers. We have provided the tools needed to apply our approach. We hope these measures will be widely tested and applied and provide insights into the structure and function of metacommunities.

Acknowledgement: We thank Jonathan Levine, Samuel M. Scheiner, Nate Sanders, and Michael Catchen for insightful discussions.

References

Alahuhta, J., Kosten, S., Akasaka, M., Auderset, D., Azzella, M. M., Bolpagni, R., Bove, C. P., Chambers, P. A., Chappuis, E., Clayton, J., de Winton, M., Ecke, F., Gacia, E., Gecheva, G., Grillas, P., Hauxwell, J., Hellsten, S., Hjort, J., Hoyer, M. V., Ilg, C., Kolada, A., Kuoppala, M., Lauridsen, T., Li, E. H., Lukács, B. A., Mjelde, M., Mikulyuk, A., Mormul, R. P., Nishihiro, J., Oertli, B., Rhazi, L., Rhazi, M., Sass, L., Schranz, C., Søndergaard, M., Yamanouchi, T., Yu, Q., Wang, H., Willby, N., Zhang, X. K. & Heino, J. (2017). Global variation in the beta diversity of lake macrophytes is driven by environmental heterogeneity rather than latitude. *Journal of Biogeography*.

Anderson, M. J., Crist, T. O., Chase, J. M., Vellend, M., Inouye, B. D., Freestone, A. L., Sanders, N. J., Cornell, H. V., Comita, L. S., Davies, K. F. *et al.* (2011). Navigating the multiple meanings of β diversity: a roadmap for the practicing ecologist. *Ecology Letters*, 14, 19–28.

Angulo, M. T., Kelley, A., Montejano, L., Song, C. & Saavedra, S. (2021). Coexistence holes characterize the assembly and disassembly of multispecies systems. *Nature Ecology & Evolution*, 5, 1091–1101.

Baselga, A. (2010). Partitioning the turnover and nestedness components of beta diversity. *Global Ecology and Biogeography*, 19, 134–143.

Baselga, A. (2012). The relationship between species replacement, dissimilarity derived from nestedness, and nestedness. *Global Ecology and Biogeography*, 21, 1223–1232.

Bennett, J. R. & Gilbert, B. (2016). Contrasting beta diversity among regions: how do classical and multivariate approaches compare? *Global Ecology and Biogeography*, 25, 368–377.

Blanchet, F. G., Cazelles, K. & Gravel, D. (2020). Co-occurrence is not evidence of ecological interactions. *Ecology Letters*, 23, 1050–1063.

Blonder, B. (2018). Hypervolume concepts in niche-and trait-based ecology. *Ecography*, 41, 1441–1455.

Blonder, B., Lamanna, C., Violle, C. & Enquist, B. J. (2014). The n-dimensional hypervolume. *Global Ecology and Biogeography*, 23, 595–609.

Boucher, F. C., Thuiller, W., Arnoldi, C., Albert, C. H. & Lavergne, S. (2013). Unravelling the architecture of functional variability in wild populations of polygonum viviparum l. *Functional ecology*, 27, 382–391.

Chao, A., Chiu, C.-H. & Jost, L. (2014). Unifying species diversity, phylogenetic diversity, functional diversity, and related similarity and differentiation measures through hill numbers. *Annual review of ecology, evolution, and systematics*, 45, 297–324.

Currie, D. J., Mittelbach, G. G., Cornell, H. V., Field, R., Guégan, J.-F., Hawkins, B. A., Kaufman, D. M., Kerr, J. T., Oberdorff, T., O'Brien, E. *et al.* (2004). Predictions and tests of climate-based hypotheses of broad-scale variation in taxonomic richness. *Ecology letters*, 7, 1121–1134.

De Cáceres, M., Coll, L., Legendre, P., Allen, R. B., Wiser, S. K., Fortin, M.-J., Condit, R. & Hubbell, S. (2019). Trajectory analysis in community ecology. *Ecological Monographs*, 89, e01350.

Deane, D. C., Xing, D., Hui, C., McGeoch, M. & He, F. (2022). A null model for quantifying the geometric effect of habitat subdivision on species diversity. *Global Ecology and Biogeography*, 31, 440–453.

Gentry, A. H. (1988). Changes in plant community diversity and floristic composition on environmental and geographical gradients. *Annals of the Missouri botanical garden*, 1–34.

Gilbert, B. & Lechowicz, M. J. (2004). Neutrality, niches, and dispersal in a temperate forest understory. *Proceedings of the National Academy of Sciences*, 101, 7651–7656.

Gonzalez, A., Cardinale, B. J., Allington, G. R., Byrnes, J., Arthur Endsley, K., Brown, D. G., Hooper, D. U., Isbell, F., O'Connor, M. I. & Loreau, M. (2016). Estimating local biodiversity change: a critique of papers claiming no net loss of local diversity. *Ecology*, 97, 1949–1960.

Gonzalez, A., Germain, R. M., Srivastava, D. S., Filotas, E., Dee, L. E., Gravel, D., Thompson, P. L., Isbell, F., Wang, S., Kéfi, S. *et al.* (2020). Scaling-up biodiversity-ecosystem functioning research. *Ecology Letters*, 23, 757–776.

Grendar, M. (2006). Entropy and Effective Support Size. *Entropy*, 8, 169–174.

Grilli, J., Adorisio, M., Suweis, S., Barabás, G., Banavar, J. R., Allesina, S. & Maritan, A. (2017). Feasibility and coexistence of large ecological communities. *Nature communications*, 8, 0.

Grman, E., Zirbel, C. R., Bassett, T. & Brudvig, L. A. (2018). Ecosystem multifunctionality increases with beta diversity in restored prairies. *Oecologia*, 188, 837–848.

Hartigan, J. A. & Hartigan, P. M. (1985). The dip test of unimodality. *The annals of Statistics*, 70–84.

Hughes, A. C., Orr, M. C., Ma, K., Costello, M. J., Waller, J., Provoost, P., Yang, Q., Zhu, C. & Qiao, H. (2021). Sampling biases shape our view of the natural world. *Ecography*, 44, 1259–1269.

Hui, C. & McGeoch, M. A. (2014). Zeta diversity as a concept and metric that unifies Incidence-based biodiversity patterns. *The American Naturalist*, 184, 684–694.

Hunter Jr, M. L. & Gibbs, J. P. (2006). *Fundamentals of conservation biology*. John Wiley & Sons.

Hutchinson, G. (1957). Concluding remarks cold spring harbor symposia on quantitative biology, 22: 415–427. *GS SEARCH*.

Janni, K. D. (2003). Global patterns of plant diversity: Alwyn h. gentry's forest transect data set. *Economic Botany*, 57, 283–283.

Jones, S. K., Ripplinger, J. & Collins, S. L. (2017). Species reordering, not changes in richness, drives long-term dynamics in grassland communities. *Ecology Letters*, 20, 1556–1565.

Jost, L. (2007). Partitioning diversity into independent alpha and beta components. *Ecology*, 88, 2427–2439.

Juhász-Nagy, P. & Podani, J. (1983). Information theory methods for the study of spatial processes and succession. *Vegetatio*, 51, 129–140.

Jurasinski, G., Retzer, V. & Beierkuhnlein, C. (2009). Inventory, differentiation, and proportional diversity: a consistent terminology for quantifying species diversity. *Oecologia*, 159, 15–26.

Kraft, N. J., Comita, L. S., Chase, J. M., Sanders, N. J., Swenson, N. G., Crist, T. O., Stegen, J. C., Vellend, M., Boyle, B., Anderson, M. J. *et al.* (2011). Disentangling the drivers of β diversity along latitudinal and elevational gradients. *Science*, 333, 1755–1758.

Legendre, P. (2014). Interpreting the replacement and richness difference components of beta diversity. *Global Ecology and Biogeography*, 23, 1324–1334.

Legendre, P. & De Cáceres, M. (2013). Beta diversity as the variance of community data: dissimilarity coefficients and partitioning. *Ecology Letters*, 16, 951–963.

Legendre, P. & De Cáceres, M. (2013). Beta diversity as the variance of community data: dissimilarity coefficients and partitioning. *Ecology Letters*, 16, 951–963.

Leinster, T. & Cobbold, C. A. (2012). Measuring diversity: the importance of species similarity. *Ecology*, 93, 477–489.

Levine, J. M., Bascompte, J., Adler, P. B. & Allesina, S. (2017). Beyond pairwise mechanisms of species coexistence in complex communities. *Nature*, 546, 56–64.

Loreau, M. & Hector, A. (2001). Partitioning selection and complementarity in biodiversity experiments. *Nature*, 412, 72–76.

Lu, M. (2021). Complex relationships between beta diversity and dispersal in meta-community models. *Ecography*, 44, 1769–1780.

Lu, M., Vasseur, D. & Jetz, W. (2019). Beta diversity patterns derived from island biogeography theory. *The American Naturalist*, 194, E52–E65.

Lu, M., Winner, K. & Jetz, W. (2021). A unifying framework for quantifying and comparing n-dimensional hypervolumes. *Methods in Ecology and Evolution*, 12, 1953–1968.

Mammola, S. (2019). Assessing similarity of n-dimensional hypervolumes: Which metric to use? *Journal of Biogeography*, 46, 2012–2023.

Maron, P.-A., Sarr, A., Kaisermann, A., Lévêque, J., Mathieu, O., Guigue, J., Karimi, B., Bernard, L., Dequiedt, S., Terrat, S. *et al.* (2018). High microbial diversity promotes soil ecosystem functioning. *Applied and Environmental Microbiology*, 84, e02738–17.

McGranahan, D. A., Hovick, T. J., Elmore, R. D., Engle, D. M. & Fuhlendorf, S. D. (2018). Moderate patchiness optimizes heterogeneity, stability, and beta diversity in mesic grassland. *Ecology and Evolution*, 8, 5008–5015.

Medeiros, L. P., Song, C. & Saavedra, S. (2021). Merging dynamical and structural indicators to measure resilience in multispecies systems. *Journal of Animal Ecology*, 90, 2027–2040.

Mittelbach, G. G. & McGill, B. J. (2019). *Community ecology.* Oxford University Press.

Mori, A. S., Isbell, F. & Seidl, R. (2018). β-diversity, community assembly, and ecosystem functioning. *Trends in ecology & evolution*, 33, 549–564.

Ohlmann, M., Miele, V., Dray, S., Chalmandrier, L., O'connor, L. & Thuiller, W. (2019). Diversity indices for ecological networks: a unifying framework using hill numbers. *Ecology Letters*, 22, 737–747.

Qian, H., Chen, S., Mao, L. & Ouyang, Z. (2013). Drivers of β-diversity along latitudinal gradients revisited. *Global Ecology and Biogeography*, 22, 659–670.

Raup, D. M. & Michelson, A. (1965). Theoretical morphology of the coiled shell. *Science*, 147, 1294–1295.

Ribando, J. M. (2006). Measuring solid angles beyond dimension three. *Discrete & Computational Geometry*, 36, 479–487.

Sablé-Meyer, M., Fagot, J., Caparos, S., van Kerkoerle, T., Amalric, M. & Dehaene, S. (2021). Sensitivity to geometric shape regularity in humans and baboons: A putative signature of human singularity. *Proceedings of the National Academy of Sciences*, 118.

Scheiner, S. M. (2019). A compilation of and typology for abundance-, phylogenetic-and functional-based diversity metrics. *BioRxiv*, 530782.

Scheiner, S. M., Kosman, E., Presley, S. J. & Willig, M. R. (2017). Decomposing functional diversity. *Methods in Ecology and Evolution*, 8, 809–820.

Scheiner, S. M. & Rey-Benayas, J. M. (1994). Global patterns of plant diversity. *Evolutionary ecology*, 8, 331–347.

Šizling, A., Keil, P., Tjørve, E., Tjørve, K., Žárskỳ, J. & Storch, D. (2022). Mathematically and biologically consistent framework for presence-absence pairwise indices. *bioRxiv*, 2021–07.

Song, C., Fukami, T. & Saavedra, S. (2021). Untangling the complexity of priority effects in multispecies communities. *Ecology Letters*, ele.13870.

Song, C., Rohr, R. P. & Saavedra, S. (2018). A guideline to study the feasibility domain of multi-trophic and changing ecological communities. *Journal of Theoretical Biology*, 450, 30–36.

Tilman, D., Isbell, F. & Cowles, J. M. (2014). Biodiversity and ecosystem functioning. *Annual review of ecology, evolution, and systematics*, 45.

van der Plas, F. (2019). Biodiversity and ecosystem functioning in naturally assembled communities. *Biological Reviews*, 94, 1220–1245.

Violle, C. & Jiang, L. (2009). Towards a trait-based quantification of species niche. *Journal of Plant Ecology*, 2, 87–93.

Whittaker, R. H. (1960). Vegetation of the siskiyou mountains, oregon and california. *Ecological Monographs*, 30, 279–338.

Whittaker, R. H. (1972). Evolution and measurement of species diversity. *Taxon*, 21, 213–251.

Xing, D. & He, F. (2021a). Analytical models for β-diversity and the power-law scaling of β-deviation. *Methods in Ecology and Evolution*, 12, 405–414.

Xing, D. & He, F. (2021b). Analytical models for β-diversity and the power-law scaling of β-deviation. *Methods in Ecology and Evolution*, 12, 405–414.