

1 *Running head:* Species interactions in Markov networks

2 **Title:** Estimating species interactions from observational data with Markov
3 networks

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

5 **Abstract**

6 Inferring species interactions from observational data is one of the most controversial tasks in
7 community ecology. One difficulty is that a single pairwise interaction can ripple through an
8 ecological network and produce surprising indirect consequences. For example, two
9 competing species would ordinarily correlate negatively in space, but this effect can be
10 reversed in the presence of a third species that is capable of outcompeting both of them
11 when it is present. Here, I apply models from statistical physics, called Markov networks or
12 Markov random fields, that can predict the direct and indirect consequences of any possible
13 species interaction matrix. Interactions in these models can be estimated from observational
14 data via maximum likelihood. Using simulated landscapes with known pairwise interaction
15 strengths, I evaluated Markov networks and several existing approaches. The Markov
16 networks consistently outperformed other methods, correctly isolating direct interactions
17 between species pairs even when indirect interactions or abiotic environmental effects largely
18 overpowered them. A linear approximation, based on partial covariances, also performed well
19 as long as the number of sampled locations exceeded the number of species in the data.
20 Indirect effects reliably caused a common null modeling approach to produce incorrect
21 inferences, however.

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

24 ***Introduction***

25 Ecologists' intense interest in drawing inferences about species interactions —especially
26 competition—from presence-absence data has a long history (MacArthur 1958, Diamond
27 1975, Connor et al. 2013). If nontrophic species interactions are important drivers of
28 community assembly patterns, then we should expect to see their influence in our data sets.
29 Despite decades of work and several major controversies, however (Lewin 1983, Strong et al.
30 1984, Gotelli and Entsminger 2003), existing methods for detecting competition's effects on
31 community structure are unreliable (Gotelli and Ulrich 2009). More generally, it can be
32 difficult to reason about the complex web of direct and indirect interactions that contribute
33 to the structure of real assemblages, especially when these interactions occur against a
34 background of other ecological processes such as dispersal and environmental filtering
35 (Connor et al. 2013). For this reason, it isn't always clear what kinds of patterns would even
36 constitute evidence of competition, as opposed to some other biological process or random
37 sampling error (Lewin 1983, Roughgarden 1983).

38 Most existing methods in this field compare the frequency with which two putative
39 competitors are observed to co-occur against the frequency that would be expected if *all*
40 species on the landscape were independent (Strong et al. 1984, Gotelli and Ulrich 2009).
41 Examining a species pair against such a “null” background, however, rules out the possibility
42 that the overall association between two species could be driven by an outside force, as
43 shown in Figure 1. Even though the two shrub species in this example compete with one
44 another for resources at a mechanistic level, they end up clustering together on the
45 landscape because they both grow best in areas that are not overshadowed by trees. Due to
46 fundamental constraints on the structure of valid covariance matrices (Brown et al. 2010),

47 this kind of positive correlation is almost inevitable, even when all the interactions are
48 negative. For this reason, knowing that the two shrubs co-occur more than expected by
49 chance under some null model (e.g. that their correlation in the data set is positive) can
50 not—by itself—tell us much about their direct effects on one another.

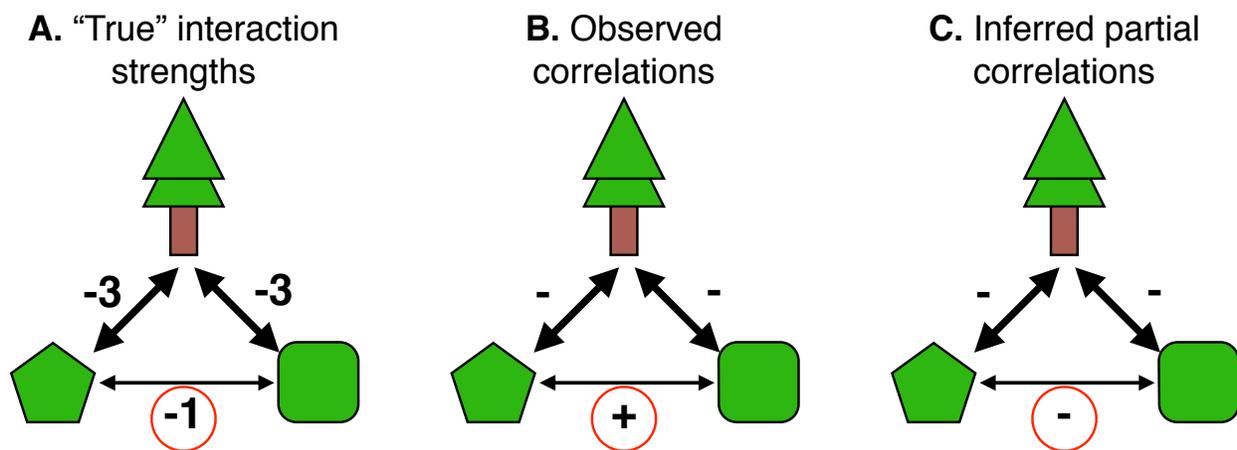


Figure 1: **A.** A small network of three competing species. The tree (top) tends not to co-occur with either of the two shrub species, as indicated by the strongly negative coefficient linking them. The two shrub species also compete with one another, as indicated by their negative coefficient (circled), but this effect is substantially weaker. **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 method such as a partial covariance or a Markov network allows us to correctly identify the negative interaction between these two species (circled).

51 While the competition between the two shrubs in the previous example does not leave the

52 commonly-expected pattern in community structure (negative association at the landscape
53 level), it nevertheless does leave a signal in the data (Figure 1C). Specifically, *among shaded*
54 *sites*, there will be a deficit of co-occurrences, and *among unshaded sites*, there will also be
55 such a deficit. These *conditional* (“all-else-equal”) relationships can be estimated with
56 partial correlations (Albrecht and Gotelli 2001), although this approach is best-suited for
57 Gaussian variables. For non-Gaussian data, a more general approach is needed.

58 In this paper, I introduce Markov networks (undirected graphical models also known as
59 Markov random fields; Murphy 2012) as a framework for understanding the landscape-level
60 consequences of pairwise species interactions, and for detecting them from observational data.
61 Markov networks, which generalize partial correlations to non-Gaussian data, have been used
62 in many scientific fields to model associations between various kinds of “particles”. For
63 example, a well-studied network called the Ising model has played an important role in our
64 understanding of physics (where nearby particles tend to align magnetically with one another;
65 Cibra 1987). In spatial contexts, these models have been used to describe interactions
66 between adjacent grid cells (Harris 1974, Gelfand et al. 2005). In neurobiology, they have
67 helped researchers determine which neurons are connected to one another by modeling the
68 structure in their firing patterns (Schneidman et al. 2006). Following recent work by Azaele
69 et al. (2010) and Fort (2013), I suggest that ecologists could similarly treat species as the
70 interacting particles in this modeling framework. Doing so would allow ecologists to simulate
71 and study the landscape-level consequences of arbitrary species interaction matrices, even
72 when our observations are not Gaussian. While ecologists explored some related approaches
73 in the 1980’s (Whittam and Siegel-Causey 1981), computational limitations had previously
74 forced researchers to rely on severe approximations that produced unintelligible results (e.g.
75 “probabilities” greater than one; Gilpin and Diamond 1982). Now that it is computationally

76 feasible to fit these models exactly, the approach has become worth a second look.

77 The rest of the paper proceeds as follows. First, I discuss how Markov networks work and
78 how they can be used to simulate landscape-level data and to predict the direct and indirect
79 consequences of possible interaction matrices. Then, using simulated data sets where the
80 “true” ecological structure is known, I compare this approach with several existing methods
81 for detecting species interactions. Finally, I discuss opportunities for extending the approach
82 presented here to larger problems in community ecology.

83 *Methods*

84 ***Conditional relationships and Markov networks.*** Ecologists are often interested in
85 inferring direct interactions between species, controlling for the indirect influence of other
86 species. In statistical terms, this implies that ecologists want to estimate *conditional*
87 (“all-else-equal”) relationships, rather than *marginal* (“overall”) relationships. The most
88 familiar conditional relationship is the partial correlation, which indicates the portion of the
89 correlation between two species that remains after controlling for other variables in the data
90 set (Albrecht and Gotelli 2001), as opposed to the simple correlation, which does not control
91 for other variables. The example with the shrubs and trees in Figure 1 shows how the two
92 correlation measures can have opposite signs, and suggests that the partial correlation is
93 more relevant for drawing inferences about species interactions (e.g. competition). To the
94 extent that our observations are not multivariate Gaussian (e.g. presence-absence data),
95 partial covariances do not describe variables’ conditional relationships exactly, and a more
96 general approach is needed. Markov networks allow researchers to generalize conditional
97 relationships to non-Gaussian networks, much as generalized linear models allow researchers
98 to extend regression to non-Gaussian response variables (Lee and Hastie 2012).

99 Markov networks give a probability value for every possible combination of presences and
100 absences in communities. For example, given a network with binary outcomes (i.e. 0 for
101 absence and 1 for presence), the relative probability of observing a given presence-absence
102 vector, \vec{y} , is given by

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

103 Here, α_i is the amount that the presence of species i contributes to the log-probability of \vec{y} ;
104 it directly controls the prevalence of species i . Similarly, β_{ij} is the amount that the
105 co-occurrence of species i and species j contributes to the log-probability, and controls how
106 often the two species will be found together (Figure 2A, Figure 2B). β thus acts as an analog
107 of the partial covariance, but for non-Gaussian networks. Because the relative probability of
108 a presence-absence vector increases when positively-associated species co-occur and decreases
109 when negatively-associated species co-occur, the model tends to produce assemblages that
110 have many pairs of positively-associated species and relatively few pairs of
111 negatively-associated species (exactly as an ecologist might expect).

112 A major feature of Markov networks is the fact that the conditional relationships between
113 species can be read directly off the matrix of β coefficients (Murphy 2012). For example, if
114 the coefficient linking two mutualist species is $+2$, then—all else equal—the odds of
115 observing either species increase by a factor of e^2 when its partner is present (Murphy 2012).
116 Of course, if all else is *not* equal (e.g. Figure 1, where the presence of one competitor is
117 associated with release from another competitor), then species' marginal association rates
118 can differ from this expectation. For this reason, it is important to consider how coefficients'
119 effects propagate through the network, as discussed below.

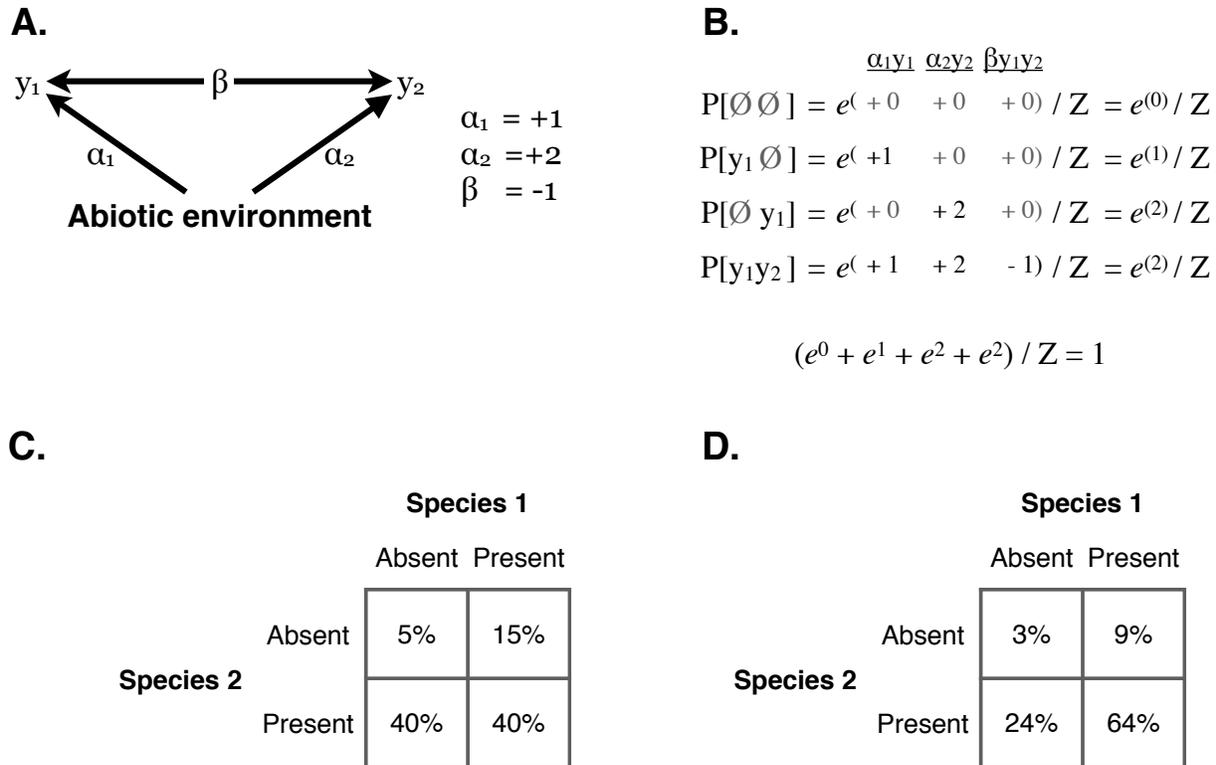


Figure 2: **A.** A small Markov network with two species. The depicted abiotic environment favors the occurrence of both species ($\alpha > 0$), particularly species 2 ($\alpha_2 > \alpha_1$). The negative β coefficient linking these two species implies that they co-occur less than expected under independence. **B.** Relative probabilities of all four possible presence-absence combinations for Species 1 and Species 2. The exponent includes α_1 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 relative probabilities sum to 1. In this case, Z is about 18.5. **C.** Using the probabilities, we can find the expected frequencies of all possible co-occurrence patterns between the two species of interest. **D.** If β 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 occurrence rate and the deficit of co-occurrences would be eliminated.

120 Estimating the marginal relationships predicted by a Markov network is more difficult than
121 estimating conditional relationships, because doing so requires absolute probability estimates.
122 Turning the relative probability given by Equation 1 into an absolute probability entails
123 scaling by a *partition function*, $Z(\alpha, \beta)$, which ensures that the probabilities of all possible
124 assemblages that could be produced by the model sum to one (bottom of Figure 2B). While
125 the analyses in this paper involve exact computations, this approach quickly becomes
126 infeasible as the number of species increases: with 2^N possible assemblages of N species the
127 number of bookkeeping operations required for exact inference quickly spirals exponentially
128 into the billions. Numerous techniques are available for working with Markov networks that
129 keep the computations tractable, either through analytic approximations (Lee and Hastie
130 2012) or Monte Carlo sampling (Salakhutdinov 2008), but these techniques are beyond the
131 scope of this paper.

132 ***Simulations.*** In order to compare different methods for drawing inferences from
133 observational data, I simulated two sets of landscapes using known parameters.
134 The first set of simulated landscapes included the three competing species shown in Figure 1.
135 For each of 1000 replicates, I generated a landscape with 100 sites by sampling exactly from
136 a probability distribution defined by the interaction coefficients in that figure. Each of the
137 methods described below (a Markov network, two correlation-based methods and a null
138 modeling approach) was then evaluated on its ability to correctly infer that all three species
139 interactions were negative, despite the frequent co-occurrences of the two shrub species.
140 I also simulated a second set of landscapes with five, ten, or twenty potentially-interacting
141 species on landscapes composed of 20, 100, 500, or 2500 observed communities (24 replicate
142 simulations for each combination). These simulated data sets span the range from small,

143 single-observer data sets to large collaborative efforts such as the North American Breeding
144 Bird Survey. As described in Appendix 1, I randomly drew the “true” coefficient values for
145 each replicate so that most species pairs interacted negligibly, a few pairs interacted very
146 strongly, and competition was three times more common than facilitation. I then used Gibbs
147 sampling to randomly generate landscapes with varying numbers of species and sites via
148 Markov chain Monte Carlo (Appendix 1). For half of the simulated landscapes, I treated
149 each species’ α coefficient as a constant, as described above. For the other half, I treated the
150 α coefficients as linear functions of two abiotic environmental factors that varied from
151 location to location across the landscape (Appendix 1). The latter set of simulated
152 landscapes provide an important test of the methods’ ability to distinguish co-occurrence
153 patterns that were generated from pairwise interactions among the observed species from
154 those that were generated by external forces like abiotic environmental filtering. This task
155 was made especially difficult because—as with most analyses of presence-absence data for
156 co-occurrence patterns—the inference procedure did not have access to any information
157 about the environmental or spatial variables that helped shape the landscape (cf Connor et
158 al. 2013, Blois et al. 2014).

159 ***Inferring α and β coefficients from presence-absence data.*** The previous sections
160 involved known values of α and β . In most real cases, however, the true values of these
161 parameters are not known, and ecologists will need to estimate them from data. When the
162 number of species is reasonably small, one can compute exact maximum likelihood estimates
163 for all of the α and β coefficients by optimizing Equation 1 (Appendix 2). Doing so allows us
164 to find the unique set of α and β coefficients that would be expected to produce exactly the
165 observed occurrence frequencies and co-occurrence frequencies. Fully-observed Markov
166 networks like the ones considered here have unimodal likelihood surfaces (Murphy 2012),

167 ensuring that this procedure will always converge on the global maximum.

168 For the analyses in this paper, I made a small modification to this maximum likelihood
169 procedure. Given the large number of parameters associated with some of the networks to be
170 estimated, I regularized the likelihood using a logistic prior distribution (Gelman et al. 2008)
171 with a scale of 1 on the α and β terms.

172 The optimization procedures described here and in Appendix 2 are implemented in the
173 *rosalia* package¹ for the R programming language (R Core Team 2015).

174 ***Other inference techniques for comparison.*** After fitting Markov networks to the
175 simulated landscapes described above, I used several other techniques for inferring the sign
176 and strength of marginal associations between pairs of species.

177 The first two interaction measures were the simple and partial covariances between each pair
178 of species' data vectors on the landscape (Albrecht and Gotelli 2001). Because partial
179 covariances are undefined for landscapes with perfectly-correlated species pairs, I used a
180 regularized estimate based on ridge regression [Wieringen and Peeters (2014); i.e. linear
181 regression with a Gaussian prior]. For these analyses, I set the ridge parameter to 0.2
182 divided by the number of sites on the landscape.

183 The third method, described in Gotelli and Ulrich (2009), involved simulating possible
184 landscapes from a null model that retained the row and column sums of the original matrix
185 (Strong et al. 1984). Using the default options in the Pairs software described in Gotelli and
186 Ulrich (2009), I simulated the null distribution of scaled C-scores (a test statistic describing
187 the number of *non-co-occurrences* between two species). The software then calculated a
188 *Z*-statistic for each species pair using this null distribution. After multiplying this statistic

¹source code available for download from <https://github.com/davharris/rosalia>

189 by -1 so that positive values corresponded to facilitation and negative values corresponded
190 to competition, I used it as another estimate of species interactions.

191 ***Method evaluation.*** For the first simulated landscape (three species), I kept the
192 evaluation simple and qualitative: any method that reliably determined that the two shrub
193 species were negatively associated passed; other methods failed.

194 For the larger landscapes, I rescaled the four methods' estimates using linear regression
195 through the origin so that they all had a consistent interpretation. In each case, I regressed
196 the "true" β coefficient for each species pair against the model's estimate, re-weighting the
197 pairs so that each landscape contributed equally to the rescaled estimate². Then, for each
198 combination of species richness, landscape size, method, and presence/absence of
199 environmental filtering, I calculated the mean squared error associated with the rescaled
200 estimate across all 12 replicates. Finally, I calculated an R^2 value for each landscape by
201 comparing this squared error estimate with the squared error associated with the null
202 hypothesis that all interactions were zero.

203 ***Results***

204 ***Three species.*** As shown in Figure 1, the marginal relationship between the two shrub
205 species was positive—despite their competition for space at a mechanistic level— due to
206 indirect effects of the dominant tree species. As a result, the covariance method falsely
207 reported positive associations 94% of the time and the randomization-based null model
208 falsely reported such associations 100% of the time. The two methods for evaluating
209 conditional relationships (Markov networks and partial covariances), however, successfully
210 controlled for the indirect pathway via the tree species and each correctly identified the

²The null model generated one Z -score outlier greater than 1000, which dominated the regression and squared error analyses. To reduce its influence on these results, I changed its value to 32.5, which was the value of the next largest Z -score in the null model's results.

211 direct negative interaction between the shrubs 92% of the time.

212 **Larger landscapes.** The Markov network consistently outperformed all of the other
 213 techniques that I evaluated (Figure 3). Of the remaining methods, partial covariances tended
 214 to be most accurate, especially when 100 or more communities had been observed on the
 215 landscape. Z -scores from the null model had the least predictive utility overall, as indicated
 216 by R^2 values. In general, all the models' accuracies declined when environmental filters were
 217 added, but the declines were smaller when environmentally-induced correlations were spread
 218 out among a larger number of pairwise interactions.

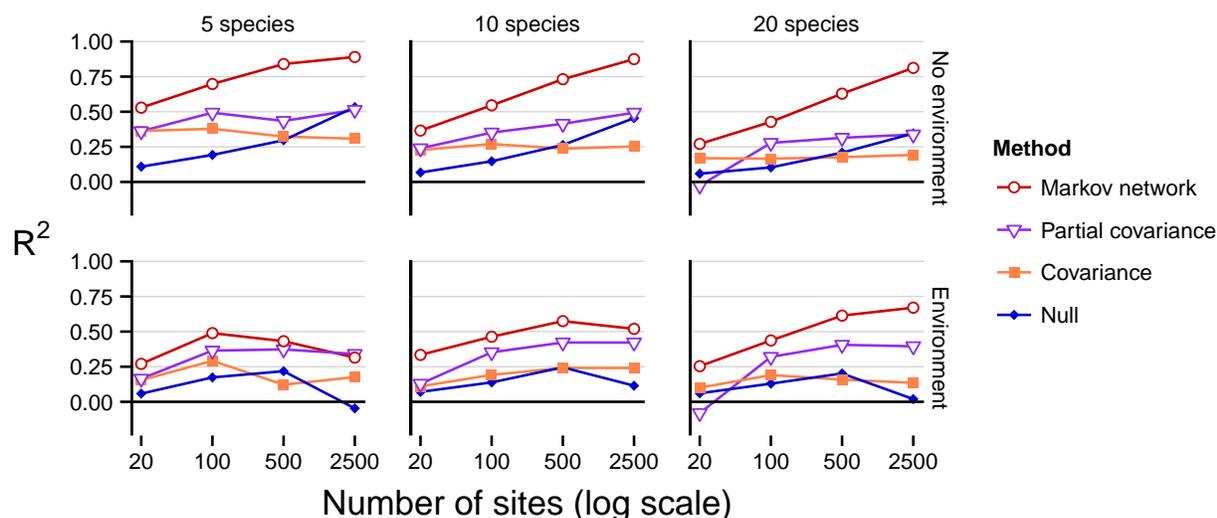


Figure 3: Proportion of variance in interaction coefficients explained by each method with 5, 10, or 20 species arrayed across varying numbers of sampled locations when environmental filtering was absent (top row) or present (bottom row).

219 **Discussion**

220 The results presented above are very promising, as they show that Markov networks can
 221 recover much of the variation in species' pairwise interaction strengths from observational
 222 data, even when direct interactions are largely overwhelmed by indirect effects (e.g. Figure 1)
 223 or environmental effects (lower panels of Figure 3). For cases where it is infeasible to fit a

224 Markov network, these results also indicate that partial covariances—which can be computed
225 straightforwardly by linear regression—can often provide an accurate approximation.

226 Apart from the environmental filters, the simulated landscapes presented here represent the
227 best-case scenario for these methods. Future research should thus examine these models’
228 performance characteristics when the “true” interaction matrices include guild structure or
229 trophic levels, which could make the β coefficients much more difficult to infer (particularly
230 for linear approximations like the partial covariance approach; Loh and Wainwright (2013)).

231 On the other hand, ecologists may often have prior information about the nature of real
232 species’ interaction patterns from natural history or ecological experiments, which could
233 substantially reduce the probability and magnitude of error. The *rosalia* package has built-in
234 mechanisms for incorporating this kind of information, if it can be expressed as a prior
235 probability distribution or a penalty on the likelihood.

236 Additionally, it is important to note that, while partial correlations and Markov networks
237 both prevent us from mistaking marginal associations for conditional ones, they cannot tell
238 us the underlying biological mechanism. Real species co-occurrence patterns will depend on
239 a number of factors—especially in taxa that emigrate in response to other species—and the
240 β coefficients in Markov networks have to reduce this to a single number. Thus, experiments
241 and natural history knowledge will generally be required to pin down the exact nature of the
242 interaction (e.g. who outcompetes whom).

243 Despite these limitations, the results with environmental filtering seem to indicate that the
244 method can be very robust. Additionally, the fact that Markov networks provide a likelihood
245 function to optimize makes them highly extensible, even when it is inconvenient to compute
246 the likelihood exactly. For example, the *mistnet* software package for joint species

247 distribution modeling (Harris 2015) can fit *approximate* Markov networks to large species
248 assemblages (>100 species) while simultaneously modeling each species' response to the
249 abiotic environment with complex, nonlinear functions. This sort of approach, which
250 combines multiple ecological processes, could help ecologists to disentangle different factors
251 behind the co-occurrence patterns we observe in nature. Numerous other extensions are
252 possible: similar networks can be fit with continuous variables, count data, or both (Lee and
253 Hastie 2012). There are even methods (Whittam and Siegel-Causey 1981, Tjelmeland and
254 Besag 1998) that would allow the coefficient linking two species in an interaction matrix to
255 vary as a function of the abiotic environment or of third-party species that could tip the
256 balance between facilitation and exploitation (Bruno et al. 2003). Fully exploring these
257 possibilities will require more research into the various available approximations to the
258 log-likelihood and to its gradient, in order to balance efficiency, accuracy, and the ability to
259 generate confidence limits for statistical inference.

260 By providing precise quantitative expectations about the results of species interactions,
261 Markov networks have the potential for addressing long-standing ecological questions. For
262 example, Markov networks can provide a precise answer to the question of how competition
263 affects species' overall prevalence, which was a major flash point for the null model debates
264 in the 1980's (Strong et al. 1984). From Equation 1, one can derive the expected prevalence
265 of a species in the absence of biotic influences ($\frac{1}{1+e^{-\alpha}}$). Any significant difference between
266 this value and the observed prevalence can be attributed to the β coefficients linking this
267 species to its facilitators and competitors (cf Figure 2D).

268 This paper only scratches the surface of what Markov networks can do for ecology. This
269 family of models—particularly the Ising model for binary networks—has been extremely

270 well-studied in statistical physics for nearly a century, and the models' properties,
271 capabilities, and limits are well-understood in a huge range of applications, from spatial
272 lattices (Gelfand et al. 2005) to haphazard neural connections in the retina and brain
273 (Schneidman et al. 2006) to complex behavioral interactions among human beings (Lee et al.
274 2013). Modeling species interactions using the same framework would thus allow ecologists
275 to tap into an enormous set of existing discoveries and techniques for dealing with indirect
276 effects, stability, and alternative stable states.

277 Finally, the null models commonly used by ecologists for drawing inferences about pairwise
278 species interactions do not seem to distinguish between direct and indirect interactions any
279 better than simple correlation coefficients. Null and neutral models can be very useful for
280 clarifying our thinking about the numerical consequences of species' richness and abundance
281 patterns (Harris et al. 2011, Xiao et al. 2015), but it is less clear whether deviations from
282 the null hypothesis can clearly implicate specific biological processes like competition.

283 **Acknowledgements:** This research was funded by a Graduate Research Fellowship from
284 the US National Science Foundation and benefited greatly from discussions with A. Sih, M.
285 L. Baskett, R. McElreath, R. J. Hijmans, A. C. Perry, and C. S. Tysor. Additionally, A. K.
286 Barner, E. Baldrige, E. P. White, D. Li, D. L. Miller, N. Golding, and N. J. Gotelli
287 provided useful feedback on an earlier draft of this work.

288 **References:**

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

291 Azaele, S., R. Muneeppeerakul, A. Rinaldo, and I. Rodriguez-Iturbe. 2010. Inferring plant
292 ecosystem organization from species occurrences. *Journal of theoretical biology* 262:323–329.

293 Blois, J. L., N. J. Gotelli, A. K. Behrensmeyer, J. T. Faith, S. K. Lyons, J. W. Williams, K.
294 L. Amatangelo, A. Bercovici, A. Du, J. T. Eronen, and others. 2014. A framework for
295 evaluating the influence of climate, dispersal limitation, and biotic interactions using fossil
296 pollen associations across the late Quaternary. *Ecography* 37:1095–1108.

297 Brown, I. H., E. Bedrick, S. Morgan, I.-L. E. C. Ernest, and I. F. Kelly. 2010. Constraints
298 on Negative Relationships: Mathematical Causes and Ecological. *The Nature of Scientific
299 Evidence: Statistical, Philosophical, and Empirical Considerations*:298.

300 Bruno, J. F., J. J. Stachowicz, and M. D. Bertness. 2003. Inclusion of facilitation into
301 ecological theory. *Trends in Ecology & Evolution* 18:119–125.

302 Cipra, B. A. 1987. An introduction to the Ising model. *American Mathematical Monthly*
303 94:937–959.

304 Connor, E. F., M. D. Collins, and D. Simberloff. 2013. The checkered history of
305 checkerboard distributions. *Ecology* 94:2403–2414.

306 Diamond, J. M. 1975. The island dilemma: Lessons of modern biogeographic studies for the
307 design of natural reserves. *Biological conservation* 7:129–146.

308 Fort, H. 2013. Statistical Mechanics Ideas and Techniques Applied to Selected Problems in
309 Ecology. *Entropy* 15:5237–5276.

310 Gelfand, A. E., A. M. Schmidt, S. Wu, J. A. Silander, A. Latimer, and A. G. Rebelo. 2005.
311 Modelling species diversity through species level hierarchical modelling. *Journal of the Royal
312 Statistical Society: Series C (Applied Statistics)* 54:1–20.

313 Gelman, A., A. Jakulin, M. G. Pittau, and Y.-S. Su. 2008. A Weakly Informative Default
314 Prior Distribution for Logistic and Other Regression Models. *The Annals of Applied
315 Statistics* 2:1360–1383.

- 316 Gilpin, M. E., and J. M. Diamond. 1982. Factors contributing to non-randomness in species
317 Co-occurrences on Islands. *Oecologia* 52:75–84.
- 318 Gotelli, N. J., and G. L. Entsminger. 2003. Swap algorithms in null model analysis.
319 *Ecology*:532–535.
- 320 Gotelli, N. J., and W. Ulrich. 2009. The empirical Bayes approach as a tool to identify
321 non-random species associations. *Oecologia* 162:463–477.
- 322 Harris, D. J. 2015. Generating realistic assemblages with a Joint Species Distribution Model.
323 *Methods in Ecology and Evolution*.
- 324 Harris, D. J., K. G. Smith, and P. J. Hanly. 2011. Occupancy is nine-tenths of the law:
325 Occupancy rates determine the homogenizing and differentiating effects of exotic species.
326 *The American naturalist* 177:535.
- 327 Harris, T. E. 1974. Contact Interactions on a Lattice. *The Annals of Probability* 2:969–988.
- 328 Lee, E. D., C. P. Broedersz, and W. Bialek. 2013. Statistical mechanics of the US Supreme
329 Court. arXiv:1306.5004 [cond-mat, physics:physics, q-bio].
- 330 Lee, J. D., and T. J. Hastie. 2012, May. Learning Mixed Graphical Models.
- 331 Lewin, R. 1983. Santa Rosalia Was a Goat Ecologists have for two decades made
332 assumptions about the importance of competition in community organization; that idea is
333 now under vigorous attack. *Science* 221:636–639.
- 334 Loh, P.-L., and M. J. Wainwright. 2013. Structure estimation for discrete graphical models:
335 Generalized covariance matrices and their inverses. *The Annals of Statistics* 41:3022–3049.
- 336 MacArthur, R. H. 1958. Population ecology of some warblers of northeastern coniferous
337 forests. *Ecology* 39:599–619.

- 338 Murphy, K. P. 2012. *Machine Learning: A Probabilistic Perspective*. The MIT Press.
- 339 R Core Team. 2015. *R: A Language and Environment for Statistical Computing*. R
340 Foundation for Statistical Computing, Vienna, Austria.
- 341 Roughgarden, J. 1983. *Competition and Theory in Community Ecology*. The American
342 Naturalist 122:583–601.
- 343 Salakhutdinov, R. 2008. *Learning and evaluating Boltzmann machines*. Technical Report
344 UTML TR 2008-002, Department of Computer Science, University of Toronto, Dept. of
345 Computer Science, University of Toronto.
- 346 Schneidman, E., M. J. Berry, R. Segev, and W. Bialek. 2006. Weak pairwise correlations
347 imply strongly correlated network states in a neural population. *Nature* 440:1007–1012.
- 348 Strong, D. R., D. Simberloff, L. G. Abele, and A. B. Thistle. 1984. *Ecological communities:
349 Conceptual issues and the evidence*. Princeton University Press.
- 350 Tjelmeland, H., and J. Besag. 1998. *Markov Random Fields with Higher-order Interactions*.
351 *Scandinavian Journal of Statistics* 25:415–433.
- 352 Whittam, T. S., and D. Siegel-Causey. 1981. *Species Interactions and Community Structure
353 in Alaskan Seabird Colonies*. *Ecology* 62:1515–1524.
- 354 Wieringen, W. N. van, and C. F. Peeters. 2014. *Ridge Estimation of Inverse Covariance
355 Matrices from High-Dimensional Data*. arXiv preprint arXiv:1403.0904.
- 356 Xiao, X., D. J. McGlinn, and E. P. White. 2015. *A strong test of the Maximum Entropy
357 Theory of Ecology*. *The American Naturalist* 185:E70–E80.