

Title: Estimating species interactions from co-occurrence data with Markov networks

Author: David J. Harris: Population Biology; 1 Shields Avenue, Davis CA, 95616

Abstract: Estimating species interactions from co-occurrence data is one of the most controversial tasks in community ecology. One difficulty is that a single pairwise interaction can ripple through an ecological network and produce surprising indirect consequences. For example, the negative correlation between two competing species can be reversed in the presence of a third species that is capable of outcompeting both of them. Here, I apply models from statistical physics, called Markov networks or Markov random fields, that can predict the direct and indirect consequences of any possible species interaction matrix.

Interactions in these models can also be estimated from observed co-occurrence rates via maximum likelihood, controlling for indirect effects. Using simulated landscapes with known pairwise interaction strengths, I evaluated Markov networks and six existing approaches. The Markov networks consistently outperformed other methods, correctly isolating direct interactions between species pairs even when indirect interactions or abiotic factors largely overpowered them. Two computationally efficient approximations, based on controlling for indirect effects with linear or generalized linear models, also performed well. Indirect effects reliably caused a common null modeling approach to produce incorrect inferences, however.

Key words: Ecological interactions; Occurrence data; Species associations; Markov network; Markov random field; Ising model; Biogeography; Presence–absence matrix; Null model

Introduction

To the extent that nontrophic species interactions (such as competition) affect community assembly, ecologists might expect to find signatures of these interactions in species composition data (MacArthur 1958, Diamond 1975). Despite decades of work and several major controversies, however (Lewin 1983, Strong et al. 1984, Connor et al. 2013), existing

methods for detecting competition's effects on community structure are unreliable (Gotelli and Ulrich 2009). In particular, species' effects on one another can become lost in the complex web of indirect effects. For example, the competitive interaction between the two shrub species in Figure 1A can become obscured by their shared tendency to occur in unshaded areas (Figure 1B). While ecologists have long known that indirect effects can overwhelm direct ones at the landscape level (Dodson 1970, Levine 1976), the vast majority of our methods for drawing inferences from observational data do not control for these effects (e.g. Diamond 1975, Strong et al. 1984, Gotelli and Ulrich 2009, Veech 2013, Pollock et al. 2014). To the extent that indirect interactions like those in Figure 1 are generally important (Dodson 1970), existing methods will thus not provide much evidence regarding species interactions.

While competition doesn't reliably reduce co-occurrence rates at the whole-landscape level (as most methods assume), it nevertheless does leave a signal in the data (Figure 1C). For example, after partitioning the data set into shaded sites and unshaded sites, there will be co-occurrence deficits in each subset that might not be apparent at the whole-landscape level. More generally, controlling for other species in the network will often be important for obtaining reliable estimates of direct (conditional, or all-else-equal) effects. This kind of precision is difficult to obtain from null models, which only simulate the most extreme possible scenario: that *all* direct and indirect interactions are exactly zero. Nevertheless, null models have dominated this field for more than three decades (Strong et al. 1984, Gotelli and Ulrich 2009).

Following Azaele et al. (2010), this paper shows that Markov networks (undirected graphical models also known as Markov random fields; Murphy 2012) can provide a framework for understanding the landscape-level consequences of pairwise species interactions, and for detecting them from observed presence-absence matrices. Markov networks have been

used in many scientific fields in similar contexts for decades, from physics (where nearby particles interact magnetically; Cipra 1987) to spatial statistics (where adjacent grid cells have correlated values; Harris 1974, Gelfand et al. 2005). While community ecologists explored some related approaches in the 1980’s (Whittam and Siegel-Causey 1981), they used severe approximations that led to unintelligible results (e.g. “probabilities” greater than one; Gilpin and Diamond 1982).

Below, I introduce Markov networks and show how they can be used to simulate landscape-level data from an interaction matrix or to make exact predictions about the direct and indirect consequences of possible interaction matrices. Then, using simulated data sets where the “true” interactions are known, I compare this approach with several existing methods. Finally, I discuss opportunities for extending the approach presented here to other problems in community ecology, e.g. quantifying the overall effect of species interactions on occurrence rates (Roughgarden 1983) and disentangling the effects of biotic versus abiotic interactions on species composition (Pollock et al. 2014).

Methods

Markov networks. Markov networks provide a framework for translating back and forth between the conditional (all-else-equal) relationships among species (Figure 1C) and the kinds of species assemblages that these relationships produce. Here, I show how a set of conditional relationships can be used to determine how groups of species can co-occur. Methods for estimating conditional relationships from data are discussed in the next section.

A Markov network defines the relative probability of observing a given vector of species-level presences (1s) and absences (0s), \vec{y} , as

$$p(\vec{y}; \alpha, \beta) \propto \exp\left(\sum_i \alpha_i y_i + \sum_{\langle ij \rangle} \beta_{ij} y_i y_j\right),$$

where the second sum is over all $\frac{1}{2}n(n-1)$ pairs of n species. In this model, α_i is an intercept term determining the amount that the presence of species i contributes to the log-probability of \vec{y} ; it directly controls the prevalence of species i . Similarly, β_{ij} is the amount that the co-occurrence of species i and species j contributes to the log-probability; it controls the conditional relationship between two species, i.e. the probability that they will be found together, after controlling for the other species in the network (Figure 2A, Figure 2B). For example, β_{ij} might have a value of +2 for two mutualists, indicating that the odds of observing one species are e^2 times higher in sites where its partner is present than in comparable sites where its partner is absent.

Because the relative probability of a presence-absence vector increases when positively-associated species co-occur and decreases when negatively-associated species co-occur, the model tends—all else equal—to produce assemblages that have many pairs of positively-associated species and relatively few pairs of negatively-associated species (exactly as an ecologist might expect). Of course, if all else is *not* equal (e.g. Figure 1, where the presence of one competitor is associated with release from another competitor), then species' overall co-occurrence patterns can be more complicated, and may require summing over the different possible assemblages (Figure 2B) to determine the overall outcomes.

Estimating α and β coefficients from presence-absence data. In the previous section, the values of α and β were known and the goal was to make predictions about possible species assemblages. In practice, however, ecologists will often need to estimate the parameters from an observed co-occurrence matrix (i.e. from a set of independent \vec{y} vectors indicating which species are present at each site on the landscape). When the number of species is reasonably small, one can compute exact maximum likelihood estimates for all of the α and β coefficients given a presence-absence matrix by optimizing $p(\vec{y}; \alpha, \beta)$. Fully-observed Markov networks

like the ones considered here have unimodal likelihood surfaces (Murphy 2012), ensuring that this procedure will always converge on the global maximum. This maximum represents the unique combination of α and β coefficients that would be expected to produce exactly the observed co-occurrence frequencies on average (i.e., maximizing the likelihood ensures that the sufficient statistics of the model distribution match the sufficient statistics of the data; Murphy 2012). I used the *rosalia* package (Harris 2015a) for the R programming language (R Core Team 2015) to optimize the Markov network parameters.

Simulated landscapes. I simulated several sets of landscapes using known parameters as a test bed for different statistical methods. The first set of landscapes included the three competing species shown in Figure 1. For each of 1000 replicates, I generated a landscape with 100 sites by sampling from a probability distribution defined by the figure’s interaction coefficients (Appendix 1). Each of the methods described below was then evaluated on its ability to correctly infer that the two shrub species competed with one another, despite their frequent co-occurrence.

I then generated landscapes with up to 20 interacting species at 25, 200, or 1600 sites using three increasingly complex models (50 replicates for each combination of size and model). As described in Appendix [[X]], I randomly drew the “true” coefficient values for each replicate landscape so that most species pairs interacted negligibly, a few pairs interacted very strongly, and competition was three times more common than facilitation.

The first set of 20-species landscapes, like the landscapes with three species, were generated directly from a Markov network to ensure that the model could recover the parameters used to generate the “observed” co-occurrence data. Then, I added two environmental factors that varied from location to location across the simulated landscapes, and simulated a new set of co-occurrence data so that species’ α coefficients depended on the

local environment. The latter set of simulated landscapes provide an important test of the methods' ability to distinguish co-occurrence patterns that were generated from pairwise biotic interactions from those that were generated by external forces like abiotic environmental filtering. This task was made especially difficult because—as with most analyses of presence-absence data for co-occurrence patterns—the inference procedure did not have access to any information about the environmental or spatial variables that helped shape the landscape (cf Connor et al. 2013). I generated the final set of landscapes by simulating the population dynamics of 20 interacting species, as described in Appendix [[X]].

Recovering species interactions from simulated data. I compared seven techniques for determining the sign and strength of the associations between pairs of species from simulated data (Appendix 3). First, I used the *rosalia* package (Harris 2015a) to fit Markov network models, as described above. For the analyses with 20 species, I added a weak regularizer (equivalent to a logistic prior with scale 2) to ensure that the model estimates were always finite [[Appendix]].

I also evaluated six alternative methods: five from the existing literature, plus a novel combination of two of these methods. The first alternative interaction metric was the sample correlation between species' presence-absence vectors, which summarizes their marginal association. Next, I used partial correlations, which summarize species' conditional relationships. This approach is common in molecular biology (Friedman et al. 2008), but is rare in ecology (see Albrecht and Gotelli (2001) and Faisal et al. (2010) for two exceptions). In the context of non-Gaussian data, the partial correlation can be thought of as a computationally efficient approximation to the full Markov network model (Loh and Wainwright 2013). Because partial correlations are undefined for landscapes with perfectly-correlated species pairs, I used a regularized estimate based on James-Stein

shrinkage, as implemented in the corpcor package’s `pcor.shrink` function with the default settings (Schäfer et al. 2014).

The third alternative, generalized linear models (GLMs), also provide a computationally efficient approximation to the Markov network (Lee and Hastie 2012). Following Faisal et al. (2010), I fit regularized logistic regression models (Gelman et al. 2008) for each species, using the other species on the landscape as predictors. This produced two interaction estimates for each species pair; one for the effect of species i on species j and one for the reverse. These two coefficients are not identifiable from the data, however (Schmidt and Murphy 2012), so I used their average as an overall measure of the overall relationship between species i and species j .

The next method, described in Gotelli and Ulrich (2009), involved simulating new landscapes from a null model that retains the row and column sums of the original matrix (Strong et al. 1984). I used the Z -scores computed by the Pairs software described in Gotelli and Ulrich (2009) as my null model-based estimator of species interactions.

The last two estimators used the latent correlation matrix estimated by the BayesComm package (Golding and Harris 2015) in order to evaluate the recent claim that the correlation coefficients estimated by “joint species distribution models” provide an accurate assessment of species’ pairwise interactions (Pollock et al. 2014, see also Harris 2015b). In addition to using the posterior mean correlation (Pollock et al. 2014), I also used the posterior mean *partial* correlation, which might be able to control for indirect effects.

Evaluating model performance. For the simulated landscapes based on Figure 1, I assessed whether each method’s test statistic indicated a positive or negative relationship between the two shrubs (Appendix 1). For the null model (Pairs), I calculated statistical significance using its Z -score. For the Markov network, I used the Hessian matrix to generate approximate confidence intervals and noted whether these intervals included zero.

I then evaluated the relationship between each method’s estimates and the “true” interaction strengths among all of the species pairs from the larger simulated landscapes. This determined which of the methods provide a consistent way to know how strong species interactions are—regardless of which species were present in a particular data set or how many observations were taken. To ensure that the different test statistics (e.g. correlations versus Z scores) were on a common scale, I rescaled them using linear regression through the origin. After rescaling each method’s estimates, I calculated squared errors between the scaled interaction estimates and “true” interaction values across the simulated data sets. These squared errors determined the proportion of variance explained for different combinations of model type and landscape size (compared with a null model that assumed all interaction strengths to be zero).

Results

Three species. As shown in Figure 1, the marginal relationship between the two shrub species was positive—despite their competition for space at a mechanistic level—due to indirect effects of the dominant tree species. As a result, the correlation between these species was positive in 94% of replicates, and the randomization-based null model falsely reported positive associations 100% of the time. Worse, more than 98% of these false conclusions were statistically significant. The partial correlation and Markov network estimates, on the other hand, each correctly isolated the direct negative interaction between the shrubs from their positive indirect interaction 94% of the time (although the confidence intervals overlapped zero in most replicates).

Twenty species. Despite some variability across contexts (Figure 3), the four methods that controlled for indirect effects (the Markov network, the generalized linear models, and the two partial correlation-based methods) outperformed those that did not. Of these, the Markov

network consistently performed the best, especially on the smaller landscapes (which are also the most representative of typical analyses in this field; Gotelli and Ulrich 2009). As shown in Appendix [[X]], the standard errors associated with the estimates in Figure 3 are all small (less than 0.01), so differences across methods should not be attributed to sampling error.

Of the methods that did not control for indirect effects, Figure 3 shows that simple correlation coefficients provided a more reliable indicator of species’ true interaction strengths than either the joint species distribution model (BayesComm) or the null model (Pairs). The estimates from these three approaches were tightly correlated (Figure 3B), suggesting that the effect sizes from the null model (and, to a lesser extent, the correlation matrices from joint species distribution models) only contain noisy versions of the same information that could be obtained more easily and interpretably by calculating correlation coefficients between species’ presence-absence vectors.

Whereas Figure 3 summarizes the accuracy of each model’s point estimates, Figure 4 compares the inferential value of the Markov network versus the null model. When the “true” β coefficient is near zero, the Markov network’s 95% confidence intervals usually include zero (but not quite 95% of the time), and it is rare for the model to confidently attribute the wrong sign to a coefficient (Figure 4A). The null model, by contrast, typically rejects the null hypothesis at the $p = 0.05$ level regardless of the underlying parameters, and the sign of its Z -scores is not a reliable indicator of the interaction’s actual sign.

Discussion

The results presented above show that Markov networks can reliably recover species’ pairwise interactions from observational data, even for cases where environmental heterogeneity and indirect interactions cause ecologists’ typical null modeling approaches to reliably fail. For cases where fitting a Markov network is computationally infeasible, these results also indicate

that partial covariances and generalized linear models can both provide useful approximations. The partial correlations' success on simulated data may not carry over to real data sets, however; Loh and Wainwright (2013) show that the linear approximations can be less reliable in cases where the true interaction matrix contains more structure (e.g. guilds or trophic levels). Similarly, the approximation involved in using separate generalized linear models for each species can occasionally lead to catastrophic overfitting with small-to-moderate sample sizes (Lee and Hastie 2012). For these reasons, it will usually be best to fit a Markov network rather than one of the alternative methods when one's computational resources allow it.

It's important to note that none of these methods can identify the exact nature of the pairwise interactions (e.g. which species in a positively-associated pair is facilitating the other; Schmidt and Murphy 2012), particularly when real pairs of species can reciprocally influence one another in multiple ways simultaneously (Bruno et al. 2003); with compositional data, there is only enough information to provide a single number describing each species pair. To estimate asymmetric interactions, such as commensalism or predation, ecologists would need other kinds of data, as from time series, behavioral observations, manipulative experiments, or natural history. These other sources of information could also be used to augment the likelihood function with an informative prior distribution, which could lead to better results on some real data sets than was shown in Figure 3A.

Despite their limitations, Markov networks have enormous potential to improve ecological understanding. In particular, they are less vulnerable than some of the most commonly-used methods to mistakenly identifying positive species interactions between competing species, and can make precise statements about the conditions where indirect interactions will overwhelm direct ones. They also provide a simple answer to the question of how competition should affect a species' overall prevalence, which was a major flashpoint for

the null model debates in the 1980's (Roughgarden 1983, Strong et al. 1984). Equation 1 can be used to calculate the expected prevalence of a species in the absence of biotic influences ($\frac{e^\alpha}{1+e^\alpha}$; Lee and Hastie 2012). Competition's effect on prevalence in a Markov network can then be calculated by subtracting this value from the observed prevalence (cf Figure 2D). This kind of insight would have been difficult to obtain without a generative model that makes predictions about the consequences of species interactions; null models (which presume *a priori* that interactions do not exist) have no way to make such predictions.

Markov networks—particularly the Ising model for binary networks—have been studied for nearly a century (Cipra 1987), and the models' properties, capabilities, and limits are well-understood. Using the same framework for species interactions would thus allow ecologists to tap into an enormous set of existing discoveries and techniques for dealing with indirect effects, stability, and alternative stable states. Numerous other extensions are possible: for example, the states of the interaction network can be modeled as a function of the local abiotic environment (Lee and Hastie 2012), which would provide a rigorous and straightforward approach to the difficult and important task of incorporating whole networks of biotic interactions into species distribution models (Pollock et al. 2014), leading to a better understanding of the interplay between biotic and abiotic effects on community structure. Alternatively, models could allow one species to affect the interactions between two other species (Tjelmeland and Besag 1998), tipping the balance between facilitation and exploitation (Bruno et al. 2003).

Finally, the results presented here have important implications for ecologists' continued use of null models for studying species interactions. Null and neutral models can be useful for clarifying our thinking about the numerical consequences of species' richness and abundance patterns (Harris et al. 2011, Xiao et al. 2015), but deviations from a particular null model

must be interpreted with care (Roughgarden 1983). Even in small networks with three species, it may simply not be possible to implicate individual species pairs or specific ecological processes like competition by rejecting a general-purpose null (Gotelli and Ulrich 2009), especially when the test statistic is effectively just a correlation coefficient [[Figure 3B]]. When a non-null backdrop is not controlled for, Type I error rates can skyrocket, the apparent sign of the interaction can change, and null models can lead to reliably incorrect inferences (Figure 1, Figure 4).

Controlling for indirect effects via simultaneous estimation of multiple ecological parameters seems like a much more promising approach: to the extent that the models' relative performance on real data sets is similar to the range of results shown in Figure 3A, scientists in this field could more than double their explanatory power by switching from null models to Markov networks (or increase it nearly as much with linear or generalized linear approximations). Regardless of the methods ecologists ultimately choose, controlling for indirect effects could clearly improve our understanding of species' direct effects on one another and on community structure.

Acknowledgements: This research was funded by a Graduate Research Fellowship from the US National Science Foundation and benefited greatly from discussions with A. Sih, M. L. Baskett, R. McElreath, R. J. Hijmans, A. C. Perry, and C. S. Tysor. Additionally, A. K. Barner, E. Baldrige, E. P. White, D. Li, D. L. Miller, N. Golding, N. J. Gotelli, C. F. Dormann, and two anonymous reviewers provided very helpful feedback on the text.

Figure 1: A. A small network of three competing species. The tree (top) tends not to co-occur with either of the two shrub species, as indicated by the strongly negative coefficient linking them. The two shrub species also compete with one another, but more weakly (circled coefficient). **B.** In spite of the competitive interactions between the two shrub species, their

shared tendency to occur in locations without trees makes their occurrence vectors positively correlated (circled). **C.** Controlling for the tree species' presence with a conditional (all-else-equal) approach such as a partial covariance or a Markov network leads to correct identification of the negative shrub-shrub interaction (circled).

Figure 2: A. A small Markov network with two species, defined by its α and β values. The abiotic environment favors the occurrence of each species ($\alpha > 0$), particularly species 2 ($\alpha_2 > \alpha_1$). The negative β coefficient linking these two species implies that they co-occur less than expected under independence. **B.** These coefficients determine the probabilities of all four possible presence-absence combinations for Species 1 and Species 2. α_1 is added to the exponent whenever Species 1 is present ($y_1 = 1$), but not when it is absent ($y_1 = 0$). Similarly, the exponent includes α_2 only when species 2 is present ($y_2 = 1$), and β only when both are present ($y_1 y_2 = 1$). The normalizing constant Z , ensures that the four probabilities sum to 1. In this case, Z is about 18.5. **C.** The expected frequencies of all possible co-occurrence patterns between the two species of interest, as calculated in the previous panel. **D.** If, as assumed in null model simulations, β_{12} equaled zero (e.g. if the species no longer competed for the same resources), then the reduction in competition would allow each species to increase its overall occurrence rate.

Figure 3: A. Proportion of variance in interaction coefficients explained by each method versus number of sampled locations. **B.** The Z -scores produced by the null model ("Pairs") for each pair of species can be predicted using the correlation between the presence-absence vectors of those same species and from the number of sites on the landscape.

Figure 4: A. Inferences from the Markov network's 95% confidence intervals on the "no-environment" simulated landscapes, as a function of the "true" coefficient values. **B.** Inferences from the Z -statistics estimated by the "Pairs" software for null modeling.

References:

- Albrecht, M., and N. J. Gotelli. 2001. Spatial and temporal niche partitioning in grassland ants. *Oecologia* 126:134–141.
- Azaele, S., R. Muneeppeerakul, A. Rinaldo, and I. Rodriguez-Iturbe. 2010. Inferring plant ecosystem organization from species occurrences. *Journal of theoretical biology* 262:323–329.
- Bruno, J. F., J. J. Stachowicz, and M. D. Bertness. 2003. Inclusion of facilitation into ecological theory. *Trends in Ecology & Evolution* 18:119–125.
- Cipra, B. A. 1987. An introduction to the Ising model. *American Mathematical Monthly* 94:937–959.
- Connor, E. F., M. D. Collins, and D. Simberloff. 2013. The checkered history of checkerboard distributions. *Ecology* 94:2403–2414.
- Diamond, J. M. 1975. The island dilemma: Lessons of modern biogeographic studies for the design of natural reserves. *Biological conservation* 7:129–146.
- Dodson, S. I. 1970. COMPLEMENTARY FEEDING NICHES SUSTAINED BY SIZE-SELECTIVE PREDATION. *Limnology and Oceanography* 15:131–137.
- Faisal, A., F. Dondelinger, D. Husmeier, and C. M. Beale. 2010. Inferring species interaction networks from species abundance data: A comparative evaluation of various statistical and machine learning methods. *Ecological Informatics* 5:451–464.
- Friedman, J., T. Hastie, and R. Tibshirani. 2008. Sparse inverse covariance estimation with the graphical lasso. *Biostatistics* 9:432–441.
- Gelfand, A. E., A. M. Schmidt, S. Wu, J. A. Silander, A. Latimer, and A. G. Rebelo. 2005. Modelling species diversity through species level hierarchical modelling. *Journal of the Royal Statistical Society: Series C (Applied Statistics)* 54:1–20.
- Gelman, A., A. Jakulin, M. G. Pittau, and Y.-S. Su. 2008. A Weakly Informative Default

- 336 Prior Distribution for Logistic and Other Regression Models. *The Annals of Applied*
337 *Statistics* 2:1360–1383.
- 338 Gilpin, M. E., and J. M. Diamond. 1982. Factors contributing to non-randomness in species
339 Co-occurrences on Islands. *Oecologia* 52:75–84.
- 340 Golding, N., and D. J. Harris. 2015. BayesComm: Bayesian Community Ecology Analysis.
- 341 Gotelli, N. J., and W. Ulrich. 2009. The empirical Bayes approach as a tool to identify
342 non-random species associations. *Oecologia* 162:463–477.
- 343 Harris, D. J. 2015a. Rosalia: Exact inference for small binary Markov networks. R package
344 version 0.1.0. Zenodo. <http://dx.doi.org/10.5281/zenodo.17808>.
- 345 Harris, D. J. 2015b. Generating realistic assemblages with a Joint Species Distribution Model.
346 *Methods in Ecology and Evolution*.
- 347 Harris, D. J., K. G. Smith, and P. J. Hanly. 2011. Occupancy is nine-tenths of the law:
348 Occupancy rates determine the homogenizing and differentiating effects of exotic species.
349 *The American naturalist* 177:535.
- 350 Harris, T. E. 1974. Contact Interactions on a Lattice. *The Annals of Probability* 2:969–988.
- 351 Lee, J. D., and T. J. Hastie. 2012, May. Learning Mixed Graphical Models.
- 352 Levine, S. H. 1976. Competitive Interactions in Ecosystems. *The American Naturalist*
353 110:903–910.
- 354 Lewin, R. 1983. Santa Rosalia Was a Goat. *Science* 221:636–639.
- 355 Loh, P.-L., and M. J. Wainwright. 2013. Structure estimation for discrete graphical models:
356 Generalized covariance matrices and their inverses. *The Annals of Statistics* 41:3022–3049.
- 357 MacArthur, R. H. 1958. Population ecology of some warblers of northeastern coniferous
358 forests. *Ecology* 39:599–619.
- 359 Murphy, K. P. 2012. Machine Learning: A Probabilistic Perspective. The MIT Press.

- 360 Pollock, L. J., R. Tingley, W. K. Morris, N. Golding, R. B. O'Hara, K. M. Parris, P. A. Vesk,
361 and M. A. McCarthy. 2014. Understanding co-occurrence by modelling species
362 simultaneously with a Joint Species Distribution Model (JSDM). *Methods in Ecology and*
363 *Evolution*:n/a–n/a.
- 364 R Core Team. 2015. R: A Language and Environment for Statistical Computing. R
365 Foundation for Statistical Computing, Vienna, Austria.
- 366 Roughgarden, J. 1983. Competition and Theory in Community Ecology. *The American*
367 *Naturalist* 122:583–601.
- 368 Schäfer, J., R. Opgen-Rhein, V. Zuber, M. Ahdesmäki, A. P. D. Silva, and K. Strimmer. 2014.
369 Corpcor: Efficient Estimation of Covariance and (Partial) Correlation.
- 370 Schmidt, M., and K. Murphy. 2012. Modeling Discrete Interventional Data using Directed
371 Cyclic Graphical Models. arXiv preprint arXiv:1205.2617.
- 372 Strong, D. R., D. Simberloff, L. G. Abele, and A. B. Thistle. 1984. Ecological communities:
373 Conceptual issues and the evidence. Princeton University Press.
- 374 Tjelmeland, H., and J. Besag. 1998. Markov Random Fields with Higher-order Interactions.
375 *Scandinavian Journal of Statistics* 25:415–433.
- 376 Veech, J. A. 2013. A probabilistic model for analysing species co-occurrence. *Global Ecology*
377 *and Biogeography* 22:252–260.
- 378 Whittam, T. S., and D. Siegel-Causey. 1981. Species Interactions and Community Structure
379 in Alaskan Seabird Colonies. *Ecology* 62:1515–1524.
- 380 Xiao, X., D. J. McGlinn, and E. P. White. 2015. A strong test of the Maximum Entropy
381 Theory of Ecology. *The American Naturalist* 185:E70–E80.