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

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

Abstract: Inferring 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 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.

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

1

24 major controversies, however (Lewin 1983, Strong et al. 1984, Connor et al. 2013), existing

25 methods for detecting competition's effects on community structure are unreliable (Gotelli

26 and Ulrich 2009). In particular, species' effects on one another can become lost in a web of

27 indirect effects. For example, the competitive interaction between the two shrub species in

28 Figure 1A is obscured by their shared tendency to occur in unshaded areas (Figure 1B).

29 While ecologists have long known that indirect effects can overwhelm direct ones (Levine

30 1976), most methods for drawing inferences from co-occurrence data do not control for these

31 effects (e.g. Diamond 1975, Strong et al. 1984, Gotelli and Ulrich 2009, Veech 2013, Pollock

32 et al. 2014). As a result, ecologists do not have tools that allow them to isolate direct

33 interactions from indirect effects.

34   While competition doesn't reliably reduce co-occurrence rates at the whole-landscape

35 level (as most methods assume), it does still leave a signal in the data (Figure 1C). After

36 controlling for the presence of the tree species (e.g. by splitting the data set into shaded and

37 unshaded sites or by using a model that estimates conditional relationships amongs species),

38 the two shrubs do have a negative association, with fewer co-occurrences in unshaded areas

39 than would be expected if they occurred independently of one another.

40   Following Azaele et al. (2010), this paper shows that Markov networks (undirected

41 graphical models also known as Markov random fields; Murphy 2012) can provide a

42 framework for understanding the landscape-level consequences of pairwise species

43 interactions, and for estimating them from observed presence-absence matrices. Markov

44 networks have been used in diverse scientific fields in similar contexts for decades, from

45 physics (where nearby particles interact magnetically; Cipra 1987) to spatial statistics (where

46 adjacent grid cells have correlated values; Harris 1974, Gelfand et al. 2005). While

47 community ecologists explored some related approaches in the 1980's (Whittam and

48 Siegel-Causey 1981), they used severe approximations that led to unintelligible results (e.g.

49 "probabilities" greater than one; Gilpin and Diamond 1982).

50 Below, I demonstrate Markov networks' ability to produce exact predictions about the

51 direct and indirect consequences of an interaction matrix, and also to make inferences about

52 the interaction matrix based on co-occurrence rates. Using simulated data sets where the

53 "true" interactions are known, I compare this approach with several existing methods.

54 Finally, I discuss opportunities for extending the approach presented here to other problems

55 in community ecology, e.g. quantifying the overall effect of species interactions on occurrence

56 rates (Roughgarden 1983) and disentangling the effects of biotic versus abiotic interactions

57 on species composition (Pollock et al. 2014).

58 **Methods**

59 **Markov networks.** Markov networks provide a framework for translating back and forth

60 between the conditional (all-else-equal) relationships among species (Figure 1C) and the

61 kinds of species assemblages that these relationships produce. Here, I show how a set of

62 conditional relationships can determine species composition. Methods for estimating

63 conditional relationships from data are discussed in the next section.

64 A Markov network defines the relative probability of observing a given vector of

65 species-level presences (1s) and absences (0s), $\vec{y}$ at a site, as

66 $$p(\vec{y}; \alpha, \beta) \propto exp(\sum_i \alpha_i y_i + \sum_{<ij>} \beta_{ij} y_i y_j),$$

67 where the second sum is over all $\frac{1}{2}n(n-1)$ pairs of n species. In this model, α_i is an

68 intercept term determining the amount that the presence of species i contributes to the

69 log-probability of $\vec{y}$; it directly controls the prevalence of species i. Similarly, β_{ij} is the

70 amount that the co-occurrence of species i and species j contributes to the log-probability; it

3

⁷¹ controls the conditional relationship between two species, i.e. the probability that they will

⁷² be found together, after controlling for the other species in the network (Figure 2A, Figure

⁷³ 2B). For example, if $\beta_{ij} = +2$, then each species' odds of occurrence would be e^2 times

⁷⁴ higher when the other one is present (as compared with otherwise equivalent sites). The

⁷⁵ relative probability of a presence-absence vector increases when positively-associated species

⁷⁶ co-occur and decreases when negatively-associated species co-occur. As a result, the model

⁷⁷ tends—all else equal—to produce assemblages where many positively-associated species pairs

⁷⁸ co-occur and few negatively-associated pairs do (just as an ecologist might expect). When all

⁷⁹ else is *not* equal (e.g. Figure 1, where the presence of one competitor is associated with

⁸⁰ release from another competitor), then predicting species' overall co-occurrence rates would

⁸¹ require normalizing the relative probabilities of possible $\vec{y}$ vectors so that they sum to one

⁸² (Figure 2B).

⁸³ **Estimating α and β coefficients from presence-absence data.** In the previous

⁸⁴ section, the values of α and β were known and the goal was to make predictions about

⁸⁵ possible species assemblages. In practice, however, ecologists will often need to estimate the

⁸⁶ parameters from an observed co-occurrence matrix (i.e. from a set of independent $\vec{y}$ vectors

⁸⁷ indicating which species are present at each site on the landscape). When Equation 1 can be

⁸⁸ normalized, one can find exact maximum likelihood estimates for α and β using numerical

⁸⁹ optimization. Fully-observed networks like the ones considered here have unimodal likelihood

⁹⁰ surfaces (Murphy 2012), so these estimates will represent the global optimum. I used the

⁹¹ rosalia package (Harris 2015a) for the R programming language (R Core Team 2015) to

⁹² calculate $p(\vec{y}; \alpha, \beta)$ and its gradients (Murphy 2012); the package passes these functions to

⁹³ the "BFGS" method in R's general-purpose optimizer, which then estimates the Markov

⁹⁴ network parameters.

95 **Simulated landscapes.** I simulated several sets of landscapes using known parameters to

96 evaluate different statistical methods' performance. The first set of landscapes included the

97 three competing species shown in Figure 1. For each of 1000 replicates, I generated a

98 landscape with 100 sites by sampling from a probability distribution defined by the figure's

99 interaction coefficients (Appendix 1). Each of the methods described below was then

100 evaluated on its ability to correctly infer that the two shrub species competed with one

101 another, despite their frequent co-occurrence.

102 I then simulated landscapes with up to 20 interacting species at 25, 200, or 1600 sites

103 using three increasingly complex models (50 replicates for each combination of size and

104 model). The simplest set of simulated landscapes were generated with Gibbs samples from

105 Equation 1, and provided a test of methods' ability to isolate direct species-level interactions

106 from a web of indirect ones. For each replicate, I randomly drew the "true" β coefficient

107 magnitudes an exponential distribution with rate 1 so that most species pairs interacted

108 negligibly but a few interactions were strong enough to propagate through the newtwork. I

109 randomly assigned 25% of the interactions to be positive; the remainder were negative.

110 The next set of landscapes provided a way to assess each method's ability to identify

111 direct interactions in the presence of environmental heterogeneity. Here, the β coefficients

112 were calculated as above, but each species' α value depended linearly on two

113 normally-distributed environmental factors, which were drawn from independent Gaussians

114 for each site. Once the local α values were calculated, independent Gibbs samplers

115 determined the species composition for each site based on Equation 1.

116 In the third set of landscapes, I simulated each species' abundance (instead of just their

117 presence/absence), and interactions between species occurred on a per-capita basis. To avoid

118 runaway mutualisms leading to infinite abundance, all interaction coefficients were negative

5

119 in these simulations. If a method performs well on these landscapes, it would indicate some

120 robustness to the biolocical mechanisms underlying species' interactiosn with one another.

121 **Recovering species interactions from simulated data.** I compared seven techniques

122 for determining the sign and strength of the associations between pairs of species from

123 simulated data (Appendix 3). First, I used the rosalia package (Harris 2015a) to fit Markov

124 network models, as described above. For the analyses with 20 species, a weak regularizer

125 (equivalent to a logistic prior with scale 2) ensured that the estimates were always finite.

126 I also evaluated six alternative methods: five from the existing literature, plus a novel

127 combination of two of these methods. The first alternative interaction metric was the sample

128 correlation between species' presence-absence vectors, which summarizes their marginal

129 association. Next, I used partial correlations, which summarize species' conditional

130 relationships. This approach, which is closely related to linear regression, is common in

131 molecular biology (Friedman et al. 2008), but is rare in ecology (see Albrecht and Gotelli

132 (2001) and Faisal et al. (2010) for two exceptions). In the context of non-Gaussian data, the

133 partial correlation can be thought of as a computationally efficient approximation to the full

134 Markov network model (Loh and Wainwright 2013). Partial correlations are undefined for

135 landscapes with perfectly-correlated species pairs, so I used the regularized estimate

136 provided by the corpcor package's `pcor.shrink` function with the default settings (Schäfer

137 et al. 2014).

138 The third alternative, generalized linear models (GLMs), also provide a computationally

139 efficient approximation to the Markov network (Lee and Hastie 2012). Following Faisal et al.

140 (2010), I fit regularized logistic regression models (Gelman et al. 2008) for each species, using

141 the other species on the landscape as predictors. This produced two interaction estimates for

142 each species pair; one for the effect of species i on species j and one for the reverse. These

6

143 two coefficients are not identifiable from the data, however (Schmidt and Murphy 2012), so I

144 used their average as an overall measure of the pairwise relationship.

145   The next method used the Pairs software described in Gotelli and Ulrich (2009). This

146 program simulating new landscapes from a null model that retains the row and column sums

147 of the original matrix (Strong et al. 1984) and calculates Z-scores to summarize a species

148 pair's deviation from this null.

149   The last two estimators used the latent correlation matrix estimated by the BayesComm

150 package (Golding and Harris 2015) in order to evaluate the recent claim that the correlation

151 coefficients estimated by "joint species distribution models" provide an accurate assessment

152 of species' pairwise interactions (Pollock et al. 2014, see also Harris 2015b). In addition to

153 using the posterior mean correlation (Pollock et al. 2014), I also used the posterior mean

154 *partial* correlation, which should control better for indirect effects.

155 **Evaluating model performance.** For the simulated landscapes based on Figure 1, I

156 assessed whether each method's test statistic indicated a positive or negative relationship

157 between the two shrubs (Appendix 1). For the null model (Pairs), I calculated statistical

158 significance using its Z-score. For the Markov network, I used the Hessian matrix to

159 generate approximate confidence intervals.

160   For the larger landscapes, I evaluated the relationship between each method's estimates

161 and the "true" interaction strengths. To ensure that the different test statistics

162 (e.g. correlations versus Z scores) were on a common scale, I rescaled them using linear

163 regression through the origin. I then calculated the proportion of variance explained for

164 different combinations of model type and landscape size (compared with a baseline model

165 that assumed all interaction strengths to be zero).

166   For the null model and the Markov network, the probability of rejecting the null

167 hypothesis of zero interaction was estimated across a range of "true" interaction strengths

168 using a generalized additive model with binomial responses (reject versus not; Appendix 4).

169 The Type I error rate was defined as this model's estimated probability of rejecting the null

170 hypothesis when the "true" interaction was zero.

171 **Results**

172 **Three species.** As shown in Figure 1, the marginal relationship between the two shrub

173 species was positive—despite their competition for space at a mechanistic level—due to

174 indirect effects of the dominant tree species. As a result, the correlation between these

175 species was positive in 94% of replicates, and the randomization-based null model falsely

176 reported positive associations 100% of the time. Worse, more than 98% of these false

177 conclusions were statistically significant. The partial correlation and Markov network

178 estimates, on the other hand, each correctly isolated the direct negative interaction between

179 the shrubs from their positive indirect interaction 94% of the time (although the confidence

180 intervals overlapped zero in most replicates).

181 **Twenty species.** In general, each model's performance was highest for large landscapes

182 with simple assembly rules and no environmental heterogeneity (Figure 3). Despite some

183 variability across contexts, the rank ordering across methods was very consistent. In

184 particular, the four methods that controlled for indirect effects (the Markov network, the

185 generalized linear models, and the two partial correlation-based methods) always matched or

186 outperformed those that did not. The Markov network consistently performed best of all. As

187 anticipated by Lee and Hastie (2012), generalized linear models closely approximated the

188 Markov network estimates (Figure 4A), especially when the data sets were very large (Figure

189 3). As reviewed in Gotelli and Ulrich (2009), however, most analyses in this field of ecology

8

190 involve fewer than 50 sites; in this context, the gap between the methods was larger. As

191 shown in Appendix 4, the standard errors associated with the estimates in Figure 3 are small

192 (less than 0.01), so the differences among methods should not be attributed to sampling error.

193 Of the methods that did not control for indirect effects, Figure 3 shows that simple

194 correlation coefficients provided a more reliable indicator of species' true interaction strengths

195 than either the joint species distribution model (BayesComm) or the null model (Pairs). 95%

196 of the variance the Pairs test statistic was explained by correlation coefficients (controlling

197 for landscape size; Figure 4B); much of the remaining variance is due to sampling error.

198 Finally, we can evaluate the models' statistical inferences (focusing on the first two

199 simulation types, where the interaction coefficients are easiest to interpret). The Markov

200 network's Type I error rate depended on the simulation type: [[0.02]] for simulations that

201 matched the model's assumptions, versus [[0.14]] for simulations that included environmental

202 heterogeneity (see Appendix 4 for confidence interval coverage across a range of β_{ij} values).

203 The null model's Type I error rates were [[0.30]] and [[0.50]] for the constant and

204 heterogeneous landscapes, respectively—far higher than the nominal 0.05 rate. Figure 4C

205 shows, across a range of true interaction strengths, the probability that the null model or the

206 Markov network will predict the wrong sign of the interaction with 95% confidence. The null

207 model makes such errors more than [[5]] times as often as the Markov network, even though

208 it only rejects the null hypothesis [[twice]] as often overall (Appendix 4). The Markov

209 network's errors were also more concentrated around 0, so it never misclassified strong

210 interactions like the null model did.

211 **Discussion**

212 The results presented above show that Markov networks can reliably recover species'

213 pairwise interactions from species composition data, even for cases where environmental

9

heterogeneity and indirect interactions cause ecologists' typical null modeling approaches to reliably fail. Partial covariances and generalized linear models can both provide computationally efficient approximations, but with somewhat lower accuracy (especially for typically-sized data sets with small numbers of sites; Gotelli and Ulrich 2009). The difference in accuracy may be larger for real data sets than for the simulated landscapes in Figure 3, however; linear approximations to the Markov network make larger errors when the interaction matrix is structured (e.g. due to guilds or trophic levels; Loh and Wainwright 2013). Similarly, the separate generalized linear models for each species can severely overfit in some cases (Lee and Hastie 2012). The full Markov network should thus be preferred to the approximations when it is computationally tractable.

Compositional data only contains enough degrees of freedom to estimate one interaction per species pair (Schmidt and Murphy 2012), so 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). To estimate asymmetric interactions, such as commensalism or predation, ecologists could use time series, behavioral observations, manipulative experiments, or natural history. These other sources of information could also be used to augment the likelihood function with a more informative prior distribution, reducing ecologists' error and uncertainty relative to Figure 3's results.

Despite their limitations, Markov networks have enormous potential to improve ecological understanding. In particular, they make many fewer errors than existing approaches, 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 has important implications for community-level modeling (Strong et al. 1984). Specifically, Equation 1 can be used to

238 calculate the expected prevalence of a species in the absence of biotic influences as

239 $e^\alpha/(e^0 + e^\alpha)$. Competition's effect on prevalence can then be estimated by comparing this

240 value with the observed prevalence (e.g. comparing Figure 2D with Figure 2C). This novel

241 quantitative result conflicts with most of our null models, which unreasonably assume that

242 prevalence would be the exactly same in the absence of competition as it is in the observed

243 data (Roughgarden 1983).

244   Markov networks—particularly the Ising model for binary networks—are very well

245 understood, having been studied for nearly a century (Cipra 1987). Tapping into this

246 framework would thus allow ecologists to take advantage of into a vast set of existing

247 discoveries and techniques for dealing with indirect effects, stability, and alternative stable

248 states. Numerous extensions to the basic network are possible as well. For example, the

249 states of the interaction network can be modeled as a function of the local abiotic

250 environment (Lee and Hastie 2012), which would help incorporate networks of biotic

251 interactions into species distribution models (Pollock et al. 2014) and lead to a better

252 understanding of the interplay between biotic and abiotic effects on community structure.

253 Alternatively, models could allow one species to alter the relationship between two other

254 species (Tjelmeland and Besag 1998, cf Bruno et al. 2003).

255   Finally, the results presented here have important implications for ecologists' continued

256 use of null models for studying species interactions. Null and neutral models can be useful

257 for clarifying our thinking (Harris et al. 2011, Xiao et al. 2015), but deviations from a

258 particular null model must be interpreted with care (Roughgarden 1983). Even in small

259 networks with three species, it may simply not be possible to implicate specific ecological

260 processes like competition by rejecting a general-purpose null (Gotelli and Ulrich 2009),

261 especially when the test statistic is effectively just a correlation coefficient (Figure 4B).

When the 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 routinely produce
misleading inferences (Figure 1, Figure 4C, Gotelli and Ulrich (2009)).

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 3,
scientists in this field could often triple 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.

References:

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

12

occurrence vectors positively correlated (circled). **C.** Controlling for trees 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). See

Appendix 1 and the results for "three species" for more details.

Figure 2: A. A small Markov network, 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_{12} coefficient is consistent with competition between the two species. **B.**

The 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 includes β_{12} 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.**

Without competition (i.e. with $\beta_{12} = 0$, each species would occur more often.

Figure 3: Proportion of variance in interaction coefficients explained by each method versus

number of sampled locations across the three simulation types. For the null model (Pairs),

two outliers with $|Z| > 1000$ were manually adjusted to $|Z| = 50$ to mitigate their

detrimental influence on R^2 (Appendix 5).

Figure 4: A. The Markov network's estimated interaction coefficients were generally very

similar to the GLM estimates. **B.** The null model's estimates typically matched the

(negative) correlation coefficient, after controlling for landscape size. **C.** For any given

interaction strength, the null model was much more likely to misclassify its sign with 95%

confidence than the Markov network was.

Albrecht, M., and N. J. Gotelli. 2001. Spatial and temporal niche partitioning in grassland ants. Oecologia 126:134–141.

Azaele, S., R. Muneepeerakul, 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.

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 Prior Distribution for Logistic and Other Regression Models. The Annals of Applied Statistics 2:1360–1383.

334 Gilpin, M. E., and J. M. Diamond. 1982. Factors contributing to non-randomness in species

335   Co-occurrences on Islands. Oecologia 52:75–84.

336 Golding, N., and D. J. Harris. 2015. BayesComm: Bayesian Community Ecology Analysis.

337 Gotelli, N. J., and W. Ulrich. 2009. The empirical Bayes approach as a tool to identify

338   non-random species associations. Oecologia 162:463–477.

339 Harris, D. J. 2015a. Rosalia: Exact inference for small binary Markov networks. R package

340   version 0.1.0. Zenodo. http://dx.doi.org/10.5281/zenodo.17808.

341 Harris, D. J. 2015b. Generating realistic assemblages with a Joint Species Distribution

342   Model. Methods in Ecology and Evolution.

343 Harris, D. J., K. G. Smith, and P. J. Hanly. 2011. Occupancy is nine-tenths of the law:

344   Occupancy rates determine the homogenizing and differentiating effects of exotic species.

345   The American naturalist 177:535.

346 Harris, T. E. 1974. Contact Interactions on a Lattice. The Annals of Probability 2:969–988.

347 Lee, J. D., and T. J. Hastie. 2012, May. Learning Mixed Graphical Models.

348 Levine, S. H. 1976. Competitive Interactions in Ecosystems. The American Naturalist

349   110:903–910.

350 Lewin, R. 1983. Santa Rosalia Was a Goat. Science 221:636–639.

351 Loh, P.-L., and M. J. Wainwright. 2013. Structure estimation for discrete graphical models:

352   Generalized covariance matrices and their inverses. The Annals of Statistics 41:3022–3049.

353 MacArthur, R. H. 1958. Population ecology of some warblers of northeastern coniferous

354   forests. Ecology 39:599–619.

355 Murphy, K. P. 2012. Machine Learning: A Probabilistic Perspective. The MIT Press.

356 Pollock, L. J., R. Tingley, W. K. Morris, N. Golding, R. B. O'Hara, K. M. Parris, P. A. Vesk,

357   and M. A. McCarthy. 2014. Understanding co-occurrence by modelling species

358 simultaneously with a Joint Species Distribution Model (JSDM). Methods in Ecology and

359 Evolution:n/a–n/a.

360 R Core Team. 2015. R: A Language and Environment for Statistical Computing. R

361 Foundation for Statistical Computing, Vienna, Austria.

362 Roughgarden, J. 1983. Competition and Theory in Community Ecology. The American

363 Naturalist 122:583–601.

364 Schäfer, J., R. Opgen-Rhein, V. Zuber, M. Ahdesmäki, A. P. D. Silva, and K. Strimmer.

365 2014. Corpcor: Efficient Estimation of Covariance and (Partial) Correlation.

366 Schmidt, M., and K. Murphy. 2012. Modeling Discrete Interventional Data using Directed

367 Cyclic Graphical Models. arXiv preprint arXiv:1205.2617.

368 Strong, D. R., D. Simberloff, L. G. Abele, and A. B. Thistle. 1984. Ecological communities:

369 Conceptual issues and the evidence. Princeton University Press.

370 Tjelmeland, H., and J. Besag. 1998. Markov Random Fields with Higher-order Interactions.

371 Scandinavian Journal of Statistics 25:415–433.

372 Veech, J. A. 2013. A probabilistic model for analysing species co-occurrence. Global Ecology

373 and Biogeography 22:252–260.

374 Whittam, T. S., and D. Siegel-Causey. 1981. Species Interactions and Community Structure

375 in Alaskan Seabird Colonies. Ecology 62:1515–1524.

376 Xiao, X., D. J. McGlinn, and E. P. White. 2015. A strong test of the Maximum Entropy

377 Theory of Ecology. The American Naturalist 185:E70–E80.