Supplementary Information for Bimler et al. 2024

S1 Supplementary Information 1: Methods - Data

S1.1 Study system and community characteristics

York gum-jam woodlands range over a 1000-km gradient spanning the Mediterranean and semi-arid climates of the South West Western Australia wheatbelt. South West Australia is one of 36 internationally recognised biodiversity hotspots (Conservation International 2023) with over 5500 vascular plant species, of which over half are not found anywhere else. The landscape is highly fragmented with the vast majority of native vegetation cleared for agriculture, and the woodlands survive predominantly as isolated fragments within a mosaic of canola and wheat fields. They are defined by a variable but generally open canopy composed almost entirely of two tree species: *Eucalyptus loxophleba* and *Acacia acuminata*. The woodlands understory is dominated by dense winter annual plant communities which typically support a diverse mix of native and exotic forbs and exotic annual grasses which are well-adapted to the nutrient poor soils (Figure S1.1).

Figure S1.1: Photos of the York gum-jam woodland wildflower study system. The York gum and jam trees form an open and sparse canopy at West Perenjori reserve (left), under which a diverse wildflower community emerges every year (top right). The bottom right photo shows one of the 50×50 cm plots used to collect the data for this study. Photo credits: Malyon D. Bimler, Victoria Reynolds, Trace E. Martyn.

West Perenjori Reserve (29^{o}28'01.3"S 116^{o}12'21.6"E) has served as a study site for community ecology and biodiversity research for over a decade, where 114 species of herbaceous plants have been cumulatively identified over the years (85 native and 29 exotic species) and herbaceous plant alpha diversity can be as high as 29 species per 0.09m^2 (J. M. Dwyer and M. M Mayfield, *unpublished data*). Extensive evidence from observational studies in Perenjori and other reserves of York gum-jam woodland indicate that soil phosphorus concentration, water availability, and tree canopy cover all structure the plant diversity in the understory (Dwyer et al. 2015; Wainwright et al. 2017) and, importantly, have been associated with non-linear variation in plant–plant interaction strength and magnitude (Bimler et al. 2018). Soils in these woodlands are relatively rich in nitrogen but have very low levels of plant-available phosphorus, though agricultural runoff from the surrounding fields leads to patchy phosphorus enrichment in the reserve, resulting in turnover in soil phosphorus evident at the 5–15m scale. Water availability on the other hand varies along a regional gradient evident at the 100-km scale (Dwyer et al. 2015), transitioning from a mesic mediterranean climate in the west of Western Australia to semi-arid climates in the east, with rainfall generally declining with distance from the coast. Lastly, canopy cover varies locally at small scales (<1m), matching the size of our study plots. We decided to bin percent canopy cover into three different categories (open, intermediate and shady) as that is the lowest grouping number which is still sufficient to identify potential non-linearity in trends over an environmental gradient. Due to the predominantly open nature of the overhead canopy, percent canopy cover was unevenly distributed between our 100 plots with the majority of plots experiencing very low shade ($< 20\%$ canopy cover). We thus decided to bin canopy cover in such a way as to have similar number of plots in each shade category, rather than an even range of percent canopy cover in each category.

	Environmental Variables:						
	Canopy cover (%)			**Soil phosphorus**		**Soil water (%)**	
	range	mean	sd.	mean	sd.	mean	sd.
Open	0 to 8	4.2	2.2	11.9	2.7	17.4	2.2
Intermediate	8 to 18	12.8	2.5	13.2	4.5	18.2	2.8
Shady	18 to 40	26	6	15.2	7.9	19	3.8
	Leaf litter cover (%)			**Wood debris cover (%)**		**Bare ground (%)**	
	mean	sd.		mean	sd.	mean	sd.
Open	3.6	7.7		5.8	7.2	11.7	12
Intermediate	5.5	8.7		12	11	13.7	20
Shady	11.8	17.6		11.4	12	21	24

	Community structure:		
	Mean plot richness	**Mean Shannon evenness**	**Mean species turnover**
Open	11.3	0.56	0.60
Intermediate	11.4	0.61	0.62
Shady	12.7	0.66	0.67

Table S1.1: Plot characteristics (abiotic, biotic) for the low, intermediate and high shade microenvironments. Means and standard deviations are given across plots. Average soil phosphorus concentration (mg/kg) was measured via a standard Colwell extraction on mixed soil samples from each plot. Soil water is reported as water holding capacity. Community structure was measured from all plots, thinned and unthinned.

The three plant interaction networks thus lie along a gradient of increasing shade. Increasing shade was weakly positively correlated with soil phosphorus ($r = 0.34$) and water availability ($r = 0.28$), as well as the amount of litter ($r = 0.27$), bare ground ($r = 0.26$) and woody debris ($r = 0.26$) covering each plot. We tested for an effect of percent canopy cover (as a continuous variable) and thinning treatment (as a three-level categorical variable) on plot-level species richness and Shannon evenness with generalised linear models applied with the *glm()* function. We used a Poisson error distribution for species richness and a binomial error distribution (and logit link) for Shannon evenness. We first ran both models with an interaction between canopy cover and thinning treatment but the interaction coefficients were not deemed as significant in either case so the interaction was removed ($p > 0.1, df = 94$). Percent canopy cover and thinning treatment had no significant effect on either plot-level species richness or plot-level Shannon evenness ($p > 0.05, df = 96$). β diversity (Sorensen index) was also similar between sites, though assemblages in the open and shady plots were slightly more similar to each other ($\beta_{open\&shady} = 0.73$) than to the intermediate plots ($\beta_{open\&inter} = 0.80$ and $\beta_{inter\&shady} = 0.80$). Plant density was on average higher in the open plots (mean densities of 353, 288 and 217 individuals across all plots in the open, intermediate and shady categories, respectively) but this pattern was mainly driven by one species (*Goodenia rosea*) which forms large, dense patches in the open sun (Table S1.2).

	Open	Inter	Shady
ARCA	**464**	**201**	**392**
CAHI	*NA*	39	*NA*
GITE	*NA*	25	*NA*
GOBE	*NA*	27	44
GOPU	152	*NA*	*NA*
HAOD	*NA*	*NA*	23
HYGL	**62**	**43**	**37**
HYPO	*NA*	*NA*	34
MEDI	32	*NA*	69
PEAI	**277**	**355**	**439**
PEDU	*NA*	*NA*	27
PLDE	**99**	**106**	**21**
POCA	**175**	**223**	**152**
POLE	*NA*	34	50
PTGA	**75**	**124**	**82**
STPA	**33**	**47**	**31**
TRCY	53	*NA*	37
TROR	*NA*	*NA*	40
VECY	*NA*	76	*NA*
VERO	**729**	**375**	**79**
WAAC	**24**	**42**	**74**
Total	**2170**	**1717**	**1631**

Table S1.2: Number of individual observations for each focal species in each shade category. We use *NA* to indicate when a species had fewer than 20 observations in that specific shade category and was hence not selected as a focal for that category. Species which were selected as focal in every shade category are in bold.

S1.2 Estimating seed production and accounting for missing seeds

While counting seeds, we made note of the number of seed heads (for *Asteraceae* and other plant families) or seed pods (specifically for species in *Goodeniaceae*) that were open and/or missing seeds for each individual plant. We also counted the number of buds and flowers remaining on each plant. We were then able to estimate the number of seeds that we did not collect in the field (either because they were already dispersed, or because we had to collect the plant before it was done being pollinated and making seeds) for each individual and include that in our estimate of reproductive output. For all species except for those in *Goodeniaceae*, we did this by calculating the mean number of seeds per seed head for those seed heads that were neither missing nor open within the plot. We then multiplied this plot-level, species-specific mean number of seeds per seed head by the number of buds/flowers remaining as well as the number of heads with missing seeds for each plant individual. If there were no other individuals within the plot with which to calculate this plot-level, species mean, then we used an overall site-level species-specific number of seeds per seed head. For example, consider an *Arctotheca calendula* individual that had 2 complete seed heads with 30 and 15 seeds, 1 seed head that had already dispersed but still had 5 undispersed seeds, and 3 flowers. In addition, we know that that plot that the average number of seeds per seed head is 23. Our calculation would be $30 + 15 + 5 + 1 \times 23 + 3 \times 23 = 142$ seeds. We acknowledge that in some cases, we might be overestimating the seed number as seed heads with missing seeds are counted as completely empty. In our seed count data, our notes indicate that most of these partially filled seed heads were at least 50% empty and more commonly at least 80% empty.

The process for estimating missing seeds differed slightly for *Goodeniaceae* species (*G. berardiana*, *G. pusilliflora*, *G. cycnopotamica*, and *G. rosea*). For individuals from these species, we counted seeds on a per stem basis due to time constraints. We averaged the number of seeds per stem for each *Goodeniaceae* species and multiplied that number by any missing stems for each plant. Unfortunately, we did not count the average number of seeds per pods for estimating seeds from buds, flowers and pods with missing seeds, nor did we count the number of flowers or pods per stem. In order to still account for some of the missing seeds, we instead counted presence/absence of flowers/buds and of pods with missing seeds. Both of these variables were then multiplied by the species-specific average number of seeds per stem, such that the presence of flowers was equivalent to one missing stem, and the presence of pods with missing seed was also equivalent to one missing stem. For example, an individual *Goodeniaceae* plant with 10 counted seeds, 3 missing stems, 2 flowers, and 1 pod with missing seeds, where the average number of seeds per stem for that species is 5, received a total estimated seed count of $10 + 3 \times 5 + 2 \times 5 + 1 \times 5 = 40$ seeds. We acknowledge that these assumptions are likely to overestimate missing seeds for some small plants and underestimate missing seeds for large plants of those four *Goodeniaceae* species.

S2 Supplementary Information 2: Methods - Model & Analysis

S2.1 A summary of the joint model framework used to estimate interaction effects

The joint model framework (Bimler et al. 2023) was designed to estimate the strength and magnitude of pairwise interactions in diverse systems by using measures of species performance (e.g. seed production, growth rate, biomass) in the presence of varying densities of their potential interaction partners. Observations of individual organism performance (or a proxy for performance) are recorded, alongside the identity and abundance of neighbouring individuals. The joint model framework then regresses the performance of a species against the density and identity of other interacting species, such that increases or decreases in a species' performance are attributed to the changing densities of its interaction partners. Note that interacting species can include members of the focal species itself, thus capturing intraspecific interactions. Importantly, the framework places no restrictions on the sign of interactions which means they can be harmful (competitive) or beneficial (facilitative) to the focal species.

A key issue with estimating interactions in diverse systems is that as the number of species S increases, the number of potential pairwise interactions subsequently increases as S^2, rendering most estimation methods both data intensive and computationally complex. Moreover, the fact that data sampling is limited in time and space implies that observational datasets are often incomplete in the sense that some species may not be observed to interact with one another or when observed to interact do so at insufficiently variable densities, rendering many such interactions statistically unidentifiable. This joint model framework allows us to nevertheless make fuller use of such incomplete datasets by estimating pairwise interactions through one of two different models depending on the data available for each interaction: a neighbour density-dependent model (NDDM) for interactions which are statistically identifiable, and a response-impact model (RIM) for unidentifiable interactions. They are implemented as generalised linear models. Both models (and thus all interaction parameters) are fitted simultaneously by requiring both models to contribute to the overall joint model log likelihood. This allows both models to provide good fits to the data, but also for interaction parameters estimated by the RIM to 'adjust' around the values inferred by the NDDM.

1. Neighbour density-dependent model (NDDM): The NDDM assumes that observations of performance p_i for a focal species i are a function $f()$ of intrinsic performance γ_i (performance in the absence of interactions) and the abundance of neighbouring interaction partners N_j:

$$f(p_i) = \gamma_i - \sum_{j=1}^{S} \beta_{ij} N_j \tag{S2.1}$$

In the NDDM, the effect of each interacting species j on i is captured by a unique, identifiable interaction parameter β_{ij}.

2. Response-impact model (RIM): When an interaction parameter is unidentifiable because of insufficiently variable neighbour densities, we can instead estimate it with the RIM rather than assuming it is negligible or equal to zero. Note that the RIM is only one of many potential solutions to measuring unidentifiable interactions. In contrast to the NDDM, the RIM assumes that a species i will typically have a singular impact e_i on and a singular response r_i to neighbours independent of neighbour identity. Each

pairwise interaction parameter is then given by the product of a focal species i's' response parameter and an interacting species j's' impact parameter. For performance data fully predicted by the RIM, it would be mathematically given by:

$$f(p_i) = \gamma_i - r_i \sum_{j=1}^{S} e_j N_j \tag{S2.2}$$

3. Joint interaction effects: The interaction parameters returned by fits based on either Equations S2.1 and S2.2 can be used to construct a community interaction matrix of size $S \times T$ where S is the number of focal species i and T the number of neighbour species j across all focals. In this community interaction matrix, the effect of neighbour species j on focal species i corresponds to the value in the i'th row and the j'th column. In the joint model, the elements of this matrix, which we refer to as α_{ij}'s, take the value of β_{ij} when the effect of j on i is identifiable and the value of $r_i e_j$ when the effect of j on i is not identifiable. The diagonal elements α_{ii} of this matrix correspond to intraspecific interaction effects. Details of how interaction identifiability is determined and how both models contribute to the log-likelihood are given in the main text and Supplementary Information of Bimler et al. 2023.

4. Transforming interaction effects into per-capita interaction strengths: For annual plant dynamics, interaction effects can be converted to interaction strengths using an individual fecundity model for annual plants with a seed bank (Levine and HilleRisLambers 2009; Mayfield and Stouffer 2017; Bimler et al. 2018). This model describes the rate of change in a focal species' *i* abundance of seeds in a seed bank from one year to the next. This model includes parameters for germination rate g_i, seed survival rate s_i and intrinsic seed production λ_i (seeds produced per germinated individual of i in the absence of neighbours), as well as interactions with individuals of the same species i or of neighbouring species j. We use this model to convert the α's returned by the joint modelling framework above into per-capita measures α'' that are directly proportional to the abundance of neighbouring seeds (Godoy and Levine 2014; Bimler et al. 2018).

We thus define α'':

$$\alpha''_{ij} = \frac{g_j \alpha_{ij}}{ln(\eta_i)} \tag{S2.3}$$

where g_j is the germination rate of the interaction partner j and $\eta_i = \frac{\lambda_i g_i}{\theta_i}$, with $\theta_i = 1 - (1 - g_i)(s_i)$. Further details are provided in Bimler et al. (2023).

Three species in the shady environment (*Trachymene ornata*, *Trachymene cyanopetala* and *Medicago sp.*) and one species in the open environment (*Trachymene cyanopetala*) had growth rates and thus scaling effects ($\frac{ln(\eta_i)}{g_i}$) below 1. Scaling rates below 1 are mathematically problematic as they can overinflate interaction strengths, and also suggest that the species in question is not able to sustain its population from year to year in monoculture. It is probable, however, that our environmental categories did not capture the full spectrum of environmental conditions which allow these particular species to grow and persist. Moreover, facilitative interactions may also allow these species to persist when growing with other species. These species are all known to consistently persist in the system over long time-scales but appear to be highly restricted to certain micro-environments and likely have complex germination requirements requiring microbial symbionts (M.M. Mayfield, T.L. Staples, personnal observations). For these reasons, we felt comfortable assigning them the next lowest growth rate in that environmental category which was above 1 instead (1.736 for the open environment, and 1.343 in the shaded environment).

S2.2 Model Convergence

Running complex models in Stan often requires some amount of user input to achieve optimal convergence. We first ran all models with 4 chains of 10000 iterations each, discarding the first 8000 iterations and setting *adapt_delta* $= 0.99$ and *max_treedepth* $= 25$. We evaluated convergence through visual inspection of traceplots and $\hat{R}$ for those relevant parameters which are not expected to sign-switch, as laid out in Bimler et al. (2023). The model on intermediate shade category data converged well with those initial settings. We increased *max_treedepth* to 35 and lowered *adapt_delta* to 0.9 when running the model on the shady category data; this improved convergence and all relevant parameters had $\hat{R} < 1.01$. Unfortunately, varying *adapt_delta* and *max_treedepth* and increasing the number of iterations was not sufficient to reach full convergence for the open shade category dataset. For that dataset, we achieved the best results by setting *adapt_delta* to 0.9 and *max_treedepth* to 15, with those settings 95.7% of all relevant parameters reached a $\hat{R}$ of 1.1 or below. We also ran the model on data from all three shade categories pooled together in order to make Figure S3.1. Optimal convergence for the full dataset was reached by setting *max_treedepth* $= 15$ and *adapt_delta* $= 0.99$, 91.5% of relevant parameters reached a $\hat{R}$ below 1.2, with a maximum $\hat{R}$ of 3.7.

As outlined in Bimler et al. (2023), the lack of convergence associated with no other warnings (and good convergence on simulated data) is likely caused by correlations arising between the latent variables used to estimate unobserved interaction effects in the joint model framework. In summary, shifts or uncertainty in one latent parameter can sometimes create shifts or uncertainty in the other latent parameters, causing diffferent chains to get "stuck" in slightly different local optima and leading to multi-model posteriors for those parameters with high $\hat{R}$ values. Indeed, diagnostic plots for those parameters with difficult convergence indicate that chains with different optima (and thus different parameter values) still had very close overlapping associated log posteriors, even when $\hat{R}$ was at its highest. In other words, the model and chains were still behaving as expected in terms of providing the best fit to the data, and that the different modes of those non-convergent parameters estimates gave only minor differences in likelihood.

$\hat{R}$ values for all relevant parameters in the open category data ranged from 0.996 to 2.786, and high $\hat{R}$'s' were associated with low effective sample size. $\hat{R}$ for the parameters related to intrinsic fitness (seed production in the absence of competitors) had a median value of 1.004 and a maximum value of 1.164, which indicated that chains were still closely overlapping and multi-modality was minimal for those parameters even when optimal convergence was not reached (Figure S2.1 top row). Multimodality became more evident for parameters with $\hat{R} > 1.2$ (Figure S2.1 bottom row), though this only accounted for 10 parameters out of a total 396 (2.5%). For those 10 parameters, the posterior distributions became increasingly bimodal as $\hat{R}$ increased, often but not always centered on 0, and the standard error of the mean was larger than the mean for 3 of those parameters. Given overall good model performance, the small number of parameters which showed bimodality, and that most of those bimodal posteriors were interaction effects with non-focal neighbours and thus not included in the interaction strength network analyses, we felt comfortable with using the parameter estimates returned by the model for the open category data.

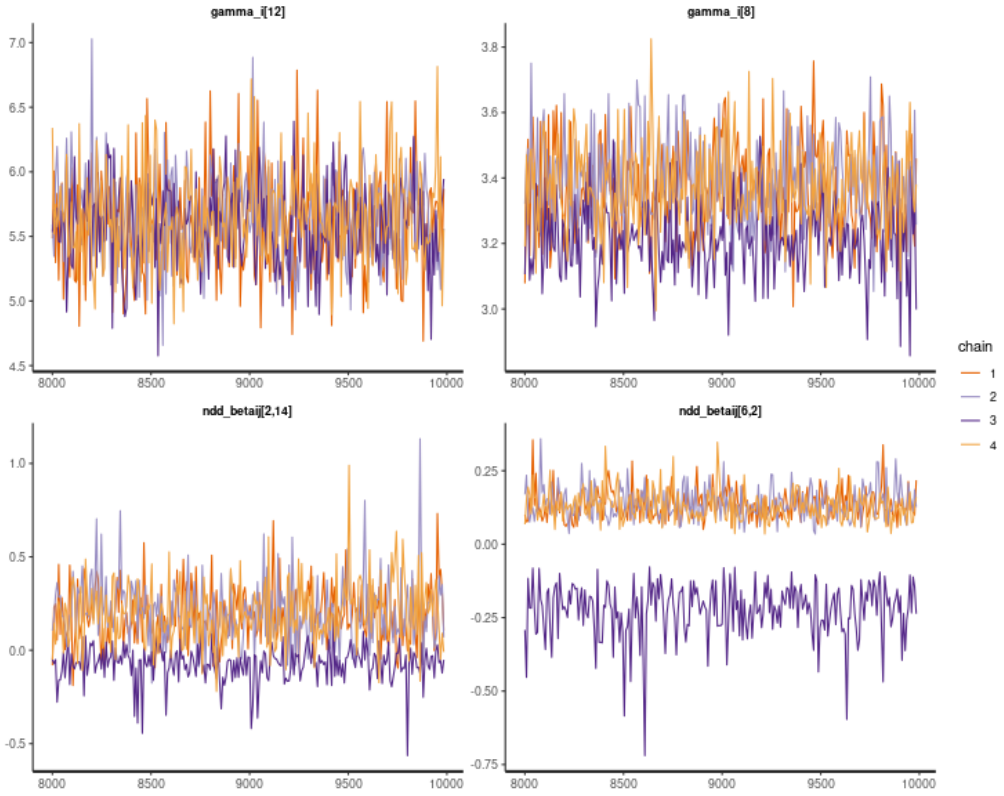

Figure S2.1: Traceplots for 4 parameters in the open shade category dataset. The x-axis indicates the number of post-warmup iterations, the y-axis indicates parameter values for each iteration. The top row shows parameters related to intrinsic fitness with $\hat{R} < 1.2$ and the bottom row shows parameters related to interaction effects with $\hat{R} > 1.2$. In the top right graph we see "optimal" convergence for a parameter with $\hat{R} = 1.004$. The top left graph shows convergence of the intrinsic fitness parameter with the highest $\hat{R}$ ($\hat{R} = 1.164$), whilst this value is not ideal the chains overlap greatly and the posterior distribution of this parameter is unimodal. On the bottom right, we see the chains start to separate at $\hat{R} = 1.327$, whereas the bottom left shows complete separation and no overlap between certain chains for the parameter with the highest $\hat{R}$ of 2.786, leading to a bimodal posterior centered on 0.

S2.3 Model Validation

In order to evaluate model fit to the data, we conducted posterior predictive checks (Kass et al. 1997) for each of the three environmental categories which the data were split into. Posterior predictive checks compare the observed data (red line in figures below) to simulated replicated data under the fitted model (grey lines). Data were simulated 1000 times by taking 1000 samples from full posterior distribution of each parameter, and using these estimates in the fitted model to replicate values of seed production for each focal individual observation. The black line shows simulated data when taking the median of each parameter posterior. Please note that the x-axis shows the log of seed production, whereas the y-axis shows the density of these seed production values.

Overall, our model predicted values of seed production which matched observations (Figures S2.2, S2.4 and S2.4). Exceptions occurred when a focal individual was surrounded by facilitative neighbours,

where our model vastly overestimated seed number. These occurrences likely point towards the existence of higher-order interactions (Mayfield and Stouffer 2017) or saturating benefits to facilitation (Stouffer 2022). These effects were not included to minimise model complexity and simplify analysis of the interaction networks; though higher-order interactions may improve model fit, their importance appears to vary widely between species in this system (Martyn et al. 2021).

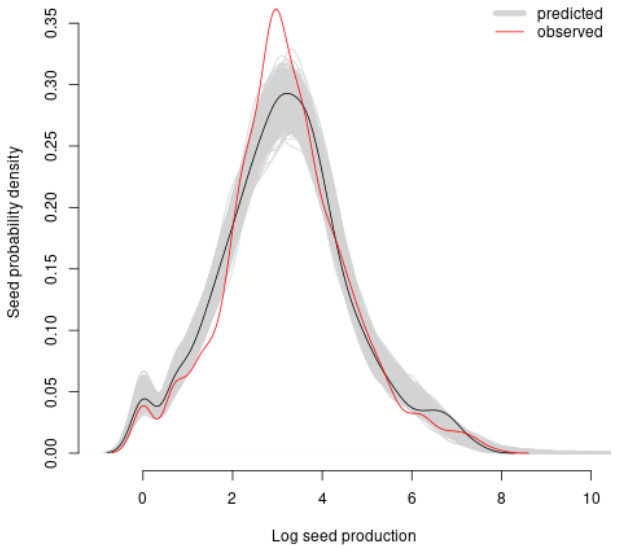

Figure S2.2: Posterior predictive check showing the density distribution of observed seed production values (red line) to simulated seed production values (light grey) for the open environment, as estimated by the joint model, on a log scale. Each grey line corresponds to simulated seed production values for 1 unique draw from the joint model posteriors, with draws covering the full posterior distribution of each parameter. The black line shows simulated values using the median of each parameter.

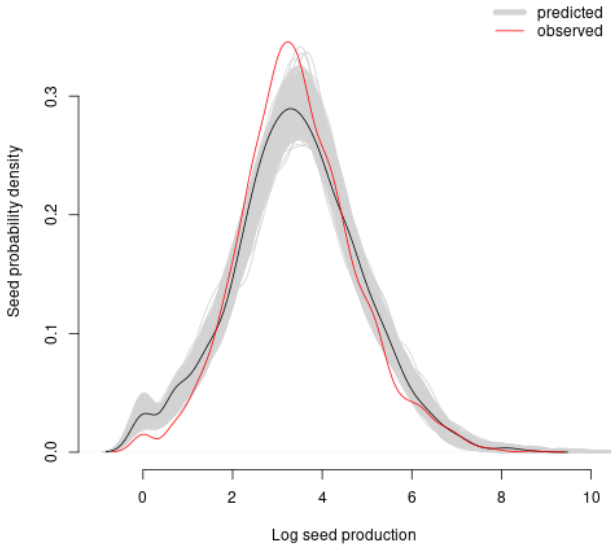

Figure S2.3: Posterior predictive check showing the density distribution of observed seed production values (red line) to simulated seed production values (light grey) for the intermediate environment, as estimated by the joint model, on a log scale. Each grey line corresponds to simulated seed production values for 1 unique draw from the joint model posteriors, with draws covering the full posterior distribution of each parameter. The black line shows simulated values using the median of each parameter.

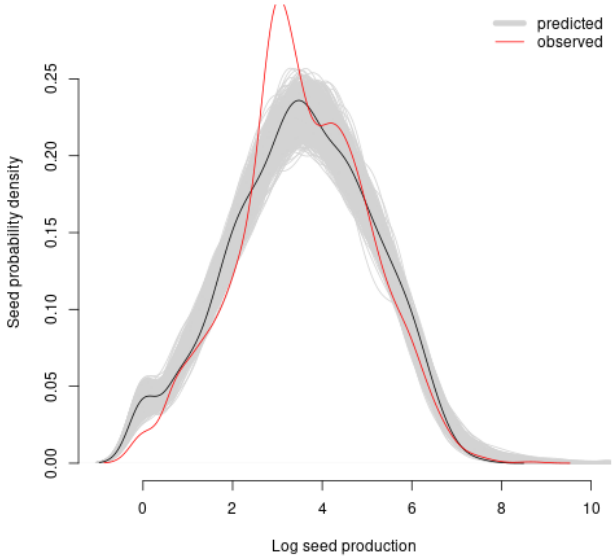

Figure S2.4: Posterior predictive check showing the density distribution of observed seed production values (red line) to simulated seed production values (light grey) for the shady environment, as estimated by the joint model, on a log scale. Each grey line corresponds to simulated seed production values for 1 unique draw from the joint model posteriors, with draws covering the full posterior distribution of each parameter. The black line shows simulated values using the median of each parameter.

S2.4 A note on significance

Throughout this study, we avoid the use of significance tests to characterise differences or relationships between interaction strengths or network properties. Both interaction strengths and networks are drawn from a large number of samples, rendering the number of 'observations' artificial and meaningless for statistical inference as p-values quickly approach zero (Lin et al. 2013). In practice, this means we only have three networks to compare. Instead, we rely on descriptive measures and figures and let readers judge for themselves whether the magnitude of differences presented here are of biological importance or interest (White et al. 2014). We also present effect sizes of the interaction effects in the Supplementary Results S3.2. Though any individual interaction effect may seem small, there is considerable evidence from the food web literature that weak interactions can have substantial impacts on the diversity and stability of a whole community (Fowler 2010; Gellner and McCann 2016; Wootton and Stouffer 2016).

S2.5 Measuring Network Properties

S2.5.1 Relative Intransitivity Index

To calculate the Relative Intransitivity Index for a network as in Laird and Schamp (2006), we first build a competitive-outcomes matrix of size $S \times S$, where S is the total number of focal species present in that network. The competitive-outcomes matrix is then populated with 1s and 0s, depending on which species outcompetes which. For two focal species i and j, the element at row i and column j of the competitive-

outcome matrix will take on the value of 1 if i outcompetes j (which counts as a 'win' below), evaluated as $\alpha''_{ji} > \alpha''_{ij}$. If this is the case, it follows that $\alpha''_{ij} < \alpha''_{ji}$ and the element at row j and column i will take on the value of 0. The sum of every row thus indicates the number of focal neighbours that a focal species outcompetes. Please note that this is a very simplified evaluation of competitive outcomes which makes several assumptions, including that 'dominance-subordinance relations within pairs of species are strictly unidirectional and deterministic' (Laird and Schamp 2006).

From this competitive-outcomes matrix, we then measure the relative variance based on the observed variance in the distribution of wins (var_{obs}), scaled by the minimum and maximum possible number of wins, var_{min} and var_{max} respectively. Relative variance, also named Relative Intransitivity Index, is thus defined as $(var_{obs} - var_{min})/(var_{max} - var_{min})$ and scales between 0 and 1. A value of 1 occurs when $var_{obs} = var_{max}$ and indicates a fully hierarchical network (where one species outcompetes all others, the second species outcompetes all but the first, the third species outcompetes all but the first and second, and so on), whereas a value of 0 occurs when $var_{obs} = var_{min}$ and indicates a fully intransitive network. In our study, the competitive-outcomes matrix and the Relative Intransitivity index were evaluated for each network sample, the median and 80% credibility interval across all samples is reported in the main text.

S2.5.2 Weighted connectance

We measured weighted connectance as in Ulanowicz and Wolff (1991) and Kinlock (2019). Weighted connectance depends on the Shannon-Weiner measure of diversity of interaction strengths for each species. For every focal species i, we first calculate the diversity of interaction strengths directed towards i, $H_{in,i}$, as well as the diversity of interaction strengths directed away from i, $H_{out,i}$:

$$H_{in,i} = -\sum_{j=1}^{S} \left(\frac{\left|\alpha''_{ij}\right|}{\sum_{j=1}^{S} \left|\alpha''_{ij}\right|} log_2 \frac{\left|\alpha''_{ij}\right|}{\sum_{j=1}^{S} \left|\alpha''_{ij}\right|} \right) \tag{S2.4}$$

$$H_{out,i} = -\sum_{j=1}^{S} \left(\frac{\left|\alpha''_{ji}\right|}{\sum_{j=1}^{S} \left|\alpha''_{ji}\right|} log_2 \frac{\left|\alpha''_{ji}\right|}{\sum_{j=1}^{S} \left|\alpha''_{ji}\right|} \right) \tag{S2.5}$$

where S is species richness. Secondly, we use these measures to calculate weighted quantitative linkage density (LD) as the effective number of species with which each species interacts, weighted by the diversity of their interaction strengths:

$$LD_{qw} = \frac{1}{2} \left(\sum_{i=1}^{S} \frac{\sum_{j=1}^{S} \left|\alpha''_{ji}\right|}{\sum \left|\alpha''\right|} 2^{H_{out,i}} + \sum_{i=1}^{S} \frac{\sum_{j=1}^{S} \left|\alpha''_{ij}\right|}{\sum \left|\alpha''\right|} 2^{H_{in,i}} \right) \tag{S2.6}$$

with $\sum |\alpha''|$ defined as the sum of the magnitudes of all interaction strengths for all species. Finally, weighted connectance is given by:

$$C_{qw} = \frac{LD_{qw}}{S} \tag{S2.7}$$

Weighted connectance is thus influenced not only by the number of species and the number of interaction strengths in a network, but also by the distribution of interaction strengths. Please note that we included both facilitative and competitive interaction strengths when calculating weighted connectance.

S2.5.3 Modularity

Broadly speaking, modularity quantifies the degree to which a network can be subdivided into modules where nodes within a module are densily connected, but nodes belonging to different modules are more sparsely connected. High modularity ($Q \rightarrow 1$) indicates strong separation of a network into different modules. There are many ways to calculate modularity, but to the authors' knowledge there is only one method which can account for both positive and negative directed links (Traag and Bruggeman 2009). The method developped by Traag and Bruggeman (2009) identifies a module as a cluster of nodes which are predominantly, positively linked to one another, but negatively linked to other clusters. Their method still accounts for link density (higher density within modules, and lower density between modules) and is based on the notion that nodes which are positively associated with one another will have similar, negative associations towards other nodes.

The modularity method developed by Traag and Bruggeman (2009) can be applied with the *cluster_spinglass()* function from the igraph package (Csárdi and Nepusz 2006), setting the *implementation* argument to 'neg'. The function uses a simulated annealing algorithm to identify modules with few negative links within a module and many negative links between modules. We set facilitative interactions in our networks to be positive and competitive interactions to be negative to stay aligned with the theoretical basis underlying this definition of modularity.

References

Bimler, M. D. et al. (2018). "Accurate predictions of coexistence in natural systems require the inclusion of facilitative interactions and environmental dependency". In: *Journal of Ecology* 106.5, pp. 1839–1852. DOI: 10.1111/1365-2745.13030.

Bimler, M. D. et al. (2023). "Estimating interaction strengths for diverse horizontal systems using performance data". In: *Methods in Ecology and Evolution*. DOI: 10.1111/2041-210X.14068.

Conservation International (2023). "Biodiversity Hotspots". In: *https://www.conservation.org/priorities/biodiversity-hotspots*. 21 Septemb.

Csárdi, G. and Nepusz, T. (2006). "The igraph software package for complex network research". In: *InterJournal Complex Systems*, p. 1695.

Dwyer, J. M. et al. (2015). "Climate moderates release from nutrient limitation in natural annual plant communities". In: *Global Ecology and Biogeography* 24.5, pp. 549–561. DOI: 10.1111/geb.12277.

Fowler, M. S. (2010). "Extinction cascades and the distribution of species interactions". In: *Oikos* 119.5, pp. 864–873. DOI: 10.1111/j.1600-0706.2009.17817.x.

Gellner, G. and McCann, K. S. (2016). "Consistent role of weak and strong interactions in high- and low-diversity trophic food webs". In: *Nature Communications* 7.11180. DOI: 10.1038/ncomms11180.

Godoy, O. and Levine, J. M. (2014). "Phenology effects on invasion success: Insights from coupling field experiments to coexistence theory". In: *Ecology* 95.3, pp. 726–736. DOI: 10.1890/13-1157.1.

Kass, R. E. et al. (1997). "Markov Chain Monte Carlo in Practice : A Roundtable Discussion". In: *The American Statistician* 52.2, p. 93. DOI: 10.2307/2685466.

Kinlock, N. L. (2019). "A meta-analysis of plant interaction networks reveals competitive hierarchies as well as facilitation and intransitivity". In: *The American Naturalist* 194.5, pp. 640–653. DOI: 10.1086/705293.

Laird, R. A. A. and Schamp, B. S. S. (2006). "Competitive Intransitivity Promotes Species Coexistence". In: *The American Naturalist* 168.2, pp. 182–193. DOI: 10.1086/506259.

Levine, J. M. and HilleRisLambers, J. (2009). "The importance of niches for the maintenance of species diversity." In: *Nature* 461.7261, pp. 254–257. DOI: 10.1038/nature08251.

Lin, M., Lucas, H. C., and Shmueli, G. (2013). "Too big to fail: Large samples and the p-value problem". In: *Information Systems Research* 24.4, pp. 906–917. DOI: 10.1287/isre.2013.0480.

Martyn, T. E. et al. (2021). "Identifying "useful" fitness models: Balancing the benefits of added complexity with realistic data requirements in models of individual plant fitness". In: *American Naturalist* 197.4, pp. 415–433. DOI: 10.1086/713082.

Mayfield, M. M. and Stouffer, D. B. (2017). "Higher-order interactions capture unexplained complexity in diverse communities". In: *Nature Ecology & Evolution* 1, pp. 1–7. DOI: 10.1038/s41559-016-0062.

Stouffer, D. B. (2022). "A critical examination of models of annual-plant population dynamics and density-dependent fecundity". In: *Methods in Ecology and Evolution* 13, pp. 2516–2530. DOI: 10.1111/2041-210X.13965.

Traag, V. A. and Bruggeman, J. (2009). "Community detection in networks with positive and negative links". In: *Physical Review E* 80.3, pp. 1–6. DOI: 10.1103/PhysRevE.80.036115.

Ulanowicz, R. E. and Wolff, W. F. (1991). "Ecosystem flow networks: Loaded dice?" In: *Mathematical Biosciences* 103.1, pp. 45–68. DOI: 10.1016/0025-5564(91)90090-6.

Wainwright, C. E., Dwyer, J. M., and Mayfield, M. M. (2017). "Effects of exotic annual grass litter and local environmental gradients on annual plant community structure". In: *Biological Invasions* 19.2, pp. 479–491. DOI: 10.1007/s10530-016-1303-2.

White, J. W. et al. (2014). “Ecologists should not use statistical significance tests to interpret simulation model results”. In: *Oikos* 123.4, pp. 385–388. DOI: 10.1111/j.1600-0706.2013.01073.x.
Wootton, K. L. and Stouffer, D. B. (2016). “Many weak interactions and few strong; food-web feasibility depends on the combination of the strength of species’ interactions and their correct arrangement”. In: *Theoretical Ecology* 9.2, pp. 185–195. DOI: 10.1007/s12080-015-0279-3.

S3 Supplementary Information 3: Results

S3.1 Supplementary Figures

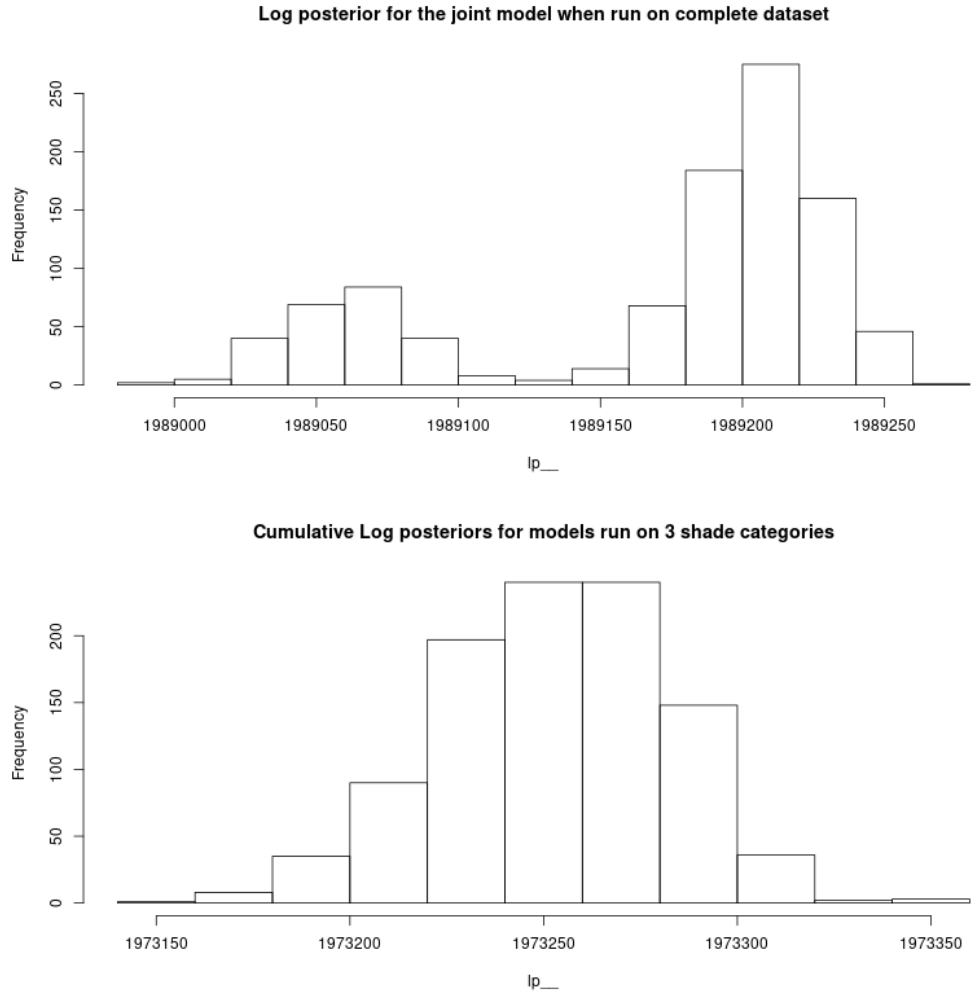

Figure S3.1: Upper panel shows the distribution of 1000 samples of the 'lp__' parameter (log posterior) returned by STAN for the joint model run on the whole dataset (no shade categories). Lower pannel shows the cumulative distribution of 1000 draws of the 'lp__' parameter for the joint model run on each of the three shade categories (for a total of 3000 samples). Upper distribution is bimodal due to imperfect convergence on the full dataset (see S2.2).

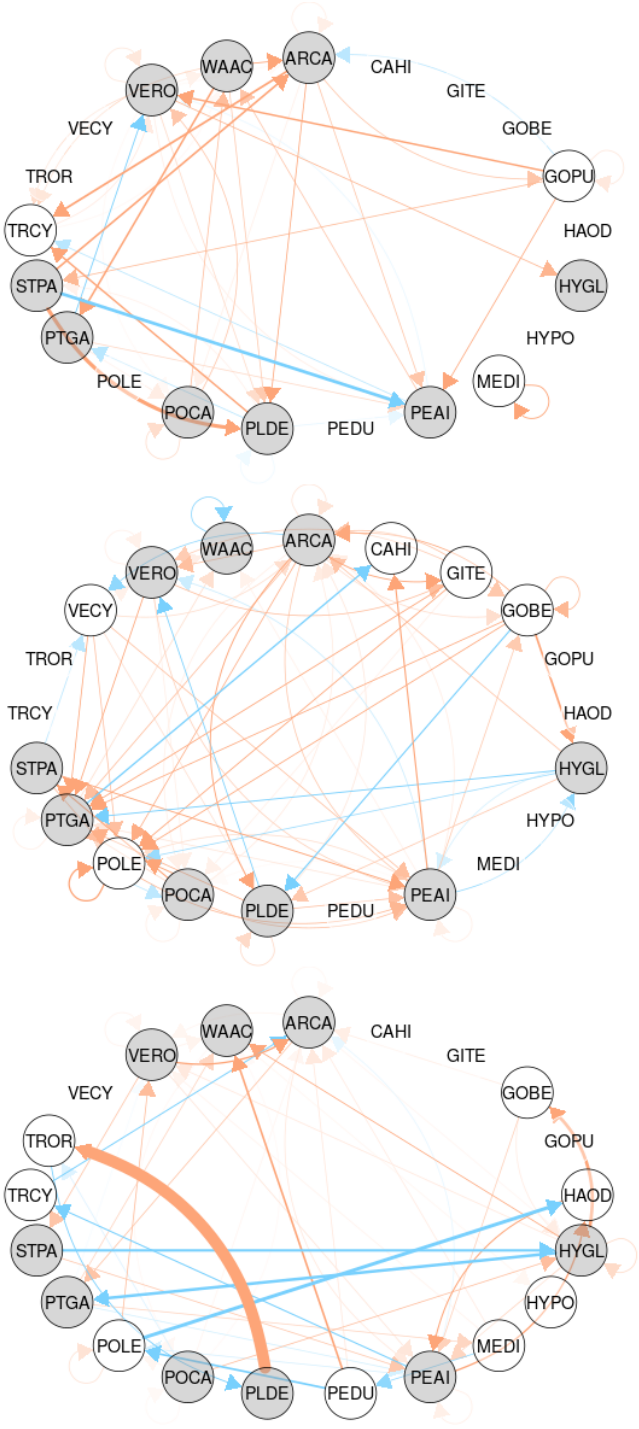
WAAC
ARCA
CAHI
VERO
GITE
VECY
GOBE
TROR
GOPU
TRCY
HAOD
STPA
HYGL
PTGA
HYPO
POLE
MEDI
POCA
PEAI
PLDE
PEDU
WAAC
ARCA
CAHI
VERO
GITE
VECY
GOBE
TROR
GOPU
TRCY
HAOD
STPA
HYGL
PTGA
HYPO
POLE
MEDI
POCA
PEAI
PLDE
PEDU
WAAC
ARCA
CAHI
VERO
GITE
VECY
GOBE
TROR
GOPU
TRCY
HAOD
STPA
HYGL
PTGA
HYPO
POLE
MEDI
POCA
PEAI
PLDE
PEDU

Figure S3.2

Figure S3.2: (Caption for previous page.) Interaction networks for each shade category (open, intermediate and shady) including only those interaction strengths whose 80% credibility interval did not overlap with 0. This accounts for 29.9%, 36.2%, and 18% of all intra- and interspecific interaction strengths respectively. Line thickness is determined by median absolute interaction strength, with competition in orange and facilitation in blue. Arrows point to species i, the species receiving the interaction. All 21 focal species are presented in each network to facilitate comparison, identified by their four-letter species code (see Table 1 in the main text). Species nodes with a border are focal species in that particular network, and grey shading indicates those species which are focal species in all three networks. Species nodes with no border are not focal species in that particular network.

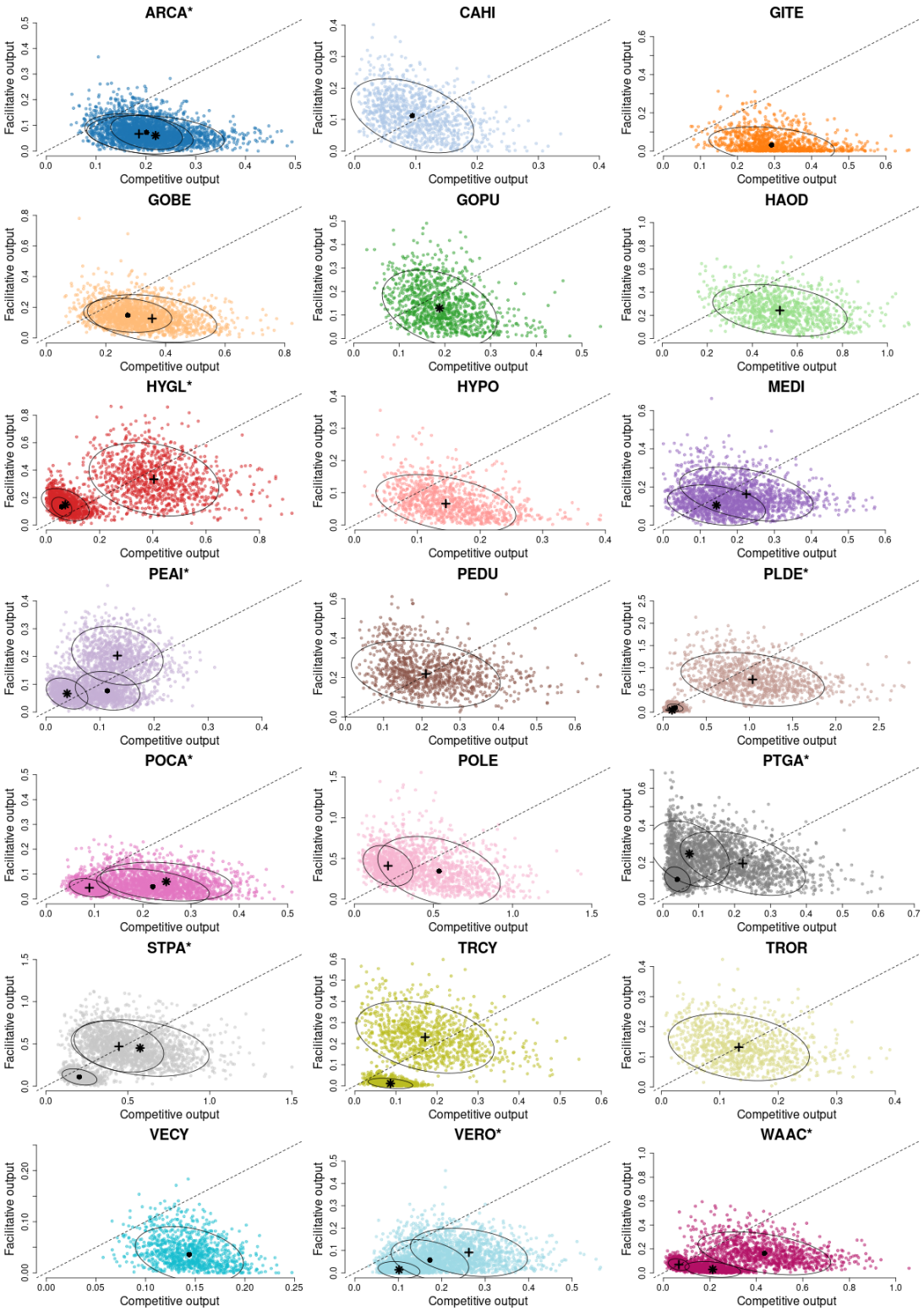
ARCA*
CAHI
GITE
GOBE
GOPU
HAOD
HYGL*
HYPO
MEDI
PEAI*
PEDU
PLDE*
POCA*
POLE
PTGA*
STPA*
TRCY
TROR
VECY
VERO*
WAAC*
Facilitative output
Competitive output

Figure S3.3

Figure S3.3: (Caption for previous page.) Many species had both competitive and facilitative effects on neighbours. Values are given as the absolute sum of scaled per-capita interaction strengths emitted by focal species (output) on neighbouring focals, either competitive (x-axis) or facilitative (y-axis). The black asterisk ($*$), circle ($\bullet$) and cross ($+$) represent species medians in the open, intermediate, and shady networks, respectively. Coloured dots correspond to samples for each species and network, the ellipses center on the median and describe the 80% normal density contour. The dashed line indicates where competitive output and facilitative output are equal. The colours used for each species are the same as in Figs. 1 and 2 of the main text. Note that the scales of the x and y-axes differ between species. Species which are focals in all three networks can be identified by the asterisk next to their four-letter code.

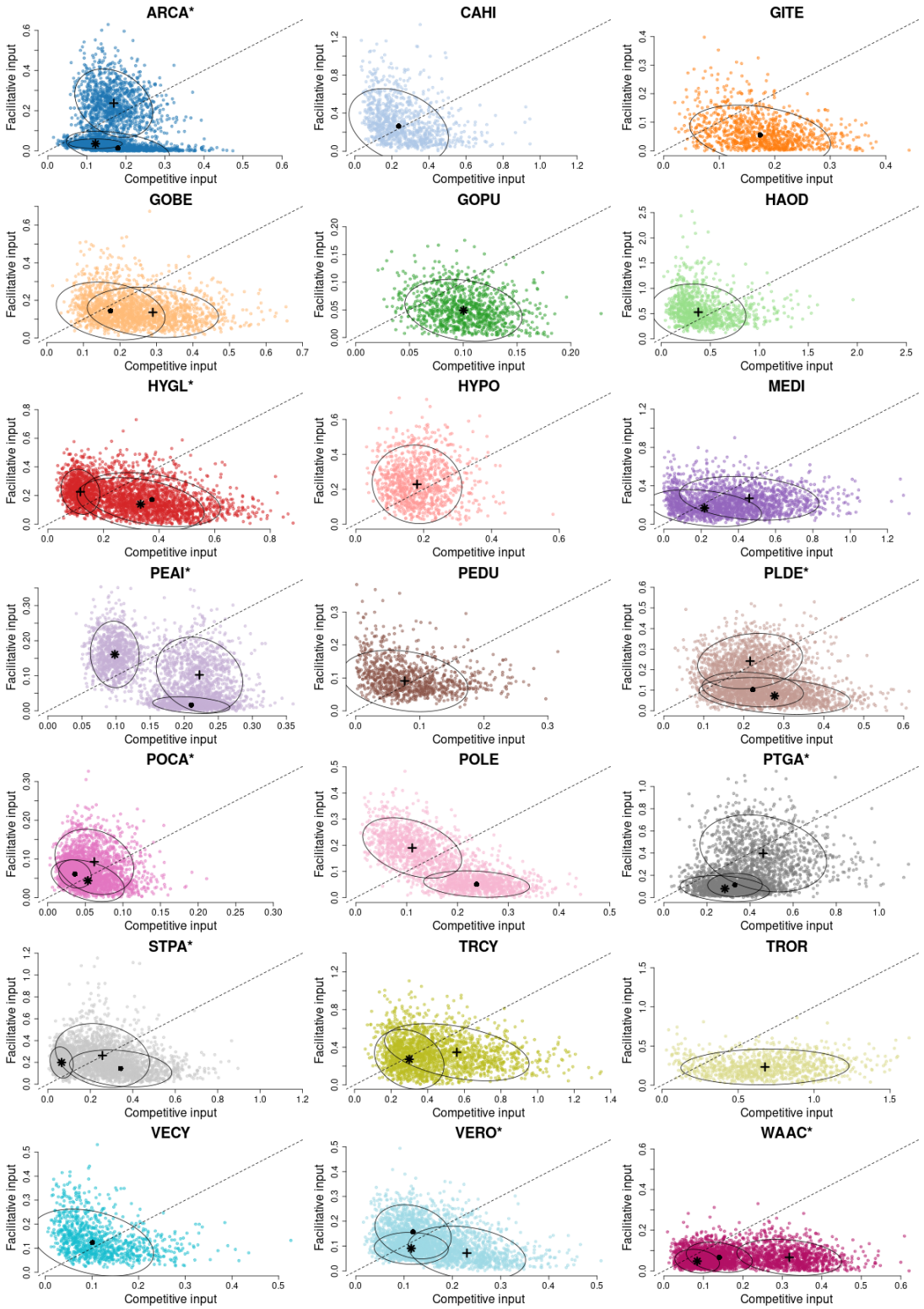
ARCA*
CAHI
GITE
GOBE
GOPU
HAOD
HYGL*
HYPO
MEDI
PEAI*
PEDU
PLDE*
POCA*
POLE
PTGA*
STPA*
TRCY
TROR
VECY
VERO*
WAAC*
Facilitative input
Competitive input

Figure S3.4

Figure S3.4: (Caption for previous page.) Many species received both competitive and facilitative interaction strengths from neighbours. Values are given as the absolute sum of scaled per-capita interaction strengths received by focal species (input) from neighbouring species, either competitive (x-axis) or facilitative (y-axis). The black asterisk ($*$), circle ($\bullet$) and cross ($+$) represent species medians in the open, intermediate, and shady networks, respectively. Coloured dots correspond to samples for each species and network, the ellipses center on the median and describe the 80% normal density contour. The dashed line indicates where competitive input and facilitative input are equal. The colours used for each species are the same as in Figs. 1 and 2 of the main text. Note that the scales of the x and y-axes differ between species. Species which are focals in all three networks can be identified by the asterisk next to their four-letter code.

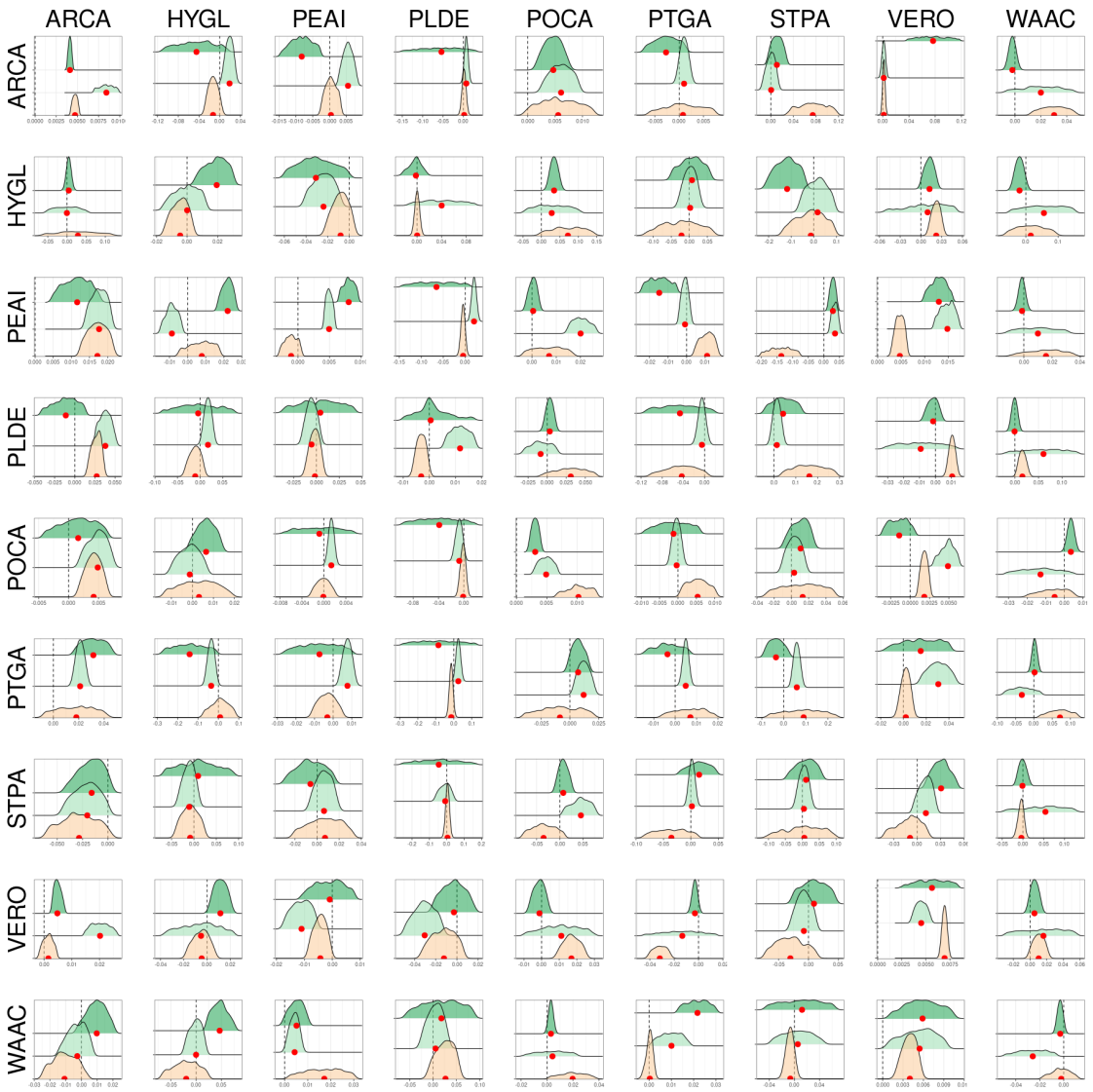

Figure S3.5: Credibility interval distributions of the scaled per-capita interaction strengths α''_{ij} between the 9 focal species which were present in all three networks. Rows correspond to species i and columns to species j, the distribution of a specicific α''_{ij} can be found by looking at the i'th row and j'th column. The diagonal shows the distributions of intraspecific interaction strengths (α''_{ii}). For each α''_{ij}, the corresponding graph plots the distribution of that interaction strength (x-axis) in the shady (dark green), intermediate (light green) and open (yellow) environments (y-axis). On the x-axis, values greater than 0 indicate competition and values lesser than 0 indicate facilitation, the red dots mark the median value of each distribution. Note that the scale of the x-axis varies for each interaction strength.

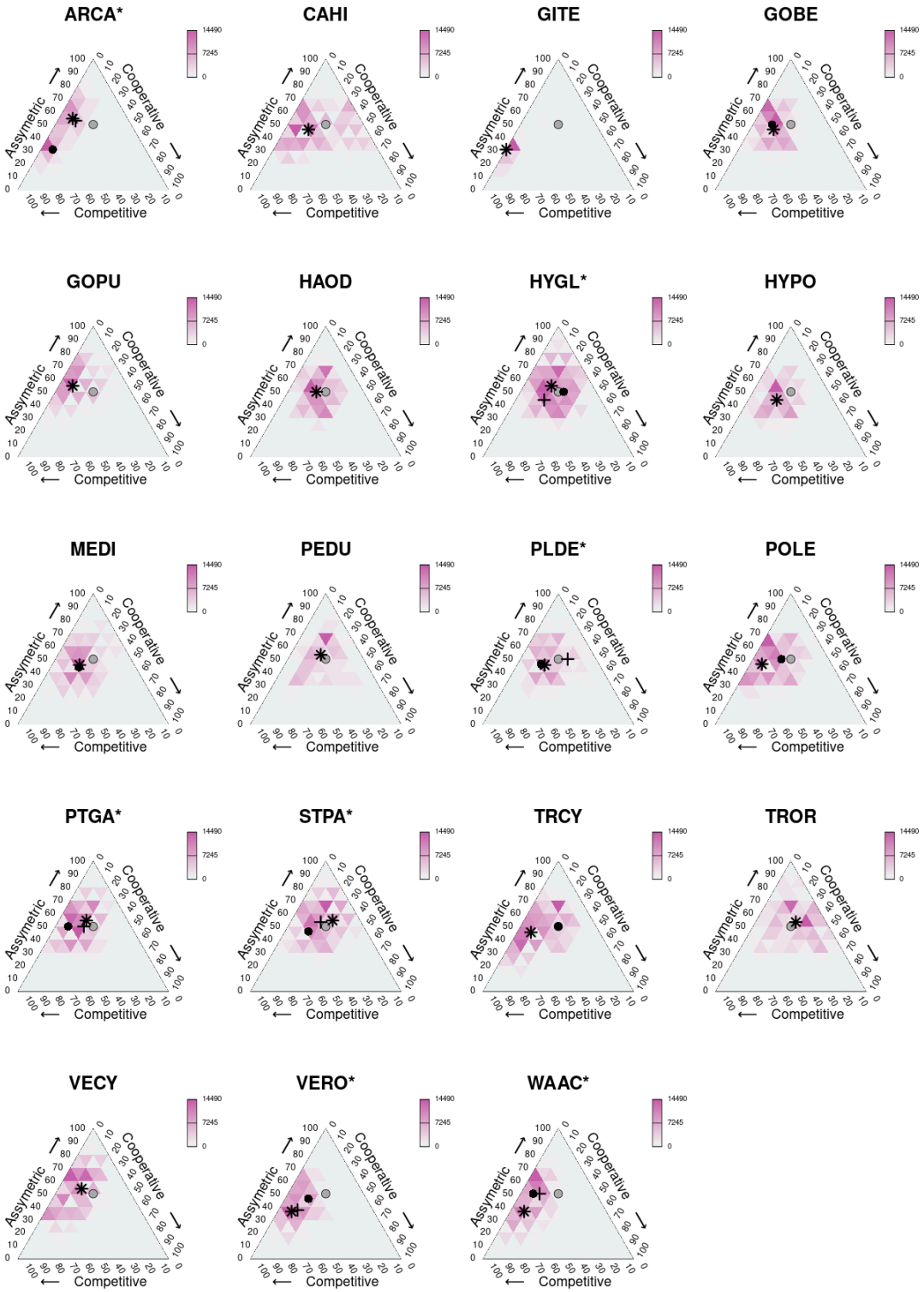
ARCA*
CAHI
GITE
GOBE
GOPU
HAOD
HYGL*
HYPO
MEDI
PEDU
PLDE*
POLE
PTGA*
STPA*
TRCY
TROR
VECY
VERO*
WAAC*
Asymmetric
Cooperative
Competitive
14490
7245
0

Figure S3.6

Figure S3.6: (Caption for previous page.) This figure shows how species participation in different types of interaction loops sometimes varied between networks, and sometimes did not. Axes represent the proportion of cooperative (+/+), competitive (-/-) and asymmetrix (+/- or -/+) pairwise interaction loops. We show the median proportion of loops which each species engaged in for each shade category: a black asterisk ($*$) for the open network, a circle ($\bullet$) for the intermediate network and a cross ($+$) for the shady network. The background colour indicates the density of values across all samples, going from light grey (low density) to magenta (high density). The grey circle centered in the top half of each plot serves as a reference point, indicating where a species would be placed if it engaged in 25% cooperative loops, 25% competitive loops, and 50% asymmetric loops. Species which are focals in all three networks can be identified by the asterisk next to their four-letter code.

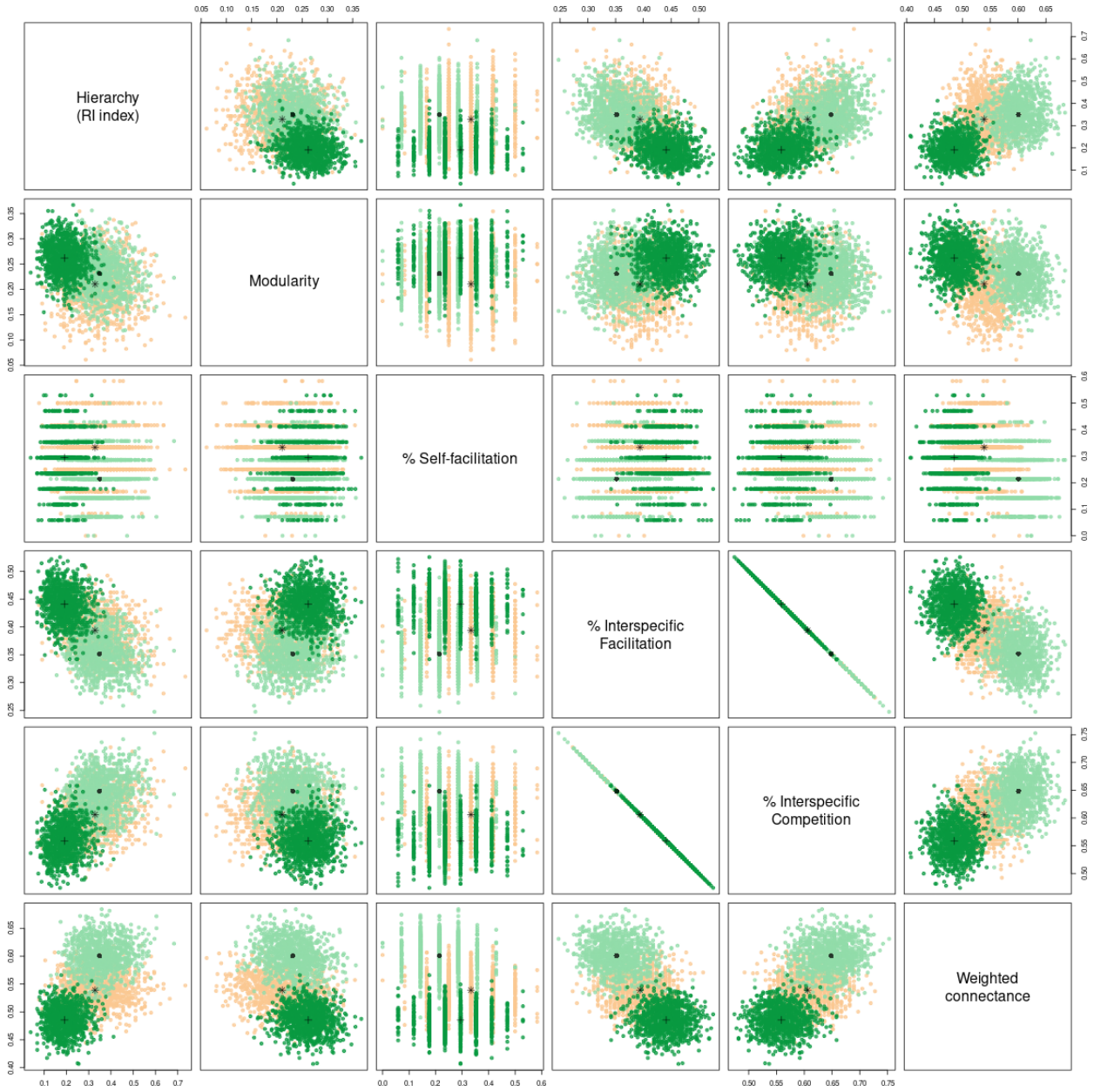

Figure S3.7: Scatterplots of the network properties used in analysis and Figure 5 of the main text. Points are coloured according to the network shade category: yellow for open, light green for intermediate and dark green for shady. The black points indicate the median value for each network: an asterisk ($*$) for the open network, a circle ($\bullet$) for the intermediate network and a cross ($+$) for the shady network.

S3.2 Interaction effect sizes

	λ_i	α_{ii}	$\widetilde{\alpha_{ij}}$	$\widetilde{\alpha_{ji}}$	$10\%\alpha_{ij}$	$90\%\alpha_{ij}$	$10\%\alpha_{ji}$	$90\%\alpha_{ji}$
median	41.52	-6.25	-5.02	-4.95	-34.38	13.76	-28.53	19.41
ARCA	69.04	-6.78	-3.25	-13.67	-31.11	25.72	-31.46	24.36
GOPU	75.59	-10.7	-2.12	-6.78	-41.45	9.28	-42.79	29.26
HYGL	21.91	4.36	-6.95	4.34	-32.18	8.75	-6.89	17.54
MEDI	17.7	-22.9	-1.16	-3.13	-9.34	6.96	-14.75	21.74
PEAI	282.25	3.53	-13.83	2.25	-45.93	12.94	-11.4	21.29
PLDE	37.09	4.22	-8.47	0.82	-30.26	12.96	-41.77	13.48
POCA	459.9	-16.69	-6.88	-10.66	-43.11	13.73	-25.6	5.17
PTGA	27.9	-5.72	-2.77	-0.32	-41.49	15.01	-12.39	57.38
STPA	41.32	-2.65	0.48	2.41	-34.24	44.15	-37.51	51.64
TRCY	10.57	-7.73	-5.37	-14.92	-34.52	45.07	-33.48	-0.88
VERO	41.72	-6.85	-4.93	-7.29	-26.37	13.79	-20.2	-0.21
WAAC	279.34	5.06	-5.11	-12.89	-41.79	27.41	-35.58	4.53

Table S3.1: **Median intrinsic seed production and interaction effect sizes for the open environment.** The first column of numbers gives λ_i, median intrinsic seed production of each focal species in the absence of neighbours. The following columns give effect sizes of unscaled interaction effects as a percentage $\%$ increase (positive) or decrease (negative) to λ_i caused by the presence of one individual plant neighbour within the interaction neighbourhood of the focal species. Columns 2 and 3 give the median of these effect sizes depending on whether the neighbour is conspecific or heterospecific, respectively. Fourth column indicates the median effect size of one individual of the focal species on intrinsic seed production of one neighbour, across all focal and non-focal neighbour species. The last four columns give the upper and lower limits of the 80% credibility interval of these effects, also as $\%$ increases or decreases to λ_i. The interaction effects used to calculate effect sizes were not scaled into per capita interaction strengths. The top row gives the median value of each column.

	λ_i	α_{ii}	$\widetilde{\alpha_{ij}}$	$\widetilde{\alpha_{ji}}$	$10\%\alpha_{ij}$	$90\%\alpha_{ij}$	$10\%\alpha_{ji}$	$90\%\alpha_{ji}$
median	52.06	-7.11	-4.72	-8.55	-30.63	20.56	-30.63	16.48
ARCA	192.04	-15.61	-9.44	-14.19	-28.67	4.45	-40.93	18.13
CAHI	11.43	-2.14	-0.18	-2.13	-18.98	2.63	-13.3	20.05
GITE	46.24	-5.62	-4.06	-21.29	-19.51	9.51	-49.92	-1.2
GOBE	146.79	-29.32	-5.43	-8.63	-30.03	27.52	-49.57	28.98
HYGL	36.23	-0.07	-4.8	7.95	-32.28	4.92	-22.27	65.91
PEAI	250.97	-10.06	-13.22	-5.36	-41.83	25.08	-27.5	29.42
PLDE	44.52	-10.62	-4.64	-2.49	-36.04	19.95	-23.04	18.72
POCA	211.5	-8.58	-0.14	-9.29	-16.19	21.17	-34.38	2.2
POLE	1669.19	-45.71	-17.92	-7.24	-56.62	32.88	-25.6	7.81
PTGA	39.82	-13.82	-7.66	-2.53	-31.66	41.21	-8.62	33.35
STPA	37.85	-5.63	-2.92	-8.75	-34.92	37.99	-48.01	11.7
VECY	57.88	-1.63	0.02	-10.56	-9.05	8.64	-33.75	2.73
VERO	34.56	-3.03	1.38	-8.46	-9.86	31.46	-21.7	3.99
WAAC	257.37	46.89	-7.08	-17.8	-31.24	1.18	-38.02	14.83

Table S3.2: **Intrinsic seed production and interaction effect sizes for the intermediate environment.** The first column of numbers gives λ_i, median intrinsic seed production of each focal species in the absence of neighbours. The following columns give effect sizes of unscaled interaction effects as a percentage $\%$ increase (positive) or decrease (negative) to λ_i caused by the presence of one individual plant neighbour within the interaction neighbourhood of the focal species. Columns 2 and 3 give the median of these effect sizes depending on whether the neighbour is conspecific or heterospecific, respectively. Fourth column indicates the median effect size of one individual of the focal species on intrinsic seed production of one neighbour, across all focal and non-focal neighbour species. The last four columns give the upper and lower limits of the 80% credibility interval of these effects, also as $\%$ increases or decreases to λ_i. The interaction effects used to calculate effect sizes were not scaled into per capita interaction strengths. The top row gives the median value of each column.

	λ_i	α_{ii}	$\widetilde{\alpha_{ij}}$	$\widetilde{\alpha_{ji}}$	$10\%\alpha_{ij}$	$90\%\alpha_{ij}$	$10\%\alpha_{ji}$	$90\%\alpha_{ji}$
median	45.98	-8.16	-0.25	-2.23	-28.26	27.04	-26.52	26.47
ARCA	66.74	-5.89	-0.59	-5.91	-35.31	36.61	-21.51	6.3
GOBE	45.98	-12.67	-0.68	-12.84	-11.98	14.6	-37.66	7.14
HAOD	11.21	5.84	-0.04	-2.72	-18.83	49.68	-49.04	18.99
HYGL	32.72	-18.8	-0.22	-2.23	-11.36	36.09	-43.65	28.67
HYPO	37.86	-1.56	-0.25	-2.67	-31.59	32.11	-25.12	0.02
MEDI	5.29	-3.36	-0.68	-6.71	-28.26	15.34	-26.64	37.42
PEAI	279.38	-11.22	-5.58	1.82	-37.33	59.49	-7.04	26.63
PEDU	57.28	-8.16	-0.05	-0.07	-7.35	2.32	-19.7	20.91
PLDE	39.66	-0.23	0.02	0.84	-12.56	27.04	-21.41	45.98
POCA	261.48	-17.09	0.67	-4.93	-28.28	24.41	-26.52	10.34
POLE	294.3	-28.4	-0.63	-0.39	-15.62	16.65	-33.55	49.2
PTGA	17	2.69	-0.05	-1.32	-33.99	32.9	-20.56	30.37
STPA	88.45	-12.21	-0.2	-8.48	-36.05	29.96	-35.39	62.9
TRCY	15.9	-19.75	-1.61	0	-32.52	22.44	-10.53	16.91
TROR	8.59	-2.62	0.05	0.06	-8.49	19.11	-11.49	26.47
VERO	70.38	-11.98	-1.13	-17.58	-20.22	13.04	-47.31	6.14
WAAC	222.6	29.59	-11.27	3.47	-46.94	32.17	-29.16	37.93

Table S3.3: **Intrinsic seed production and interaction effect sizes for the shady environment.** The first column of numbers gives λ_i, median intrinsic seed production of each focal species in the absence of neighbours. The following columns give effect sizes of unscaled interaction effects as a percentage $\%$ increase (positive) or decrease (negative) to λ_i caused by the presence of one individual plant neighbour within the interaction neighbourhood of the focal species. Columns 2 and 3 give the median of these effect sizes depending on whether the neighbour is conspecific or heterospecific, respectively. Fourth column indicates the median effect size of one individual of the focal species on intrinsic seed production of one neighbour, across all focal and non-focal neighbour species. The last four columns give the upper and lower limits of the 80% credibility interval of these effects, also as $\%$ increases or decreases to λ_i. The interaction effects used to calculate effect sizes were not scaled into per capita interaction strengths. The top row gives the median value of each column.