Title: Ecological networks over the edge: hypergraph trait-mediated indirect interaction (TMII)
structure

Keywords: hypergraph; interaction modification; trait-mediated indirect interaction; nontrophic
interaction; food web; ecological network

Authors and Addresses:

Corresponding author: Golubski AJ (agolubsk@kennesaw.edu)

Antonio J. Golubski
Kennesaw State University
Department of Ecology, Evolution, and Organismal Biology
Science Building, Building #370, MD # 1202
370 Paulding Ave NW
Kennesaw, GA 30144

Erik E. Westlund
Kennesaw State University
Department of Mathematics
Math Building- Bldg. D, MD # 9085
1100 South Marietta Pkwy
Marietta, GA 30060

John Vandermeer
University of Michigan
Department of Ecology and Evolutionary Biology
2019 Kraus Natural Science Building
830 North University
Ann Arbor, MI 48109-1048

Mercedes Pascual
University of Chicago
Department of Ecology & Evolution
1101 E 57th Street
Chicago, IL 60637

Abstract:

Analyses of ecological network structure have yielded important insights into the functioning of complex ecological systems. However, such analyses almost universally omit non-pairwise interactions, many classes of which are crucial to system structure, function, and resilience. Hypergraphs are mathematical constructs capable of considering such interactions: we discuss their utility for studying ecological networks containing diverse interaction types, and associated challenges and strategies. We demonstrate the approach using a real-world coffee agroecosystem, where resistance to agricultural pests depends upon a large number of trait-mediated interactions. A hypergraph representation successfully reflects both the importance of species imposing such effects, and the context-dependency of that importance in terms of how it is affected by removal of other species from the system.

Ecological complexity and network topology

There has been, and still is, great interest in understanding what determines the stability and resilience of ecological systems. What allows one system to bounce back from a disturbance when another does not? Why do some systems contain higher diversity of coexisting species than others? What determines the efficient provisioning of ecosystem services, or the severity of a system's response to a given species loss or invasion? The tremendous complexity of ecological systems makes the search for generalizable answers to these questions, based on fundamental system properties, extremely challenging.

Ecologists have gained insight into such questions by conceptualizing complex ecological systems as networks (mathematically described as graphs; see Glossary). An important benefit

of this approach is the ability to quantify various aspects of network structure (or topology): patterns in how interactions between actors, typically species, are arranged. The ability to dissect a system's complexity by examining the effects of its structure separately from other elements is a valuable tool in the search for general and/or fundamental patterns and properties [1-3]. A detailed review of such results for ecological systems is beyond the scope of this paper, but to give a few examples: prior work has suggested that communities might be more sensitive to the loss of species with high degree [4], in- or out-degree [5-6], or eigenvector centrality [7]. Global network metrics including connectedness, clustering, degree distribution, average shortest-path distance between species [4, 8-10], network nestedness and/or compartmentalization [11-14], 'generalization' [15], or the relative frequencies of various 'motifs' [16-17], have been proposed to influence a system's sensitivity to perturbations, as measured by different stability measures including persistence. Grouping models based on interaction patterns [18-20] have been suggested as ways to identify species playing unique roles in a system, or discern organizational elements such as multiple 'channels' of energy flow. Space constraints preclude detailed description of the network quantities referenced in the preceding list, but our point is simply that such works demonstrate that quantitative descriptions of ecological network structure are meaningful. Importantly, analyses of network structure allow insight even in systems for which limited information is available: many of the results cited above for example are based on unweighted and/or undirected graphs (i.e. representations lacking information on interaction magnitudes and/or directions).

We do not mean to suggest that these approaches are superior to others: merely complementary. Indeed, several of the above examples evaluate network metrics' relevance via their power to predict the behaviors of dynamic models with specified functional forms.

Prevalence and importance of TMIIIs

Almost universally, studies of ecological network topology have considered only pairwise interactions. This approach is inadequate for considering many important ecological phenomena, which are fundamentally properties of more than two species. Trait-mediated indirect interactions (TMIIIs) or interaction modifications (IMs) associated with adaptive behaviors [21-22] provide key examples. Consider a forager which drops an unprofitable prey item from its diet when an alternate, more profitable, prey species is sufficiently abundant. In that scenario, the latter prey species affects the former via forager behavior: an inherently tripartite interaction. Similarly, when foragers reduce their activity (and thus their rate of prey consumption) in response to fear of a predator, that predator affects the foragers' prey via forager behavior. We will use the shorthand TMII to refer to non-pairwise ecological interactions in which a direct interaction between two species is modified by one or more others, even though some important examples do not involve clear changes in traits (such as effects of cover provided by vegetation).

The importance of TMIIIs is widely appreciated [23-25]. They have repeatedly been shown to be crucial for the structure and function of ecological systems [26-29], including provisioning of key system services [30-32], and recovery of systems from disturbance [33]. They have also been implicated as mediators of major unforeseen consequences of anthropogenic impacts on ecological systems. For example, fishing is suspected to have contributed to sea otter declines in the 1990s via effects on killer whale behavior [34], and TMIIIs have been implicated in multiple important changes in Yellowstone ecosystems following the reintroduction of wolves [35]. Understanding how TMIIIs influence large, complex ecological

systems in a more general and mechanistic way is necessary for predicting how such systems will respond to perturbations, and requires conceptual and analytical frameworks better able to consider the structure of interactive effects in large systems.

There is clear need to develop common theoretical frameworks capable of accommodating diverse interaction types including TMIs [36-40]. TMIs are generally highly structured, and there is reason to expect that interactions among them are central to their community-wide implications [39, 41]. However, the complexity introduced by such effects has hindered study of that structure. Some models explore effects of one [41-43] or more [44] specific adaptive behaviors, effectively containing IMs and structure in them, but only implicitly. Others introduce IMs randomly, without the structural constraints implied by many TMI-generating mechanisms [45-47]. Most models that consider TMIs in great detail (i.e. using differential equations) consider relatively few species. Even then, fully characterizing TMIs can be quite challenging, and signs of effects can change depending on functional forms and parameter values [48-49]. Quantitative analyses of ecological network structure which include TMIs would represent valuable tools for dissecting their complex effects and roles in large communities. However, conventional approaches must approximate TMIs with pairwise edges [20, 49], which fail to reflect their inherently non-pairwise nature.

Hypergraphs

Hypergraphs [50-51] are mathematical constructs capable of representing non-dyadic interactions such as TMIs: we argue here that they hold great promise for considering diverse classes of ecological interactions within a common theoretical framework. The key distinction between a graph and a hypergraph is that in the former, each edge is associated with precisely

two nodes, while the equivalent of an edge in the latter- a hyperedge- can connect any number of nodes, and thus represent an interaction involving any number of species. Hypergraphs are increasingly being used to study interactive effects in social [52-54] and biomolecular [55-56] networks. We have found only two applications of hypergraphs to community ecology [57, 53]. Each of these consider 'competition hypergraphs' which, while offering new insights, were derived from directed food web graphs, that is from information on only pairwise trophic interactions. We discuss and illustrate some of the potential hypergraphs hold for considering new, inherently non-pairwise, types of interactions which graphs alone cannot satisfactorily represent.

As a starting point in investigating the value of hypergraph representations of ecological communities, we construct a hypergraph based on a real-world coffee agroecosystem (Box 1, Figure I). This illustrative example facilitates discussion of some general factors and open areas to be considered when constructing and analyzing ecological hypergraphs containing TMIs. Box 2 and Figure II present some technical details of hypergraph representations, which help illustrate: 1) the straightforward way in which hyperedges can represent TMIs, 2) how hypergraph topology captures the context-dependency of TMIs, and 3) reasons that network metrics differ in the hypergraph case. This facilitates the biological interpretation of key elements of hypergraph structure.

One common way to extract insight from a network representation of a system is to calculate node properties- namely centrality metrics- as potential indicators of node importance. Hypergraph analogues of such centralities, which incorporate information regarding participation in non-pairwise interactions, have yielded new insights across diverse types of systems [52-57]. We ask how four such indicators of relative node importance in the hypergraph of the coffee

agroecosystem differ from those reached via two other representations: a graph containing only direct effects, and a 'clique multigraph' in which non-pairwise interactions are approximated by the insertion of additional pairwise edges (Box 2). We compare centralities of each species in each representation, both in the full system and when *Azteca* ants are deleted. Calculation of these centrality metrics in the hypergraph and clique multigraph cases generally involved applying traditional graph algorithms to the hypergraph and clique multigraph matrices (see Online Supplementary Material Code), although there are added considerations (see below).

Species centralities in our illustrative example

We focus on eigenvector and betweenness centrality results (Figure 1, Online Supplementary Material Tables S4 and S5). Degree and closeness centrality showed generally similar patterns, and are presented in Online Supplementary Material Figures S1-S3 and Tables S1-S3. It is intuitive that the graph representation we consider, which omits any representation of TMIs, might inadequately describe a system containing many such effects. We therefore focus on comparisons between the hypergraph and clique multigraph, as the latter does model the system's TMIs, albeit using pairwise interactions.

Hypergraph eigenvector centrality clearly flags the importance of *Azteca* ants and phorids, two species at the heart of many of the behavioral interactions previously identified as critical to this system's structure and function [30-32]. For the full system, hypergraph and clique multigraph eigenvector centralities are identical. However, the hypergraph representation suggests that phorids will be of low importance should *Azteca* ants be lost from the system, which makes biological sense given that most phorid interactions consist of modification of *Azteca* ant behavior. By contrast, clique multigraph eigenvector centrality suggests continued

phorid importance when *Azteca* ants are deleted. More pronounced differences were evident in the path-based metric of betweenness centrality, which in the clique multigraph inappropriately suggests that phorids become important only when *Azteca* ants are removed.

Although we focus on single nodes here, it is worth noting that hypergraph topology can also give a different image of global network properties, and that it is also possible to examine the relative centralities of each hyperedge [52]. The latter provides another promising route for hypergraph representations to contribute unique insights, with centrality metrics potentially serving as indicators of the relative roles and/or importance of specific individual TMIs.

Challenges and strategies

Hypergraph frameworks present unique challenges, which necessitate substantial biological interpretation at all stages of hypergraph construction and analysis. Multiple reasonable ways to encode a given interaction into a hyperedge will often exist, and TMIs associated with adaptive behaviors present novel challenges for hypergraph construction. Furthermore, graph metrics often have multiple hypergraph analogues. We discuss examples of each of these issues, and suggest some strategies for addressing them, to introduce some of the considerations that arise when constructing and interpreting ecological hypergraphs.

Encoding interactions into hyperedges

Drawing the appropriate hypergraph for a particular system of interest is a nontrivial task, which must be informed by knowledge of the system's biology. Consider an herbivore N which consumes plant P , and whose activity (and thus feeding rate) is reduced by each of three predators, X , Y and Z (we will ignore any potential modifications of the herbivore:predator

interactions). Encoding this information as a single hyperedge between all 5 species would be undesirable, because the loss of any one predator would remove that hyperedge, and with it any representation of the remaining two predators' effects on the $N:P$ interaction. This is similar to the approach used to construct competition hypergraphs in previous studies [53, 57], where the effects of node losses were not a main focus. We suggest that each hyperedge representing a TMII or IM should intuitively include both nodes involved in the basal direct interaction being modified. A second approach would be to draw all hyperedges that are combinatorially possible given this restriction - namely $\{X,N,P\}$, $\{Y,N,P\}$, $\{Z,N,P\}$, $\{X,Y,N,P\}$, $\{X,Z,N,P\}$, $\{Y,Z,N,P\}$, and $\{X,Y,Z,N,P\}$. This has the merit of reflecting the expected interactive nature of IMs [39]. The hyperedge $\{X, Y, N, P\}$ for example might reflect species X 's and Y 's reciprocal modifications of one another's effects on herbivore behavior. However, at least in the unweighted and undirected case, several of these hyperedges will generally serve to artificially inflate the centralities of nodes X , Y and Z : each predator's expected importance in the system will increase in the presence of additional similar predators, which is precisely the opposite of what is expected in reality [39]. Intermediate schemes which avoid the pitfalls of each extreme are possible: for example, in this case of combining qualitatively similar TMIIIs, one could insert only a single hyperedge for each unique combination of a direct interaction and a single species modifying that interaction. In the above example, these would be: $\{X,N,P\}$, $\{Y,N,P\}$, and $\{Z,N,P\}$. As discussed next, the ideal choice will likely change in weighted and/or directed hypergraphs.

Weakening modifications and weighted hypergraphs

The distinction between 'weakening' and 'strengthening' modifications [39] presents

unique challenges for hypergraph construction. In our agroecosystem web example, the phorid-mediated hyperedge f shown in Figure II offsets the *Azteca* ant-mediated hyperedge e , and it could thus be argued that phorids reduce *Azteca* ants' importance in the system. However, in the undirected and unweighted hypergraph representation we present, the presence of hyperedge f makes *Azteca* ants more, not less, important in the system based on the centrality metrics we examine. Generally, similar IMs affecting a common direct effect are expected to combine antagonistically [39]. This could suggest the insertion of a large number of hyperedges, corresponding at minimum to each species imposing one second-order modification reducing each other species' first-order modification of the basal direct link. In an unweighted and undirected hypergraph this would cause each new modifier to (potentially dramatically) increase the relative importance of all other modifiers, whereas in the real world the presence of many similar modifiers would reduce the impact of any one.

Expanding into weighted hypergraphs presents opportunities to accommodate the distinction between weakening and strengthening modifications, via appropriate interpretations of hyperedge weight. For example, with hyperedge weights which combine additively (i.e. the cumulative hyperedge weight experienced by beetles in Figure II would be the sum of the weights of hyperedges b , e , and f), weakening modifications could be assigned negative weights (which are not uncommon in graph contexts), and strengthening modifications positive weights. Such an approach would preserve the intuitive way in which changes to the web, such as species losses, could be considered within the hypergraph framework. Representing the weakening vs. strengthening aspect of hyperedges in this way will often require the distinction between multiple levels of modifier node within a given hyperedge. The effect represented by hyperedge f in Figure II, for example, weakens *Azteca* ants' modification of the direct scale:beetle

interaction, but in doing so strengthens that underlying direct interaction. If species' importance in the system is related to the total effective magnitude of modifications they impose and direct interactions they participate in, it might thus be appropriate for hyperedge f to be assigned weights that are positive for phorid (the ultimate modifier), negative for *Azteca* ants (the next-level-down modifier whose effective impact is reduced), and positive again for scale and beetle (the participants in the basal direct interaction whose effective magnitude is increased).

Online Supplementary Material Figure 4 presents a preliminary example of this approach, although further development and refinement is required, particularly for cases of pervasive reciprocal weakening modifications as would occur for the antagonistically combining IMs discussed earlier. The potential for weakening modifications to reduce rather than inflate species' centralities in a weighted hypergraph suggests that different topologies will be appropriate for representing a given system as an unweighted vs. weighted hypergraph. In the latter, the most faithful representation will likely generally contain a large number of hyperedges representing high-order, often weakening, modifications, and developing satisfactory schemes for assigning and combining hyperedge weights in large systems will be a nontrivial task. Interest among ecologists in interaction strengths [58] and the demonstrated relevance of properties of their distribution such as variance and correlation structure [69], further motivates the development of weighted hypergraph models of ecological networks, given that TMIs are one major determinant of these strengths.

Multiple hypergraph analogues of graph metrics

Hypergraphs are extremely general mathematical structures, and even simple, fundamental graph objects can have multiple different analogues in a hypergraph. Perhaps the

simplest example is that of degree. In a graph, a node's degree (the number of edges that are incident upon it) coincides exactly with the number of other nodes that it is adjacent to; a focal species which participates in n interactions interacts with n other species. In a hypergraph, these two numbers can differ greatly. For example, in Figure II, every species is adjacent to all three others, while the numbers of hyperedges in which phorids, *Azteca*, scales, and beetles appear are 2, 4, 4, and 3, respectively. Thus, hypergraph representations allow description of separate 'incidence degree' and 'adjacency degree' metrics for each node. Directed hypergraphs containing TMIs also present the possibility of sub-dividing a node's out-degree into multiple categories, based upon multiple types of effects that a species can impose (e.g., as a prey item, vs. directly modifying a direct interaction, vs. imposing some higher-order modification of a modification).

The fundamental graph structure of a path provides another example of a concept which must be re-interpreted in the hypergraph context, with alternate interpretations of a 'hyperpath' having different biological implications (see Box 3).

Conclusions

Currently, few studies of ecological network structure consider TMIs. We suggest that doing so via pairwise interactions alone fails to faithfully represent important aspects of TMII structure, and we argue that hypergraph approaches, which avoid this shortcoming, have the potential to yield valuable new insights, complementary to those provided by either dynamic models or graph representations.

Like any model, hypergraph representations of ecological communities constitute an attempt to capture essential features of the system while distilling complexity down to

manageable levels. As with any abstraction, there will be unique ways in which hypergraph representations omit or distort important information, leading to unique challenges associated with applying them appropriately and effectively. However, such challenges are offset by equally unique opportunities, which warrant further development of hypergraph approaches and more extensive comparisons between alternate methods of studying TMII structure. Theoretical advances in this area might also encourage collection and synthesis of more comprehensive empirical data on TMII structure within and across ecological communities.

Acknowledgements

This work was supported by a Kennesaw State University IDROP award to A Golubski, E Westlund, & Selena He, and subaward 3002942167 to A Golubski from National Science Foundation Theory in Biology Program Grant F020952 to M Pascual and Stefano Allesina.

References

- 1 Strogatz, S.H. (2001) Exploring complex networks. *Nature* 410, 268-276.
- 2 Pascual, M. and Dunne, J. A., eds (2006) *Ecological Networks: Linking Structure to Dynamics in Food Webs*, Oxford University Press.
- 3 Proulx, S.R. et al. (2006) Network thinking in ecology and evolution. *Trends Ecol. Evol.* 20, 345-353.
- 4 Solé, R.V. and Montoya, J.M. (2001) Complexity and fragility in ecological networks. *Proc. R. Soc. B* 268, 2039-2045.
- 5 Curtsdotter, A. et al. (2011) Robustness to secondary extinctions: comparing trait-based sequential deletions in static and dynamic food webs. *Basic Appl. Ecol.* 12, 571-580.

- 312 6 Berg, S. et al. (2015) Ecological communities are vulnerable to realistic extinction sequences.
313 Oikos 124, 486-496.
- 314 7 Allesina, S. and Pascual, M. (2009) Googling food webs: can an eigenvector measure
315 species' importance for coextinctions? PLoS Comput. Biol. 5(9):e1000494.
- 316 8 Montoya, J.M. and Solé, R.V. (2002) Small world patterns in food webs. J. Theor. Biol. 214,
317 405-412.
- 318 9 Williams, R.J. et al. (2002) Two degrees of separation in complex food webs. Proc. Natl.
319 Acad. Sci. USA 99, 12913-12916.
- 320 10 Dunne, J.A. et al. (2004) Network structure and robustness of marine food webs. Mar. Ecol.-
321 Prog. Ser. 273, 291-302.
- 322 11 Bascompte, J. et al. (2003) The nested assembly of plant-animal mutualistic networks. Proc.
323 Natl. Acad. Sci. USA 100, 9383-9387.
- 324 12 Krause, A.E. et al. (2003) Compartments revealed in food web structure. Nature 426, 282-
325 285.
- 326 13 Thebault, E. and Fontaine, C. (2010) Stability of ecological communities and the architecture
327 of mutualistic and trophic networks. Science 329, 853-856.
- 328 14 Stouffer, D.B. and Bascompte, J. (2011) Compartmentalization increases food web
329 persistence. Proc. Natl. Acad. Sci. USA 108, 3648-3652.
- 330 15 Pocock, M.J.O. et al. (2012) The robustness and restoration of a network of ecological
331 networks. Science 335, 973-977.
- 332 16 Bascompte, J. and Stouffer, D.B. (2009) The assembly and disassembly of ecological
333 networks. Phil. Trans. R. Soc. B 364, 1781-1787.
- 334 17 Stouffer, D.B. and Bascompte, J. (2010) Understanding food-web persistence from local to

335 global scales. *Ecol. Lett.* 13, 154-161.

336 18 Allesina, S. and Pascual, M. (2009) Food web models: a plea for groups. *Ecol. Lett.* 12, 652-
337 662.

338 19 Baskerville, E.B. et al. (2011) Spatial guilds in the Serengeti food web revealed by a
339 Bayesian group model. *PLoS Comput. Biol.* 7(12):e1002321.

340 20 Sander, E.L. et al. (2015) What can interaction webs tell us about species roles? *PLoS*
341 *Comput. Biol.* 11(7):e1004330.

342 21 Wootton, J.T. (1993) Indirect effects and habitat use in an intertidal community: interaction
343 chains and interaction modifications. *Am. Nat.* 141, 71-89.

344 22 Abrams, P.A. (1995) Implications of dynamically variable traits for identifying, classifying,
345 and measuring direct and indirect effects in ecological communities. *Am. Nat.* 146, 112-134.

346 23 Schmitz, O.J. et al. (2003) Linking individual-scale trait plasticity to community dynamics.
347 *Ecology* 84, 1081-1082.

348 24 Creel, S. and Christianson, D. (2008) Relationships between direct predation and risk effects.
349 *Trends Ecol. Evol.* 23, 194-201.

350 25 Beckerman, A. et al. (2010) Adaptive foragers and community ecology: linking individuals
351 to communities and ecosystems. *Funct. Ecol.* 24, 1-6.

352 26 McNaughton, S.J. (1978) Serengeti ungulates: feeding selectivity influences the effectiveness
353 of plant defense guilds. *Science* 199, 806-807.

354 27 Kotler, B.P. et al. (1992) Predator facilitation: the combined effects of snakes and owls on
355 the foraging behavior of gerbils. *Ann. Zool. Fenn.* 29, 199-206.

356 28 Heithaus, M.R. et al. (2008) Predicting ecological consequences of marine top predator
357 declines. *Trends Ecol. Evol.* 23, 202-210.

358 29 Peacor, S.D. et al. (2011) Scaling-up anti-predator phenotypic responses of prey: impacts
359 over multiple generations in a complex aquatic community. *Proc. R. Soc. B* 279, 122-128.

360 30 Vandermeer, J. et al. (2010) Ecological complexity and pest control in organic coffee
361 production: uncovering an autonomous ecosystem service. *Bioscience* 60, 527-537.

362 31 Perfecto, I. et al. (2014) Complex ecological interactions in the coffee agroecosystem. *Annu.*
363 *Rev. Ecol. Evol. Syst.* 45, 137-158.

364 32 Vandermeer, J. et al. (2014) Qualitative dynamics of the coffee rust epidemic: educating the
365 intuition with theoretical ecology. *BioScience* 64, 210-218.

366 33 Dixon, D.L. et al. (2014) Chemically mediated behavior of recruiting corals and fishes: a
367 tipping point that may limit reef recovery. *Science* 345, 892-897.

368 34 Estes, J.A. et al. (1998) Killer whale predation on sea otters linking oceanic and nearshore
369 ecosystems. *Science* 282, 473-476.

370 35 Dobson, A.P. (2014) Yellowstone wolves and the forces that structure natural systems. *PLoS*
371 *Biol.* 12(12):e1002025.

372 36 Ings, T.C. et al. (2009) Ecological networks-beyond food webs. *J. Anim. Ecol.* 78, 253-269.

373 37 Olf, H. et al. (2009) Parallel ecological networks in ecosystems. *Phil. Trans. R. Soc. B* 364,
374 1755-1779.

375 38 Fontaine, C. et al. (2011) The ecological and evolutionary implications of merging different
376 types of networks. *Ecol. Lett.* 14, 1170-1181.

377 39 Golubski, A.J. and Abrams, P.A. (2011) Modifying modifiers: what happens when
378 interspecific interactions interact? *J. Anim. Ecol.* 80, 1097-1108.

379 40 Kéfi, S. et al. (2012) More than a meal... integrating non-feeding interactions into food webs.
380 *Ecol. Lett.* 15, 291-300.

381 41 Kamran-Disfani, A.R. and Golubski, A.J. (2013) Lateral cascade of indirect effects in food
382 webs with different types of adaptive behavior. *J Theor. Biol.* 339, 58-69.

383 42 Kondoh, M. (2007) Anti-predator defence and the complexity-stability relationship of food
384 webs. *Proc. R. Soc. B* 274, 1617-1624.

385 43 Kondoh, M. and Ninomiya, K. (2009) Food-chain length and adaptive foraging. *Proc. R. Soc.*
386 *B* 276, 3113-3121.

387 44 Uchida, S. and Drossel, B. (2007) Relation between complexity and stability in food webs
388 with adaptive behavior. *J. Theor. Biol.* 247, 713-722.

389 45 Arditi, R. et al. (2005) Rheagologies: modelling non-trophic effects in food webs. *Ecol.*
390 *Complex.* 2, 249-258.

391 46 Goudard, A. and Loreau, M. (2008) Nontrophic interactions, biodiversity, and ecosystem
392 functioning: an interaction web model. *Am. Nat.* 171, 91-106.

393 47 Lin, Y. and Sutherland, W.J. (2014) Interaction modification effects on ecological networks
394 are affected by ratio dependence and network topology. *J. Theor. Biol.* 363, 151-157.

395 48 Abrams, P.A. (1992) Predators that benefit prey and prey that harm predators: unusual effects
396 of interacting foraging adaptation. *Am. Nat.* 140, 573-600.

397 49 Dambacher, J.M. and Ramos-Jiliberto, R. (2007) Understanding and predicting effects of
398 modified interactions through a qualitative analysis of community structure. *Q. Rev. Biol.* 82,
399 227-250.

400 50 Berge, C. (1973) *Graphs and Hypergraphs*, North-Holland Publishing Co.

401 51 Bretto, A. (2013) *Hypergraph theory: an introduction*, Springer International Publishing.

402 52 Bonacich, P. et al. (2004) Hyper-edges and multidimensional centrality. *Soc. Networks* 26,
403 189-203.

- 53 Estrada, E. and Rodríguez-Velázquez, J.A. (2006) Subgraph centrality and clustering in
complex hyper-networks. *Physica A* 364, 581-594.
- 54 Cattuto, C. et al. (2007) Network properties of folksonomies. *AI Commun.* 20, 245-262.
- 55 Klamt, S. et al. (2009) Hypergraphs and cellular networks. *PLoS Comput. Biol.*
5(5):e1000385.
- 56 Ritz, A. et al. (2014) Signaling hypergraphs. *Trends Biotechnol.* 32, 356-362
- 57 Sonntag, M. and Teichert, H.M. (2004) Competition hypergraphs. *Discrete Appl. Math.* 14,
324-329.
- 58 McCann, K.S. (2012) *Food webs*, Princeton University Press.
- 59 Allesina, S. et al. (2015) Predicting the stability of large structured food webs. *Nat. Commun.*
6, 7842.
- 60 Gallo, G. et al. (1993) Directed hypergraphs and applications. *Discrete Appl. Math.* 42, 177-
201.

Figure captions

Figure 1. Two measures of the centrality of each node in the interaction network shown in Figure I, based on either a graph containing only direct interactions, a hypergraph incorporating IMs as non-pairwise interactions, or a 'clique multigraph' which uses pairwise interactions to approximate IMs (see main text and Box 2). Eigenvector centrality (**a**) increases with the number of interactions in which the focal species participates, as well as the scores of the species with which it interacts. Betweenness centrality (**b**) reflects the frequency with which a focal node lies on shortest paths between pairs of other nodes. Larger and redder nodes indicate higher centrality (relative to other nodes within the same web diagram, not across diagrams; data and details regarding calculation of each metric are presented in Online Supplementary Material Tables S4 and S5). For visual clarity, only edges corresponding to direct interactions in the original graph are drawn. Dashed lines encircle and highlight two key players in the system: phorid flies (3:00 clock position) and *Azteca* ants (3:36 position). In the hypergraph, each centrality metric indicates relatively high importance of phorids when *Azteca* ants are present in the web, but not in a web from which *Azteca* ants have been deleted. Phorid centralities are consistently low in the graph representation. In the clique multigraph, eigenvector centrality inappropriately continues to suggest phorid importance when *Azteca* ants are deleted; betweenness centrality inappropriately suggests that phorids are relatively unimportant with *Azteca* ants present, but increase in importance when *Azteca* ants are removed.

(Figure captions which are also included within corresponding Text Boxes)

Figure I. Coffee agroecosystem web. Black, blue, and red arrows indicate direct effects,

modifications of those direct effects, and modifications of those modifications, respectively. Square nodes indicate agricultural pests. Positive and negative effects (or strengthening and weakening IMs) are respectively indicated by triangular and circular arrow heads: our analyses omit this information to consider the simpler case of an undirected network.

Figure II. Key interactions within the system depicted in Figure I represented as edges and hyperedges, with corresponding hypergraph adjacency and incidence matrices. Direct effects, represented by solid black lines in (a), include beetles (*Azya orbiger*) preying on scale insects (b), *Azteca* ants consuming energy from scales (c), and phorid flies parasitizing *Azteca* ants (d). Indirect effects consist of *Azteca* ants reducing the magnitude of the scale:beetle interaction by protecting beetles (e, dashed blue line in a) and phorid flies reducing the magnitude of *Azteca* ants' effect on the scale:beetle interaction by eliciting defensive behavior (f, dashed red line in a).

Figure III. Examples of potential shortest hyperpaths from *Linyphiidae* to either *Diomus* (a, b) or White Halo Fungus (c, d). Paths are taken from Figure I: dashed borders indicate origin and destination nodes, and each gray letter indicates a single hyperedge. Hyperedge shapes are indicative of the direct interactions and first and second order modifications from which each arises. Red and blue borders indicate nodes that would be considered to be 'on' these hyperpaths under the alternate definitions discussed above.

Glossary

Adjacency: Any nodes which share incidence to at least one common edge or hyperedge are said to be adjacent to one another, as are any edges or hyperedges which share incidence to at least one common node.

Adjacency matrix: A square matrix $A = [a_{ij}]$, whose rows and columns are indexed by nodes. In an unweighted and undirected graph or multigraph or hypergraph, $a_{ij} = k$ if node i and node j appear in exactly k edges or hyperedges together. In the graph case all entries will equal either zero or one, because any pair of nodes may share at most one edge.

Betweenness centrality: A potential indicator of node importance which is based on how often the node appears on the shortest path(s) between two others, reflecting its potential role in the rapid and/or efficient transmission of information and/or effects between pairs of other nodes.

Closeness centrality: A potential indicator of node importance which is maximized by a low average shortest-path distance (i.e. low 'degrees of separation') to other nodes.

Degree: A node's degree equals the number of edges that it is incident to (the number of interactions in which it participates). In an unweighted graph, this also equals that node's degree centrality.

Edge: An element in a network which typically represents some interaction between two nodes: in this case an ecological interaction between members of different populations.

Eigenvector centrality: A potential indicator of node importance which is maximized by participating in many interactions with other high-scoring nodes.

Graph: A mathematical description of a network.

Hyperedge: The equivalent of an edge in a hypergraph: a hyperedge may represent an interaction which involves any number of nodes.

Hypergraph: A generalization of a graph in which the equivalents of edges (termed hyperedges) may each contain (be incident to; 'link') any number of nodes.

Hyperpath: The analogue of a path in a hypergraph.

Incidence: A node and any edge or hyperedge involving that node (formally 'containing' it) are said to be incident to one another.

Incidence matrix: A binary matrix $E = [e_{ij}]$ whose rows and columns are indexed by nodes and edges or hyperedges, respectively, and having $e_{ij} = 1$ if and only if node i is incident to edge or hyperedge j .

Interaction modification (IM): A modification of an interaction by a species (the 'modifier') not involved in that interaction.

In-degree: In a directed graph, the number of edges directed into a node.

Node or vertex: An object or actor in a network: in this case a population in an ecological community.

Out-degree: In a directed graph, the number of edges directed out of a node.

Path: A path between two nodes v_1 and v_{k+1} may be defined as an sequence of k distinct adjacent edges within the sequence, $v_1, e_1, v_2, e_2, \dots, v_k, e_k, v_{k+1}$, such that edge e_i is the edge joining (incident to both) nodes v_i and v_{i+1} .

Topology: In this context, the pattern of edge or hyperedge and node arrangement defining a network's structure.

Trait-mediated indirect interaction or effect (TMII): An interaction between two species, or effect of one species on another, which is mediated by some trait (often behavior) of a third species.

Undirected graph: A graph which does not contain information regarding directions of effects.

Unweighted graph: A graph which does not contain information regarding magnitudes of effects.

Box 1. An example: TMIs and pest control in a coffee agroecosystem

The example interaction network upon which we will focus (Figure I) was assembled based on field studies spanning 15 years of a 300-hectare organic farm in southern Mexico, which has been in production for nearly 100 years [30-32]. A complex web of interacting trophic and behavioral interactions has been shown to be crucial for controlling agricultural pests (an economically important ecosystem service, which has in recent times received exaggerated attention due to the coffee rust disease currently devastating coffee production in the Americas). These protective effects are particularly attributable to a large number of IMs imposed by ants of the genus *Azteca*, several of which are further modified by ant-parasitizing phorid flies. See [30-32] and references therein for further descriptions of the system and interactions within it.

Box 2 and Figure II describe how the information within Figure I was used to construct a hypergraph representation of this system. The networks we generate are undirected and unweighted; thus information within Figure I concerning signs and directions of interactions is ignored. We do this in order to consider as simple a first case as possible for our illustrative example. However, in prior studies metrics calculated from undirected and unweighted network structure have yielded insight into complex systems [4, 14-15, 52-54]. We do not foresee major conceptual challenges in constructing analogous directed hypergraphs, although this might complicate analyses and interpretations. Challenges in constructing weighted hypergraphs are discussed in the main text.

We construct three representations of this system: a graph which omits TMIs, a hypergraph, and a 'clique multigraph' which represents TMIs but utilizing only pairwise interactions (see Box 2 for details). We include a graph representation without any TMIs partly because omitting TMIs is exactly what most studies of ecological network structure do, under the implicit assumption that TMIs exist but might not modify the network conclusions, or simply because they cannot be incorporated. However, we focus on comparisons between the hypergraph and clique multigraph representations, in order to evaluate specifically the advantages of the former's capacity to capture the non-pairwise nature of TMIs.

Figure I. Coffee agroecosystem web. Black, blue, and red arrows indicate direct effects, modifications of those direct effects, and modifications of those modifications, respectively. Square nodes indicate agricultural pests. Positive and negative effects (or strengthening and weakening IMs) are respectively indicated by triangular and circular arrow heads: our analyses omit this information to consider the simpler case of an undirected network.

Box 2. Representing a food web with TMILs as a hypergraph

Constructing a hypergraph representation of the system from the information contained within Figure I is straightforward. We draw one hyperedge for each unique IM, with a node set that contains the modifier plus all species involved in any interactions being modified. Each direct interaction is also included as a hyperedge containing two nodes. The presence of any hyperedge requires the presence of all adjacent nodes, allowing hyperedges to topologically encode the context-dependency of TMILs. For example, removing *Azteca* ants from the system would mean the loss of hyperedge *f* in Figure II, thus removing the link between phorids and either scales or beetles.

Two matrix representations of this hypergraph, from which multiple metrics of interest may be calculated, are depicted below. The hypergraph adjacency matrix indicates how many hyperedges a pair of nodes share, but does not contain information regarding how those hyperedges are structured (e.g., how many of the three hyperedges involving scale and beetle in Figure II also contain *Azteca*). Because of this, the hypergraph incidence matrix is required for many analyses.

The graph representation to which we compare our hypergraph results is obtained by simply omitting any modifications, and encoding only pairwise direct effects. To further investigate the hypergraph representation's value, we also construct and analyze a 'clique multigraph' representation, which represents TMILs by containing one edge between a pair of species for each interaction (pairwise or non-pairwise) in which both species participate, equivalent to one edge for each hyperedge they share. Each hyperedge is thus replaced by a complete sub-multigraph, i.e. a clique, consisting of one edge between each pair of nodes it contains. For example, hyperedge *f* in Figure II would be approximated by the edges: {phorid, *Azteca*}, {phorid, scale}, {phorid, beetle}, {*Azteca*, scale}, {*Azteca*, beetle}, and {scale, beetle}). The clique multigraph indicates each species' aggregate participation in non-pairwise interactions, but does not encode how those interactions depend on other species in the network. Adjacency matrices of the hypergraph and 'clique multigraph' representations are identical, but node deletions behave differently between the two: the clique multigraph will retain all edges except those that are incident with the deleted node (i.e. the {phorid, scale}, {phorid, beetle}, and {scale, beetle} edges introduced as part of approximating hyperedge *f* in Figure II would persist after the deletion of *Azteca*, even though the interaction represented by hyperedge *f* biologically requires the *Azteca* ants' presence).

Figure II. Key interactions within the system depicted in Figure I represented as edges and hyperedges, with corresponding hypergraph adjacency and incidence matrices. Direct effects, represented by solid black lines in (a), include beetles (*Azya orbigera*) preying on scale insects (b), *Azteca* ants consuming energy from scales (c), and phorid flies parasitizing *Azteca* ants (d). Indirect effects consist of *Azteca* ants reducing the magnitude of the scale:beetle interaction by protecting beetles (e, dashed blue line in a) and phorid flies reducing the magnitude of *Azteca* ants' effect on the scale:beetle interaction by eliciting defensive behavior (f, dashed red line in a).

Figure 1

Box 3. Hyperpaths in an ecological hypergraph

It is often of particular interest to determine the distance between a pair of nodes as defined by the shortest path that connects them. Betweenness centrality reflects the frequency with which a given node appears on shortest paths between pairs of other nodes, as an indicator of the degree to which that node might thus facilitate efficient traversal of the graph. In a graph, each pair of adjacent edges along a path is connected by a single node, which makes identifying the nodes that lie on any given shortest path unambiguous. This is not the case in a hypergraph, and several slightly different analogues of a path (known as a hyperpath) exist (Figure III), with different biological interpretations. One option is to only consider those nodes through which a hyperpath enters and exits each hyperedge (red borders in Figure III) to be on that hyperpath. Where adjacent hyperedges share multiple incident nodes, as hyperedges *g* and *y* share incident nodes Phorid and Azteca in Figure III c, this perspective could call for defining a hyperpath as a distinct sequence of nodes and hyperedges, resulting in a greater number of hyperpaths being distinguished (for example, Figure III c would illustrate two shortest hyperpaths, one passing through Phorid and the other passing through Azteca).

However, consider phorid flies and beetles in the small web illustrated in Figure II. If hyperedge *f* in Figure II is considered to represent the shortest hyperpath (of length one) between those two nodes, it is reasonable to argue that *Azteca* ants and scales should be considered part of that hyperpath, since the existence of that hyperedge requires their presence. This suggests that it might be most appropriate to quantify betweenness based on which species each shortest hyperpath 'involves' rather than which species it specifically 'passes through': considering any node that is incident with any hyperedge in a hyperpath (blue borders in Figure III) to be on that hyperpath. This latter concept of a hyperpath was defined in [60] in the context of directed hypergraphs, and we adopt this perspective for the betweenness centralities we calculate.

Figure III. Examples of potential shortest hyperpaths from *Linyphiidae* to either *Diomus* (a, b) or White Halo Fungus (c, d). Paths are taken from Figure I: dashed borders indicate origin and destination nodes, and each gray letter indicates a single hyperedge. Hyperedge shapes are indicative of the direct interactions and first and second order modifications from which each arises. Red and blue borders indicate nodes that would be considered to be 'on' these hyperpaths under the alternate definitions discussed above.