Appendix S1

Supplementary Material for

A geometric approach to beta diversity

in *Ecological Monographs*

Chuliang Song, Muyang Lu, Joseph R. Bennett, Benjamin Gilbert,
Marie-Josée Fortin, Andrew Gonzalez

S1 Examples of metacommunities with more species than communities

Consider metacommunity I with two sites (1 and 2) and three species (S_1 and S_2):

$$\begin{array}{r c} & \begin{array}{cc} S_1 & S_2 \end{array} \\ \begin{array}{r} \text{Community 1} \\ \text{Community 2} \end{array} & \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \end{array}$$

Now add a new common species, S_3, to create metacommunity II:

$$\begin{array}{r c} & \begin{array}{ccc} S_1 & S_2 & S_3 \end{array} \\ \begin{array}{r} \text{Community 1} \\ \text{Community 2} \end{array} & \begin{pmatrix} 1 & 0 & 1 \\ 0 & 1 & 1 \end{pmatrix} \end{array}$$

Now add many new common species to create metacommunity III:

$$\begin{array}{r c} & \begin{array}{ccccc} S_1 & S_2 & S_3 & \dots & S_k \end{array} \\ \begin{array}{r} \text{Community 1} \\ \text{Community 2} \end{array} & \begin{pmatrix} 1 & 0 & 1 & \dots & 1 \\ 0 & 1 & 1 & \dots & 1 \end{pmatrix} \end{array}$$

Let's compare how classic and geometric approaches differ in beta diversity values:

1. classic approach (Whittaker's ratio):
   - Metacommunity I: $\beta_I = 2/1 = 3$
   - Metacommunity II: $\beta_{II} = 3/2 = 1.5$
   - Metacommunity III: $\beta_{III} = \frac{k}{k-1}$ (approches 1 with higher k)
2. Geometric approach:
   - Metacommunity I: $\beta_I = 2 \times (1/2)^{1/2} = \sqrt{2} \approx 1.44$
   - Metacommunity II: $\beta_{II} = 2 \times (1)^{1/2} = 2$
   - Metacommunity III: $\beta_{III} = \frac{6\sqrt{k-2}}{k}$ (approches 0 with higher k)

The classic approach, exemplified by Whittaker's ratio, shows a decrease in beta diversity when a new common species is added to the metacommunity. This is expected as the classic approach prioritizes the distinctiveness of communities based on species turnover. This is true regardless of the metacommunity structure (as $\frac{\gamma}{\alpha}$ is always larger than $\frac{\gamma+1}{\alpha+1}$).

In contrast, our geometric approach exhibits a more nuanced response. It recognizes that adding novel combinations of species has the potential to increases spatial variations (metacommunity I vs II), while at the same time, it recognizes the potential homogenizing effect of overwhelming shared species on community composition (metacommunity II vs III).

S2 Examples of the definition of convex hull

Here we show how the abstract definition of the convex hull (3) operates with two concrete examples.

For the example of Metacommunity I in Figure 1, the convex hull P is not just these three points, but rather the entire triangle region bounded by these points $((1,0),(0,1))$ and the origin $(0,0)$. This is mathematically expressed as:

$$\begin{aligned} P = \{&\text{all points } (x,y) \text{ that can be written as:} && \text{(S1)} \\ &\lambda_1(1,0) + \lambda_2(0,1) + (1-\lambda_1-\lambda_2)(0,0)\} && \text{(S2)} \\ = \{&(x,y) = (\lambda_1, \lambda_2)\} && \text{(S3)} \end{aligned}$$

where $\lambda_1, \lambda_2 \geq 0$ and $\lambda_1 + \lambda_2 \leq 1$. This corresponds to exactly the triangle region we see in Figure 1C.

For the example of Metacommunity II in Figure 1, the convex hull P is not just these three points, but rather the entire triangle region bounded by these points $((1,0),(0,1),(1,1))$ and the origin $(0,0)$. This is mathematically expressed as:

$$\begin{aligned} P = \{&\text{all points } (x,y) \text{ that can be written as:} && \text{(S4)} \\ &\lambda_1(1,0) + \lambda_2(0,1) + \lambda_3(1,1) + (1-\lambda_1-\lambda_2-\lambda_3)(0,0)\} && \text{(S5)} \\ = \{&(x,y) = (\lambda_1+\lambda_3, \lambda_2+\lambda_3)\} && \text{(S6)} \end{aligned}$$

where $\lambda_1, \lambda_2, \lambda_3 \geq 0$ and $\lambda_1 + \lambda_2 + \lambda_3 \leq 1$. This corresponds to exactly the square region we see in Figure 1C.

S3 Scaling of beta diversity with unique community compositions

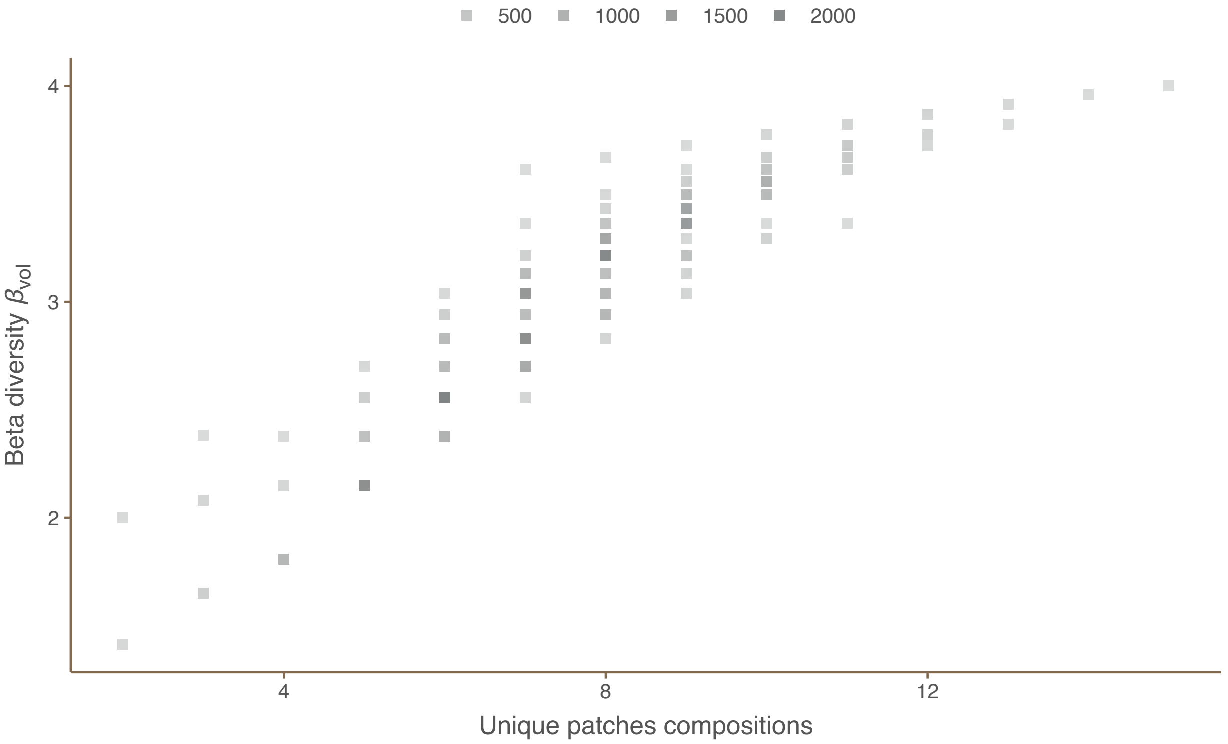

Figure S1: Same as Figure 2 except considering metacommunities with 4 species. The maximum number of unique patch compositions is $2^4 - 1 = 15$. The gray level corresponds to the number of communities with the same coordinates.

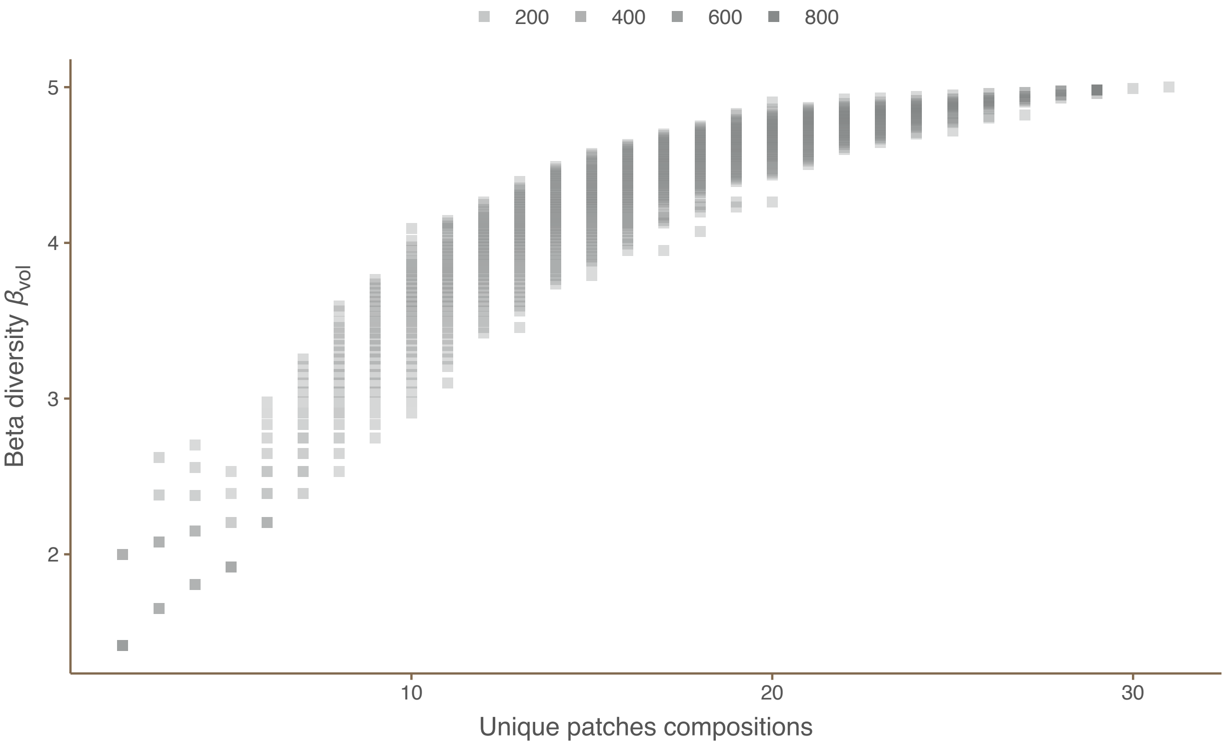

Figure S2: Same as Figure 2 except considering metacommunities with 5 species. The maximum number of unique patch compositions is $2^5 - 1 = 31$. The gray level corresponds to the number of communities with the same coordinates.

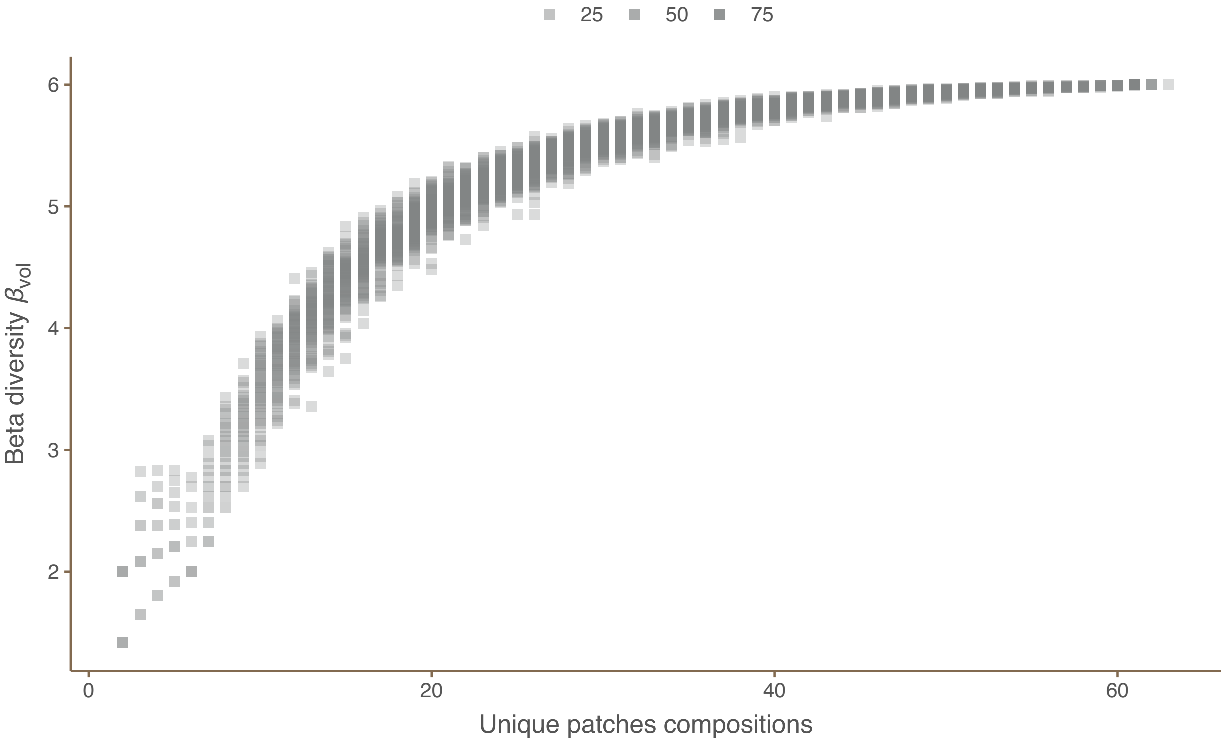

Figure S3: Same as Figure 2 except considering metacommunities with 6 species. The maximum number of unique patch compositions is $2^6 - 1 = 63$. The gray level corresponds to the number of communities with the same coordinates.

S4 Variation of geometric beta diversity with the same number of unique species combinations

Our geometric framework considers both the **number** of unique species combinations and the **richness** of those combinations. When the number of unique species combinations is equal, our approach generally favors metacommunities with higher species richness within those combinations (Figure 2 and **Section** S3). This is ecologically relevant because species-rich communities often have complex interactions and dynamics that are not captured by simply counting distinct species.

Our approach's emphasis on species richness becomes more apparent when considering metacommunities with more than two species (as is always the case in nature). For example, consider three sites and three communities. The metacommunity I has one distinct species at each site, and its metacommunity matrix reads as

$$\begin{array}{r c} & \begin{array}{ccc} A & B & C \end{array} \\ \begin{array}{r} \text{Community 1} \\ \text{Community 2} \\ \text{Community 3} \end{array} & \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \end{array}$$

The metacommunity II has two species at each site but with different combinations, and its metacommunity matrix reads as

$$\begin{array}{r c} & \begin{array}{ccc} A & B & C \end{array} \\ \begin{array}{r} \text{Community 1} \\ \text{Community 2} \\ \text{Community 3} \end{array} & \begin{pmatrix} 1 & 1 & 0 \\ 0 & 1 & 1 \\ 1 & 0 & 1 \end{pmatrix} \end{array}$$

Classic measures would favor Metacommunity I, as it has complete species turnover between sites. However, our geometric measure assigns a higher beta diversity to Metacommunity II ($\beta^{I}_{\text{vol}} = 1.65$ and $\beta^{II}_{\text{vol}} = 2.08$). This is because, while both metacommunities have three unique combinations, our approach recognizes that species-rich communities can contribute substantially to spatial variation, even when they share some species with other communities.

S5 Complementarity of our geometric measure to classic measure

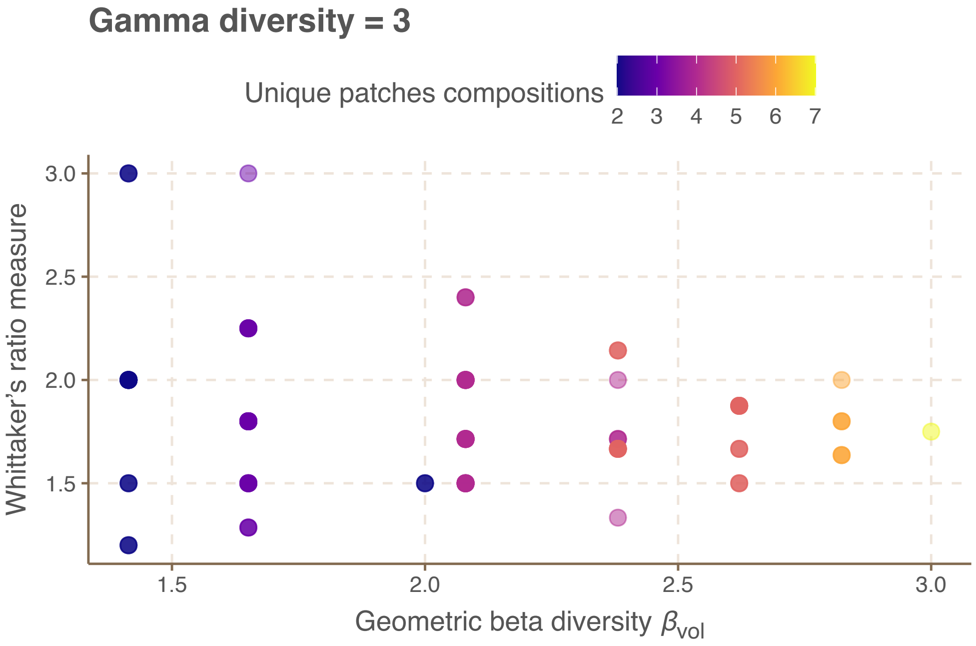

Figure S4: We study all possible configuration of metacommunities with 3 species. The horizontal axis shows the geometric beta diversity, while the vertical axis shows the classic Whittaker's ratio measure. Each point shows a different metacommunity, and its color indicates how many unique patch compositions it has.

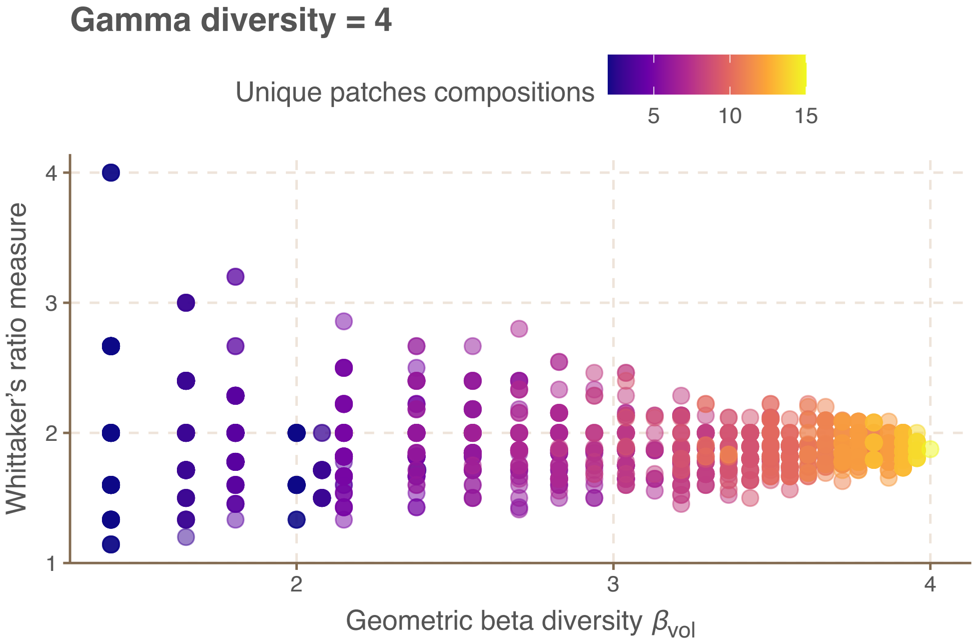

Figure S5: Same as Figure S4 except the metacommunities have 4 species.

S6 Maximize beta diversity with weighted embedding

Here we derive the maximum beta diversity with weighted embedding.

S6.1 Gamma diversity = 2

Following Eqn. (5), for a 2-species metacommunity, we have the weighted metacommunity matrix,

$$\begin{array}{cc} & \begin{array}{cc} A & B \end{array} \\ \begin{array}{c} \text{Community 1} \\ \text{Community 2} \\ \text{Community 3} \end{array} & \begin{pmatrix} 3p_1 & 0 \\ 0 & 3p_2 \\ 3(1-p_1-p_2) & 3(1-p_1-p_2) \end{pmatrix} \end{array} \tag{S7}$$

where p_1 represents the proportion of communities with only species A, while p_2 represents the proportion of communities with only species B.

The area vol of the embedded metacommunity is

$$\mathsf{vol} = \frac{1}{2} \times 3p_1 \times 3(1-p_1-p_2) + \frac{1}{2} \times 3p_2 \times 3(1-p_1-p_2) = \frac{9}{2}(p_1+p_2)(1-p_1-p_2) \tag{S8}$$

The area is maximized when $p_1 + p_2 = 1/2$.

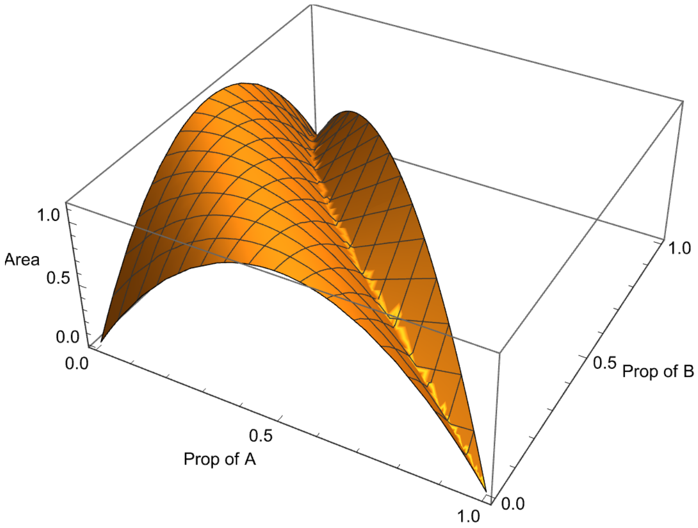

Figure S6: The two horizontal axes represent the proportion of patches with only species A or B, respectively. The vertical axis represents the area of the corresponding embedded metacommunities. We observe that the area is maximized when the proportion of patches containing both species A and B is half.

S6.2 Generalized variance and multivariate Bernoulli entropy

It is easy to show from equations S12 and S14 that the maximum beta diversity using the generalized variance and the multivariate Bernoulli entropy when gamma diversity = 2 are achieved when all community compositions have equal prevalence (see more details in S8.1

and S8.2). Here we will just give heuristic explanations about why this is the case. The generalized variance is a multivariate version of the variance. Its maximal value is achieved when the data show the highest dispersion. When gamma diversity = 2, this condition is achieved when the numbers of each unique composition are equal. For the multivariate Bernoulli entropy, maximum entropy is always achieved with a uniform distribution, which means that the prevalence of each community composition is the same. This differs from the conditions of maximum beta diversity using the geometric embedding.

S7 Nestedness-turnover decomposition

We discussed the interpretation of the nestedness-turnover decomposition using β_{vol} in the main text. Here, we discuss how to interpret the nestedness-turnover patterns in the other geometric measures β_{VAR} and β_{info} (Box 1 in the main text).

S7.1 Generalized variance

The nestedness and turnovers among communities are naturally captured by the correlation matrix X in equation S13: The distribution of correlation coefficients of a fully nested metacommunity are strictly positive (Fig. S7); the correlation distribution of a metacommunity with full turnover is strictly negative; and the correlation distribution of a metacommunity with randomly associated species is symmetric around 0.

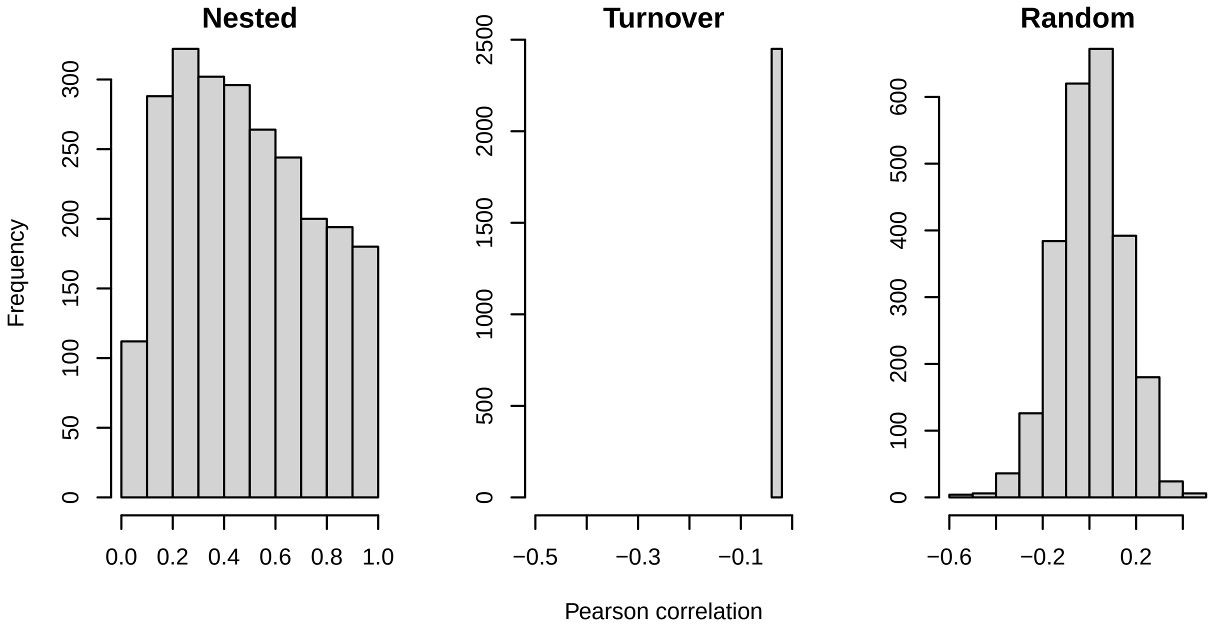

Figure S7: The histogram of the Pearson correlation coefficients of the correlation matrix X in equation S13 with 50 sites and 50 species.

Equation S13 states that both nestedness and turnover will reduce the beta diversity hypervolume. And their relative importance should be examined by the influence on the $\mathsf{det}(X)$.

S7.2 Information theory

In the entropy measure, the nestedness and turnovers among communities are reflected by the distribution of species richness in a metacommunity (Fig. S8). In a fully nested metacommunity, the species richness distribution is uniformly distributed; in a metacommuity with complete turnover, the species richness distribution is concentrated in the low richness part (at least less than 50% of the maximum richness); in a metacommunity with random species association, the species richness should be roughly normally distributed.

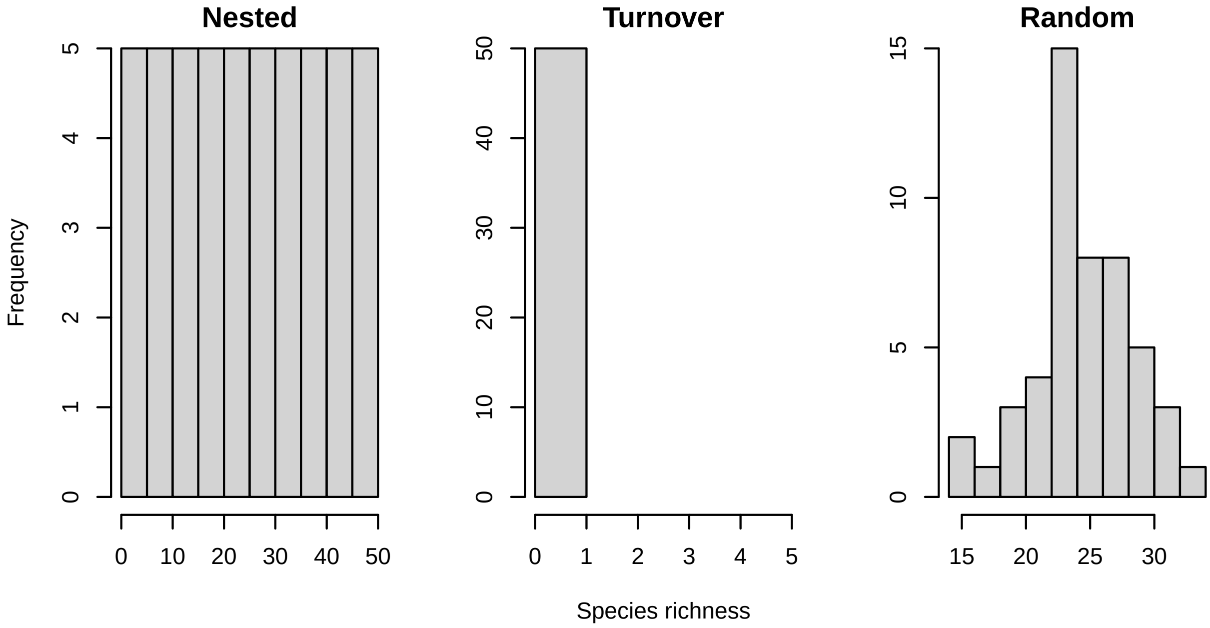

Figure S8: The histogram of the species richness of different communities with 50 sites and 50 species.

S8 Linking the geometric approach to classic formalisms

S8.1 Generalized variance

Building upon Legendre & De Cáceres (2013), it is straightforward to derive a hypervolume definition of beta diversity based on the covariance matrix of the metacommunity matrix. One natural extension of the concept of variance in the high-dimensional case is the generalized variance, which is calculated as the determinant of the covariance matrix (Lu *et al.*, 2021).

To unveil the insights in the generalized variance, we first look at metacommunities with only 2 species. The distribution in a 2-species metacommunity matrix can be characterized by a multivariate Bernoulli distribution in 2 dimensions. Let us denote the occurrence of the 2-species community as $X = [X_A, X_B]$, the proportion of communities with only species A as p_A, the proportion of communities with only species B as p_B, the proportion of communities with both species as p_{AB}. Then $p_A + p_B + p_{AB} = 1$ (excluding empty communities; Anderson *et al.*, 2006), and the probability mass function of a community composition is:

$$\mathbb{P}(X = x) = \mathbb{P}(X_A = x_A, X_B = x_B) = p_{AB}^{x_A x_B} p_A^{(x_A(1-x_B))} p_B^{(1-x_A)x_B}. \tag{S9}$$

Given equation S9, the covariance matrix of the bivariate Bernoulli distribution is:

$$\begin{aligned}\mathsf{VAR}(X) &= \mathbb{E}(XX^T) - \mathbb{E}(X)\mathbb{E}(X)^T \\ &= \begin{bmatrix} (1-p_B)p_B & -p_A p_B \\ -p_A p_B & (1-p_A)p_A \end{bmatrix}.\end{aligned} \tag{S10}$$

Note that the covariance between species A and species B is always ≤ 0 because of the exclusion of double-absence states, and this term is just the negative C-score for pairwise species association (Keil *et al.*, 2021).

Here, we define the hypervolume of the metacommunity as the generalized variance of this covariance matrix, which is the determinant $\mathsf{det}[\mathsf{VAR}(X)]$. The geometric interpretation is the hypervolume of the corresponding ellipse formed by the embedded metacommunity. Specifically, the volume of the ellipse is defined by a certain quantile (e.g. 95 %) of the normal distribution with the given covariance matrix (Lu *et al.*, 2021). Figure 8A-B illustrates two examples of metacommunity with β_{VAR}.

With this elliptic volume, we can define the corresponding measure of beta diversity β_{VAR} (the underscript highlights the use of variance):

$$\beta_{\mathsf{VAR}} = 2 \times (\mathsf{det}[\mathsf{VAR}(X)])^{(1/2)}. \tag{S11}$$

The beta diversity β_{VAR} of a 2-species metacommunity is the product of species-level variances and a cross-species association. The cross-species association is the ratio between the co-occurrence probability, p_{AB} and the expected co-occurrence probability given the marginal occurrence probabilities of A and B (which are respectively $1 - p_B$ and $1 - p_A$), while the products of species-level variances is a monotonic transformation of the dispersion-based beta diversity (Legendre & De Cáceres, 2013).

To interpret the terms more clearly, we take the log of the beta diversity β_{VAR} to make the terms additive:

$$\begin{aligned}\mathsf{log}(\beta_{\mathsf{VAR}}) &= \mathsf{log}(2) + \frac{1}{2}(\mathsf{log}(p_A) + \mathsf{log}(p_B) + \mathsf{log}(p_{AB})) \\ &= \underbrace{\mathsf{log}(2) + \frac{1}{2}}_{\gamma \text{ diversity}}(\underbrace{\mathsf{log}[p_A(1-p_A)] + \mathsf{log}[p_B(1-p_B)]}_{\text{species-level variances}} + \underbrace{\mathsf{log}[\frac{p_{AB}}{(1-p_A)(1-p_B)}]}_{\text{species association}}).\end{aligned} \tag{S12}$$

Therefore, the generalized variance definition of beta diversity naturally partitions the composition variation into a classic beta diversity measure and a spatial association component (Keil *et al.*, 2021).

Under the constraint that $p_A + p_B + p_{AB} = 1$, the generalized variance is maximized when $p_A = p_B = p_{AB}$. This is equivalent to saying that the generalized variance is maximized when each combination of community compositions are equally represented in the metacommunity matrix, which gives less weight to the mixed community compared to the geometric embedding approach.

We can generalize β_{VAR} (Eqn. S11) to higher dimensions (Lu *et al.*, 2021):

$$\beta_{\mathsf{VAR}} = d \times (\prod_{i=1}^{n} \sigma_i \, \mathsf{det}(X))^{1/d}, \tag{S13}$$

where X is the correlation matrix calculated from pairwise co-occurrence of all species in the metacommunity. Similar geometric partitioning as Eqn. S12 is straightforward.

Because of the dual representation of a metacommunity matrix as a multivariate function of either species or communities. Equations S9-S13 are equally applicable to a multivariate Bernoulli distribution with community occupancy as a random variable. Then equation S13 will partition the total beta diversity hypervolume into site-level contributions and a site association component, which is similar to the partitioning of dispersion-based beta diversity (Legendre & De Cáceres, 2013).

S8.2 Information theory

Information theory has long assisted biodiversity measurement (see a historical overview in Chao *et al.* 2014). Entropy measures, such as Shannon's and Renyi's entropy, have been traditionally used to quantify the unevenness of species abundance distributions within a local community (Jost, 2007). More recently, it has been shown that the relative entropy (Kullback–Leibler divergence) between the local and regional abundance profiles belongs to the general family of unevenness measures defined by the Hill numbers (Chao & Ricotta, 2019), which are directly linked to conventional beta diversity metrics such as the Jaccard and Sørensen indices (Chao *et al.*, 2014). Given the close connection between entropy and species diversity (Jost, 2007; Godsoe *et al.*, 2022), it is natural to extend the use of entropy to quantify the beta diversity hypervolume. Although not originally designed as a geometric measure, the entropy has a geometric interpretation as the effective size of the support of a random variable (Grendar, 2006). In the case of a multi-dimensional discrete probabilistic distribution, Shannon's entropy can be seen as a measure of the combinatorial volume of the distribution (Jankovic, 2009). The joint entropy of the multivariate Bernoulli distribution was proposed as a beta diversity metric long ago (Juhász-Nagy & Podani, 1983), but has been largely neglected until recently (Tsakalos *et al.*, 2022).

Following equation S9, the joint entropy $H(X)$ of the bivariate Bernoulli distribution is used as a measure of the hypervolume beta diversity β_{info}:

$$\beta_{\mathsf{info}} = H(X) = -[p_A \log(p_A) + p_B \log(p_B) + p_{AB} \log(p_{AB})]. \tag{S14}$$

The joint entropy is also maximized when $p_A = p_B = p_{AB}$ (a well-known property that the maximum entropy is always achieved with a uniform distribution). The contribution of species A and B to the total beta diversity are given by the entropies of the marginal Bernoulli distributions (Juhász-Nagy & Podani, 1983):

$$H(X_A) = -[p_A \log(p_A) + (1 - p_A) \log(1 - p_A)], \tag{S15}$$

$$H(X_B) = -[p_B \log(p_B) + (1 - p_B) \log(1 - p_B)]. \tag{S16}$$

The spatial association component is given by the mutual information between species A and B:

$$I(X) = H(X_A) + H(X_B) - H(X). \tag{S17}$$

When generalized to high dimensions, Eqn. S17 becomes:

$$I(X) = \sum_{i=1}^{\gamma} H(X_i) - H(X). \tag{S18}$$

Eqn. S14 becomes:

$$\beta_{\mathsf{info}} = H(X) = -\sum_{\alpha=1}^{2^\gamma - 1} \sum_{i \in \alpha} p_i^{(\alpha)} \log(p_i^{(\alpha)}), \tag{S19}$$

where γ denotes the species richness of the metacommunity, α denotes the species richness of a community, and $p_i^{(\alpha)}$ denotes the occurrence probability of community i whose species richness equals to α. The summation $-\sum_{i \in \alpha} p_i^{(\alpha)} \log(p_i^{(\alpha)})$ gives the contribution of communities with α diversity to the total beta diversity hypervolume.

It is also worthwhile to look at the simplifying scenario when all communities with the same α diversity have an equal occurrence probability $p^{(\alpha)}$. This assumes that the density dependent effects are equal among interacting species. Equation S19 then becomes:

$$\beta_{\mathsf{info}} = -\sum_{\alpha=1}^{\gamma} \binom{\gamma}{\alpha} p^{(\alpha)} \log(p^{(\alpha)}). \tag{S20}$$

If the multivariate Bernoulli distribution is treated as the distribution of community occupancy, then p_i in eqn. S20 is interpreted as the probability that a given set of communities will be occupied by a species. Again, by assuming that p_i is the same when the number of communities occupied by a species is fixed, Eqns. S19 and S20 can be respectively rewritten as:

$$\beta_{\mathsf{info}} = -\sum_{n=1}^{N} \sum_{i \in n} p_i^{(n)} \log(p_i^{(n)}), \tag{S21}$$

$$\beta_{\mathsf{info}} = -\sum_{n=1}^{N} \binom{N}{n} p^{(n)} \log(p^{(n)}), \tag{S22}$$

where N denotes the total number of communities, n denotes the number of communities occupied by a species, and $p_i^{(n)}$ denotes the probability that n selected communities will be occupied by species i. Empirically, $p_i^{(n)}$ is estimated by the relative frequency of species that occupies the n selected communities in the metacommunity. Therefore, Eqn. S21 partitions total beta diversity hypervolume into contributions of species with different prevalence in the metacommunity.

$p_i^{(n)}$ has a close relationship with zeta diversity ζ_n, which measures the mean number of species shared by n communities (Hui & McGeoch, 2014). The sum of all $p_i^{(n)}$ that occupy the same number of communities is a linear transformation of standardized ζ_n:

$$\sum_{i \in n} p_i^{(n)} = \binom{N}{n} \sum_{w=1}^{N-n+1} (-1)^{(w+1)} \binom{N-n}{w-1} \zeta_{n+w-1}. \tag{S23}$$

When $p_i^{(n)}$ are equal for all different site compositions, the zeta diversity of order n is:

$$\zeta_n = \frac{2^{\gamma-n}}{2^{\gamma}-1}, \tag{S24}$$

which justifies the use of an exponential function of zeta diversity to detect stochastic assembly (Hui & McGeoch, 2014).

S9 Dealing with continuous data

In addition to the presence/absence data typically analyzed in metacommunities, continuous data, such as species abundances, are also commonly encountered. A significant challenge with continuous data is the "point-in-the-middle" phenomenon in the convex hull approach. This phenomenon occurs when the embedded representation of a community lies within the convex hull defined by the other communities. This happens because the convex hull is only sensitive to the peripheral points, ignoring internal variations. Notably, this issue does not occur with presence/absence data.

To illustrate, consider three sites characterized by different compositions of species A and B, with abundance vectors of (10, 0), (0, 10), and (10, 10). When adding a fourth site with an intermediate abundance vector (5, 5), this point is positioned within the interior of the convex hull of the original three sites. Consequently, the convex hull method fails to fully capture the nuances of community differences because it only accounts for the outermost points, disregarding internal diversity patterns.

To more effectively deal with continuous data, we recommend employing a generalized variance approach. This approach better captures the impact of interior points by directly accounting for changes in multivariate variability, rather than merely relying on spatial positioning within a convex hull. We demonstrate the advantage of this approach below using the same example.

For the first metacommunity (sites with abundance vectors (10, 0), (0, 10), (10, 10), and (0, 0)):

$$\mathbf{Z} = \underbrace{\begin{pmatrix} 10 & 0 \\ 0 & 10 \\ 10 & 10 \\ 0 & 0 \end{pmatrix}}_{\text{Metacommunity matrix}} \Rightarrow \mathrm{VAR}(\mathbf{Z}) = \underbrace{\begin{pmatrix} 33.3 & 0 \\ 0 & 33.3 \end{pmatrix}}_{\text{Covariance matrix}} \Rightarrow \beta_{\mathrm{VAR}} = 2 \times (\mathsf{det}(\mathrm{VAR}(\mathbf{Z}))) = 66.67 \tag{S25}$$

Upon adding the interior point (5, 5) to create a new metacommunity:

$$\mathbf{Z} = \underbrace{\begin{pmatrix} 10 & 0 \\ 0 & 10 \\ 10 & 10 \\ 5 & 5 \\ 0 & 0 \end{pmatrix}}_{\text{Metacommunity matrix}} \Rightarrow \mathrm{VAR}(\mathbf{Z}) = \underbrace{\begin{pmatrix} 25 & 0 \\ 0 & 25 \end{pmatrix}}_{\text{Covariance matrix}} \Rightarrow \beta_{\mathrm{VAR}} = 2 \times (\mathsf{det}(\mathrm{VAR}(\mathbf{Z}))) = 50 \tag{S26}$$

As demonstrated, the generalized variance approach appropriately captures the effect of adding the intermediate point, resulting in a decrease in beta diversity from 66.67 to 50. This result aligns with ecological intuition, as the addition of an intermediate point contributes differently to beta diversity compared to extreme values.

S10 Computation of hypervolume

We have provided an R package `betavolume` (`https://github.com/clsong/betavolume`) that provides a user-friendly interface to compute beta diversity using the methods discussed below.

S10.1 Geometric embedding

Hypervolume of geometric shapes in high dimension is notoriously difficult to estimate. Fortunately, we do not need to compute the hypervolume of *arbitrary* geometric shape (e.g., this is typically required for fundamental niches). Instead, given the geometric embedding we have here, we are tasked with a specific geometric shape: *convex polytope* (Eqn. 3 in the main text). Convex polytope is the most fundamental and well-studied shape in computational geometry (Preparata & Shamos, 2012). Nonetheless, ruled by the curse of dimensionality, its hypervolume estimation is still a hard computational problem (Dyer & Frieze, 1988). We summarize some available computational methods, which offered a range of trade-offs between accuracy and efficiency.

A deterministic method is through the Delaunay triangulation. R package `geometry` (Habel *et al.*, 2019) provides an interface to the Matlab library `qhull` (Barber *et al.*, 1996). This method can provide almost exact estimates of the hypervolume, and it is efficient for relative low dimensions (< 20). However, its run-time grows exponentially with the dimension, which makes it unpractical when the metacommunity has many communities.

A probabilistic method is more preferred for higher dimensions (especially > 20). In brief, a statistical method known as Multiphase Monte Carlo allows efficient geometric random walks. This has led to many practical probabilistic methods, such as Sequence of Balls (Emiris & Fisikopoulos, 2018), Cooling Gaussians (Cousins & Vempala, 2016), and Cooling convex Bodies (Chalkis *et al.*, 2019). R package `volesti` (Fisikopoulos *et al.*, 2021) provides an interface to the C++ library `VolEsti` (Chalkis & Fisikopoulos, 2020), which implements the aforementioned methods. These methods are possible even in hundreds of dimensions. Notably, its computation time is not ignorable, which makes it unpractical when we deal with many metacommunities.

For even faster computations (at the expense of accuracy), a heuristic method based on multivariate normal distributions is available (Lu *et al.*, 2021). In brief, it first estimates the covariance matrix from community data, then the determinant of the covariance matrix is used to calculate the hypervolume of the high-dimensional ellipsoid.

A caveat though is that a n-dimensional cube would have a estimated hypervolume of 4^{-n}. To make the hypervolume comparable across dimensions, we re-scale the hypervolume by 4^n (which is a common practice, e.g., see Grilli *et al.* 2017). As a side note, we do not use the widely used R package `hypervolume` (Blonder & Harris, 2019), because it is not best suited to provide accurate nor fast estimation of convex polytopes (Mammola, 2019).

S10.2 Generalized variance

For the numerical implementation of equation S13 in high dimensions, we recommend first calculating the PCA of the correlation matrix R and then take the product of eigenvalues

λ_j that capture at least 99% of the total variance for dim ≥ 4, or when singular matrix is encountered:

$$\prod_{i=1}^{n} \sigma_i \prod_{j=1}^{k} \lambda_j \tag{S27}$$

An advantage of the PCA implementation is that the determinant of the correlation matrix can be further partitioned into individual eigenvalues, which quantifies the contributions of associations among species groups on the PCA axes.

S10.3 Information measure

In calculating the high dimensional information measure (equation S19), if it is of more interest to only investigate the richness-dependent effect of different communities (such as when species are believed to be functionally equivalent or similar), implementing equation S20 can reduce the computational time by only counting the number of communities with different richness.

S10.4 Computational complexity

We benchmark the computational time to calculate the geometric beta diversity on a computer with the generalized variance approach. Timings were carried out on a computer with an Apple M1 processor.

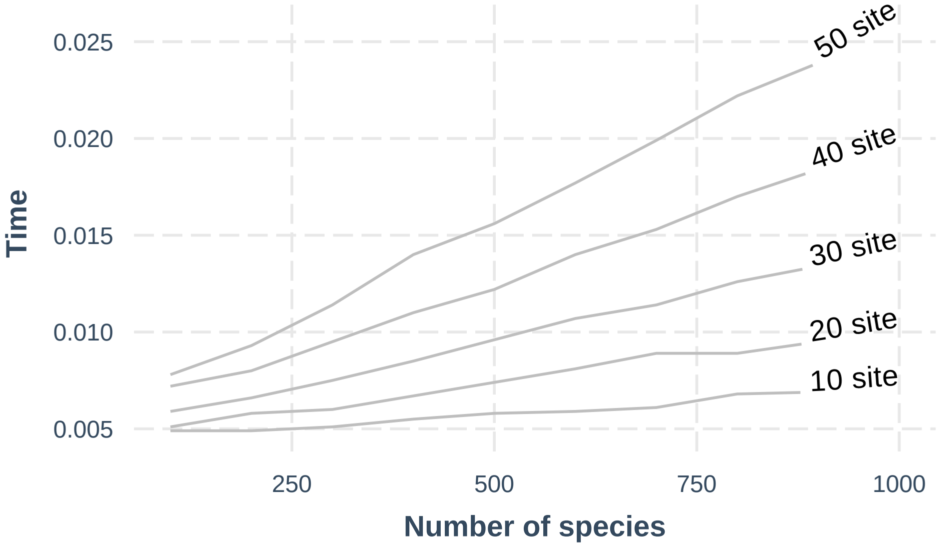

Figure S9: The horizontal axis shows the number of species (gamma diversity) in the metacommunity. The lines show different numbers of sites. The vertical axis shows the time (unit is second) to compute the geometric beta diversity of corresponding metacommunities with the generalized variance approach.

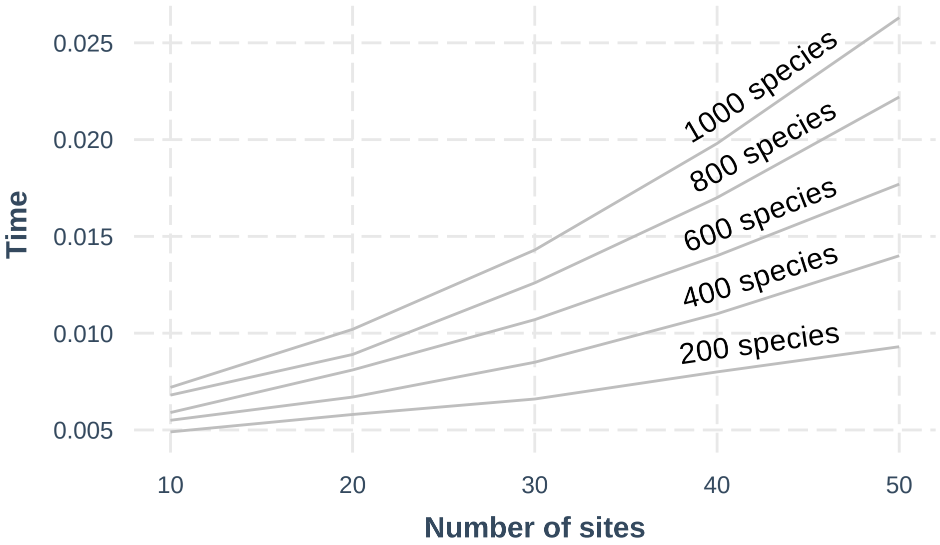

Figure S10: The horizontal axis shows the numbers of sites in the metacommunity. The lines show different numbers of species (gamma diversity). The vertical axis shows the time (unit is second) to compute the geometric beta diversity of corresponding metacommunities with the generalized variance approach.

References

Anderson, M. J., Ellingsen, K. E. & McArdle, B. H. (2006). Multivariate dispersion as a measure of beta diversity. *Ecology Letters*, 9, 683–693.

Barber, C. B., Dobkin, D. P. & Huhdanpaa, H. (1996). The quickhull algorithm for convex hulls. *ACM Transactions on Mathematical Software (TOMS)*, 22, 469–483.

Blonder, B. & Harris, D. J. (2019). *hypervolume: High Dimensional Geometry and Set Operations Using Kernel Density Estimation, Support Vector Machines, and Convex Hulls.*

Chalkis, A., Emiris, I. Z. & Fisikopoulos, V. (2019). A practical algorithm for volume estimation based on billiard trajectories and simulated annealing. *arXiv preprint arXiv:1905.05494*.

Chalkis, A. & Fisikopoulos, V. (2020). volesti: Volume approximation and sampling for convex polytopes in r. *ArXiv*, abs/2007.01578.

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.

Chao, A. & Ricotta, C. (2019). Quantifying evenness and linking it to diversity, beta diversity, and similarity. *Ecology*, 100, 1–15.

Cousins, B. & Vempala, S. (2016). A practical volume algorithm. *Mathematical Programming Computation*, 8, 133–160.

Dyer, M. E. & Frieze, A. M. (1988). On the complexity of computing the volume of a polyhedron. *SIAM Journal on Computing*, 17, 967–974.

Emiris, I. Z. & Fisikopoulos, V. (2018). Practical polytope volume approximation. *ACM Transactions on Mathematical Software (TOMS)*, 44, 1–21.

Fisikopoulos, V., Chalkis, A. & contributors in file inst/AUTHORS (2021). *volesti: Volume Approximation and Sampling of Convex Polytopes.*

Godsoe, W., Bellingham, P. J. & Moltchanova, E. (2022). Disentangling Niche Theory and Beta Diversity Change. *The American Naturalist*, 199, 510–522.

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.

Habel, K., Grasman, R., Gramacy, R. B., Mozharovskyi, P. & Sterratt, D. C. (2019). *geometry: Mesh Generation and Surface Tessellation.*

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.

Jankovic, M. V. (2009). Geometrical Interpretation of Shannon's Entropy Based on the Born Rule. *arXiv:0909.4995 [physics].*

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.

Keil, P., Wiegand, T., Tóth, A. B., McGlinn, D. J. & Chase, J. M. (2021). Measurement and analysis of interspecific spatial associations as a facet of biodiversity. *Ecological Monographs*, 91, 1–22.

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

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.

Preparata, F. P. & Shamos, M. I. (2012). *Computational geometry: an introduction*. Springer Science & Business Media.

Tsakalos, J. L., Chelli, S., Campetella, G., Canullo, R., Simonetti, E. & Bartha, S. (2022). comspat: an R package to analyze within-community spatial organization using species combinations. *Ecography*, n/a, e06216.