

Reply to “Comment on Geodesic Cycle Length Distributions in Delusional and Other Social Networks”

Alex Stivala*

Abstract

Martin (2020) describes a misinterpretation of exponential random graph (ERGM) parameters in my contribution (Stivala 2020), with the use of this parametric model obscuring, rather than illuminating, the data. He suggests that this is symptomatic of a trend in the social networks community towards a methodological monoculture focussed on the use of ERGMs. In this Reply I try to clarify how this situation arose in this specific case, and address some more general issues Martin raises, including the use of nodal covariates, what we can learn from ERGMs, and methodological monoculturalism in social network research.

Keywords: Geodesic cycle, Exponential random graph model, ERGM, dk-series random graphs, Social networks, Fictional networks, Dissociative identity disorder

*Università della Svizzera italiana, Lugano, Switzerland, alexander.stivala@usi.ch

Introduction

I am grateful to John Levi Martin for both his fascinating original paper (Martin, 2017) in JoSS which inspired my contribution (Stivala 2020), and for his thought-provoking (and entertaining) Comment (Martin 2020). As Martin notes in his Comment, he had the sense from the original version of my manuscript which he saw as a reviewer, that part of my motivation was to establish the superiority of the exponential random graph model (ERGM) approach over others, and specifically the *dk*-series (Orsini et al., 2015) technique he used in Martin (2017). Another (anonymous) reviewer¹ suggested a direct comparison of ERGM and *dk*-series models rather than only using ERGM in my work, which, in its original form, meant that the only results from the *dk*-series were those for two of the “Patricia” networks described by Martin (2017). So I did this additional work and revised the manuscript accordingly, finding that the choice of which of the two methods (*dk*-series or ERGM) makes little difference to the results. Hence, to some extent, the manuscript in its published form pre-emptively concedes the truth of some of Martin’s critique: specifically that there was, both in principle and in fact, no need to use ERGM for the purposes of examining whether geodesic cycle length distributions of these networks are “unexpected” or not, and that the *dk*-series method as used in Martin (2017) is entirely suitable.

There are, however, several more issues raised in Martin’s Comment. Here I will explain my original reasoning, some of which lead to the errors as described in the Comment. I will also address some of the other issues Martin raises as to nodal covariates, and what we can learn from ERGM, specifically from failures of convergence or goodness-of-fit, and conclude with a response to his ideas about methodological monoculturalism and the pursuit of invariants.

Orthodoxy and ritualism

Martin argues that the social networks community is moving towards a monoculture of doing everything with ERGM and describes a kind of “ritualism” or “goal displacement”, in which what might be more boringly and bureaucratically called “standard operating procedure” for achieving some end becomes the end in itself. I will try to address this point, but I can only speak for myself, and will now attempt to explain (although perhaps not defend) this particular instance of ERGM ritualism. My motivation in choosing initially to re-do everything with ERGM rather than replicating Martin’s (2017) results using *dk*-series and extending them to the distributions of geodesic cycle lengths and to other network data was not (consciously, at least) an attempt defend ERGM “orthodoxy” or even to demonstrate any supposed superiority of ERGM over *dk*-series. It was, more prosaically, a case of my being familiar with ERGM and being confident that I would be able to estimate ERGM models for these networks, being familiar with the relevant methods and software. In contrast, I was not previously familiar with *dk*-series, and so the difficulty (or otherwise) of using it in practice was unknown to me. And so (motivated partly by convenience) I sought to demonstrate that ERGM *could* be used for these data (not that it *must* or even *ought to* be used). As it turned out, on being motivated to

¹ To whom I also grateful. An impression of Martin’s initial thoughts on the manuscript was also relayed to me by the editor, further suggesting to me that using only the ERGM was not the best way to proceed and inspiring me to reproduce all the results with both families of models.

also use the *dk*-series method by reviewer comments on the manuscript, I found the RandNetGen program² (Mahadevan et al., 2006; Colomer-de-Simón and Boguñá, 2014; Colomer-de-Simón et al., 2013) for generating *dk*-series networks was readily available, and simple to compile and use, and so there was actually little difficulty involved.³

In addition, as some of the networks I wished to use⁴ had node attribute data which on the face of it would likely affect structure significantly (gender, classroom, etc.), I had assumed that a model (such as ERGM) in which such nodal covariates could be incorporated would be beneficial (or even required). However, this was mistaken, as Martin’s Comment makes clear. Since the purpose was only to generate random networks sharing certain structural features (degree distribution, clustering, etc.) of the observed networks, the values of the estimated parameters were not of interest in themselves, and so there was no need for a parameterized model such as ERGM. The structural effects of structurally relevant nodal attributes are observed in the structure itself (by definition), and so a method such as *dk*-series is completely appropriate.

ERGM parameter interpretation

I can only agree with Martin’s observations that my statement that an “[ERGM] parameter tells us about the process occurring ...” is not correct. However I do not believe that a statistically significant ERGM parameter can never give us any insight into processes: it indicates a structure that may represent an underlying social process (Lusher and Robins, 2013). Of course, as Martin points out, such structures may have been formed by some other, perhaps very different, process: the best we can say is that the data are consistent with the process we hypothesize. If we had appropriate longitudinal data, a stochastic actor-oriented model (SAOM) could be used to model process rather than just reproduce a distribution in which the observed network statistics are central.⁵

² <https://github.com/polcolomer/RandNetGen>

³ Far too many times I have wanted to reproduce results from some paper, or use its methods elsewhere, only to find that the required method has no readily available software, or the software that is supposedly available is impossible (or at least too difficult for me) to compile or install, and so I would have to go to some (sometimes very large and time-consuming) effort to write my own software for the purpose, and more often than not have abandoned the attempt as too much work for too little potential reward. The availability of PNet (now MPNet) and statnet and related R packages encourages the use of ERGM; I suspect very many (probably most) potential users of a statistical (or generally computational) method for social networks (or sociology more generally) research will not use it if it requires them to download and compile source code themselves, rather than installing a Windows binary or R package.

⁴ The selection of these networks was of course somewhat arbitrary, subject to some constraints. I wanted social network data, both empirical and fictional (and not networks from other domains such as internet connectivity, protein interactions, publication citations, etc.), that were publicly available and not too different in size from the “Patricia” networks. So I could have chosen networks that did not have relevant node attributes, but this seemed too restrictive.

⁵ For more on the relationship between SAOM and ERGM, in addition to Leifeld and Cranmer (2019) cited by Martin, see Block et al. (2018, 2019) and Block, Stadtfeld and Snijders (2019).

The “chains of affection” network (Bearman, 2004) with its spanning tree-like structure, large geodesic cycle, and the hypothesized proscription against four-cycles, is an excellent example here, as Martin suggests. In Bearman, Moody, and Stovel (2004) and Rolls (2015), only the final cross-sectional view of the network is considered. Martin points out that such structures can also arise from a distribution of the actors in a space of “likeness”, low average degree, and a strong tendency for relations based on proximity in the space. However without longitudinal data, we cannot say what sort of process formed the observed structures. Fortunately, such data do exist, in the form of time-stamped romantic tie creation and dissolution events, and are analyzed in Stadtfeld, Hollway, and Block (2017a) using the dynamic network actor model (DyNAM), an actor-based model of coordination ties.⁶ Stadtfeld et al. explicitly test the hypothesis that “Individuals will tend to avoid forming cycles of length four in the romantic partnership network...” (p. 21). However, the corresponding parameter in their DyNAM model is not significant, although it is positive rather than negative as hypothesized. They hypothesize another process that the apparent lack of four-cycles is created by vacancy chain mechanisms creating few paths of length three. So we are still without any definite conclusion as to the hypothesized four-cycle avoidance process, although as noted by Stadtfeld et al. there are many shortcomings in the data available to them, and more complete data might yield less limited interpretations.

I am conscious that, in the previous paragraph, it might look as though I have fallen in to what Martin calls the “exhausted linear model” way of thinking, that every problem can be solved by putting enough parameters in the model (although in this case a dynamic actor-based model), and we can only learn from statistically significant parameter estimates in a converged model. Perhaps another way forward would be, as Martin suggests in the context of ERGM, to examine goodness-of-fit of different models to see what could be learned by failures to fit particular structures well. Unfortunately such a goodness-of-fit procedure does not seem to be available for the particular DyNAM used in Stadtfeld et al. (2017a), although the possibility of extending another DyNAM (for directed event sequences) to include such a procedure is suggested by Stadtfeld and Block (2017).

Getting back to the specific case of the “Patricia” networks, it was, as Martin notes, unwarranted to move from the GWDEGREE parameter, which describes nothing but degree distribution, to any process interpretation.⁷ Indeed nothing was gained by attempting to interpret the ERGM parameters for the “Patricia” networks and it would perhaps have been

⁶ For the relationships between DyNAM, relational event models (REM) and SAOM, see Butts (2017), Snijders (2017), Stadtfeld, Hollway and Block (2017b), and for a survey of tie-oriented dynamic network models, Fritz, Lebacher, and Kauermann. (2020).

⁷ Possibly illustrating Martin’s point even further, the substantive part of the phrase “there is a significant tendency *against* preferential attachment [by] degree (the GWDEGREE parameter is positive and significant)” [emphasis added, *against* was intended rather than the *towards* Martin corrected it to] was not a typographical error, although the sentence as a whole is regrettably confused, and, evidently, confusing. The unfortunate use of “preferential attachment” here (and Table B2 describing the parameters) is derived from Hunter (2007) where it is stated that GWDEGREE “may be thought of as a sort of anti-preferential attachment model term” (p. 221). In fact, a negative GWDEGREE parameter indicates centralization (edges accruing among a small number of high-degree nodes); note that this is “opposite” to the interpretation of the alternating-*k*-star parameter as used in PNet (where a positive value indicates centralization). Of course these parameters also interact with GWESP (or alternating-*k*-triangles), further complicating matters. Interpretation of the GWDEGREE parameter is notoriously confusing (Levy, 2016).

better not to even attempt to do so (as I did not for the 1990 Patricia network, or any of the other networks in which the parameters are relegated to an Appendix).

Nodal covariates in ERGMs

Martin notes the potential circularity (or regress from a structural model to a model of individual attributes) of adding membership of the “Sphere of the blue flame” as a covariate (as I did in Model 3 of Table 3 and Models 3 and 4 of Table 4), given that the Sphere is a structural (and perhaps spatial) feature, which resembles a network community. He imagines the consequences if Patricia had labeled each of the components or easily separable communities and these labels were included as nodal covariates, reducing actual structure to individual attributes. Of course, I did not do this, using only the Sphere (and, for the 1993 network, the separate component) that Patricia herself had, in fact, labeled, but nevertheless it was a poor modeling choice to do so, as this argument shows (it makes little difference to the geodesic cycle lengths of the networks simulated from these models in any case).

Of nodal covariates, or “actor-relation effects” (Lusher and Robins, 2013) in ERGM more generally, used for modeling sex homophily, for example, Martin states that such covariates “did not necessarily have any privileged theoretical position and not all were happy with the idea that the models should be stretched to include parameters that broke the deductive link to the Hammersley–Clifford theorem”, and, if we were to maintain the connection to the underlying mathematics, we ought to relax the homogeneity assumption by having different structural parameters for different sets of nodes. Whether or not such covariates have a privileged theoretical position, I do not think it is true that their inclusion (necessarily) breaks the deductive link to the Hammersley–Clifford theorem: Robins, Elliott and Pattison (2001) derive network models of social selection⁸ in the framework of endogenous network processes, explicitly from a variant of the Hammersley–Clifford theorem, in which nodal covariates (binary and continuous) are described as having the function of loosening the homogeneity assumption.

This explicit construction of a more general model to include attributes based on partial conditional independence (partially dependent attribute models) is not necessarily how we think about nodal covariates in present-day ERGM usage, however. For example, Robins and Daraganova (2013) state that “we prefer to consider attribute effects as indicative of exogenous processes that operate alongside endogenous self-organizing mechanisms” (pp. 91-92). This is presumably the interpretation Martin has in mind when he writes that it “is not impossible that there is a second stochastic process of sex homophily that happens to be independent of the structural processes ... but it is hardly obvious that this is frequently the case. If it is not the case, then all the ERGM can do is give us precise estimates of the wrong model.” However, it is hardly obvious that it is never (or even rarely) the case either, and as Martin also notes, predictors in ERGMs cannot generally be termed “independent” anyway. And, since Martin has already introduced Tukey’s “precisely wrong” and “approximately right”, I feel justified in quoting another well-worn aphorism that “all models are wrong, but some are useful”⁹, and

⁸ Referred to as a generalization of p* models in that paper.

⁹ Usually attributed to George Box.

I think we can agree that ERGMs can certainly be useful, even (or perhaps especially) when they include nodal covariates. For example, Models 2 and 3 in Gondal and McLean (2013) which Martin cites as an example of how to best to learn from ERGMs.

Gondal and McLean (2013) show three models, with Model 1 including only structural parameters, and Models 2 and 3 incorporating nodal attribute parameters. As Martin notes, the parameters added in subsequent models were chosen carefully, based on substantive understanding and ERGM convergence and goodness-of-fit, leading to the final (rather complex) model. The general procedure of starting with a purely structural model, and adding attribute parameters, checking if they improve model fit, is quite common in ERGM modeling, and is one I followed for the two Patricia networks which nodal attributes, and most¹⁰ of the other networks. However as already discussed, including parameters based on structural features such as the Sphere (or the component labeled “(behind)”) was a poor choice, so I should have stopped at Model 2 in Tables 3 and 4.

Convergence, goodness-of-fit, and learning from failure with ERGMs

I agree with Martin’s proposition that, rather than treating ERGM as an “exhausted linear model”, presenting a final converged and well-fitting model (perhaps a very complex one with many parameters, both structural and attribute-related) and then interpreting the parameters as “the truth”, we might be better served by presenting a series of increasingly complex nested models, in which the addition of parameters is guided by theory (and/or substantive understanding of the data) and failures of convergence or fit (most likely the latter, as I will discuss below). Of course the actual development of the final model might well have proceeded this way (with more or less understanding or justification for the added parameters), but a published paper usually obscures this process by presenting only the final product (or a small subset of the models developed along the way).

I will now address the question of whether recent developments allowing problems of convergence to be overcome in more cases are “cause for celebration”, or not (as Martin suggests). First, Martin notes that I use the (rather boilerplate) phrase that “only models that show acceptable convergence and goodness-of-fit ... are included”. Indeed this is because not all models could be adequately estimated, and in the Appendix to this Reply I detail the failed models. I do not think we learn much from these, with the possible exception of the high school friendship network where we find that a parameter for classroom homophily is required for model convergence (as noted in the paper).

Here I think it is important to distinguish between two modes of “failure”: first, a failure of the model to converge, and second, poor goodness-of-fit. The first mode, failure of the estimation procedure to converge, can be further classed as either model mis-specification, or algorithmic failure. The failure of MCMC or Bayesian ERGM estimation to converge “*most likely* is a sign not of algorithmic deficiencies but of the inability of the specified model to represent the

¹⁰ The exceptions being the law firm friendship network, where there is only a single model, including all the attributes shown (based on prior modeling experience with this network), and the Grey’s Anatomy network where heterophily on sex was clearly a very important feature.

observed network” (Pattison and Snijders, 2013, p. 289; emphasis added). Perhaps the model cannot represent the observed network because the network was formed by some process inconsistent with our dependence assumptions, in which case indeed we cannot expect to fit any model (with parameters allowed by those assumptions) and infer anything from it. However, if our dependence assumptions are reasonable, this is where the use of the geometrically weighted effects, for example, is useful to overcome problems related to near-degeneracy. Sometimes, however, the problem actually *is* in the estimation algorithm, and it is possible that a different algorithm can find converged estimates for the same model, e.g. the *E. coli* transcriptional regulation network in Hummel, Hunter and Handcock (2012).

So, from a non-converged model, we may learn that either our model is mis-specified, or the algorithm we used for estimation is incapable of estimating that model (but we might not know which). I suggest that what we learn from a non-converged model is usually somewhat limited. For example, we are no longer surprised when models specified under the dyadic independence or Markov assumptions fail to converge, and so starting with something like what Martin terms the “out of the box” specification I used, which is made under the social circuit dependence assumption and includes (in the undirected case) GWESP (or alternating- k -triangles) is, as he suggests, the right thing to do for the first model of some arbitrary social network, without taking its specific features into account yet.

It is from the second mode of failure, poor goodness-of-fit, that more may be learned. This, of course, is part of Martin’s argument that we should treat ERGM model fitting as a means to an end, and not fall into the trap of thinking that finding the model with the “best” fit is an end in itself. However to get to the point of worrying about fit, we need a converged model. When a model fails to converge, we do not have valid parameter estimates, so cannot simulate networks and see which features are reproduced or not (as we do in a goodness-of-fit test). We might not even know whether the failure of the estimation procedure to converge is due to our model being mis-specified or is just some artifact of the estimation procedure itself.

For example Gondal and McLean (2013) find that Markov models do not converge, from which nothing much can be learned. They perhaps learn more from the failure to converge of some slightly more parsimonious versions of the final model (see their footnote 10). However, as Martin points out, they learn something substantive from a converged Bernoulli model (and note that estimation of a Bernoulli model should always converge, no matter how poor its goodness-of-fit): that including reciprocity leads to much better goodness-of-fit and so there is likely a positive tendency towards reciprocity in the data.

For such reasons I think that an increased chance of getting a converged ERGM model (whether due to improved algorithms, or geometrically weighted model terms, etc.) is, if not necessarily cause for celebration, at least an advance, in that it potentially allows more to be learned about the data.

Are we really at risk of a monoculture?

Martin states that part of his motivation in writing the first paper (Martin, 2017) was “to try to help us avoid the monoculture that I see developing with the use of ERGMs”. The potential dangers of such a monoculture he illustrates by the interpretive errors I made in the use of ERGM parameters, and the interpretation of these parameters confusing or obscuring, rather than illuminating, the nature of the “Patricia” networks, for example. I agree that such an

orthodoxy or monoculture would be a bad thing¹¹, but it does raise the question of whether such a situation is, in fact, developing (although Martin is perhaps not suggesting that it is, but just that he foresees that such a situation might start developing).

There is some ambiguity (at least in my mind) as to what the potential monoculture is precisely: is it the use of ERGMs (however (mis-)used or (mis-)interpreted) displacing other methods, or specifically ERGMs as “exhausted” linear models (“ERGM-as-ELM”)? The former is suggested by the first paragraph of Martin’s Comment and also his comment on the dangers of conceptual monoculture that “had we only the data, Pajek and statnet, and were we permitted only to employ ERGMs...”; however the latter is suggested by his statement in the concluding section of his Comment that we need to “re-awaken our interest in the ERGM-not-as-ELM.”

While I would not attempt to predict the future, I do not get the impression (from the annual INSNA Sunbelt conferences, and reading journals such as *Connections*, *Journal of Social Structure*, and *Social Networks*) that ERGMs are taking over from other methods in a way that suggests a monoculture is developing. Certainly they do not seem to me to be approaching the status of something as ubiquitous as GLM (and in particular ordinary least squares regression). So while it is difficult to test if the second interpretation of the monoculture Martin sees as developing is actually already starting to occur (as it would require something like a systematic review), I thought it might be instructive to try a small (and rather crude) bibliometric test of the first interpretation.

To this end, I used Web of Science¹² to count the number of ERGM papers (counted as papers that contain “ERGM” or “exponential random graph” in their Web of Science “topic”, which includes title, abstract, and keywords) published in *Social Networks* from 2009 to 2020.¹³ The results are shown in Figure 1, and it seems there is no significant increase in the fraction of papers published in *Social Networks* each year with ERGM in the topic, over the last 10 years. If an ERGM monoculture is developing, it is not apparent in the pages of *Social Networks*, at least¹⁴. But this does not mean necessarily that this is not happening at all, or that it would not happen, and I agree that such a development would be unfortunate.

¹¹ And as Martin notes, the opposite situation, in which every paper is both a substantive claim and a methodological innovation, is also problematic.

¹² Clarivate Analytics <http://www.webofknowledge.com>.

¹³ I chose *Social Networks* as an appropriate journal since it seems to me that, if an ERGM monoculture were to develop, it would likely first be apparent in *Social Networks*. It contains both methodological and substantive applications of ERGMs, and also according to Web of Science it is, as we might expect, the journal with the largest share of ERGM papers: nearly 13% as opposed to the next highest 3.6% (*PLOS ONE*). I chose 2009 as the starting year as it is a time after ERGM became common terminology and software such as statnet and PNet became available, and yet 10 years should be long enough to get at least some impression of any trend.

¹⁴ Although the fact that just over 13% of articles published in *Social Networks* in the last 10 years mention ERGM as a topic might be considered evidence that a lack of methodological diversity has *already* developed. However to test this properly a systematic review (or at least comparison with mentions of other methods which could have been used) would have to be done in order to determine whether or not this is close to the highest fraction of papers in which ERGM could reasonably have been used at all.

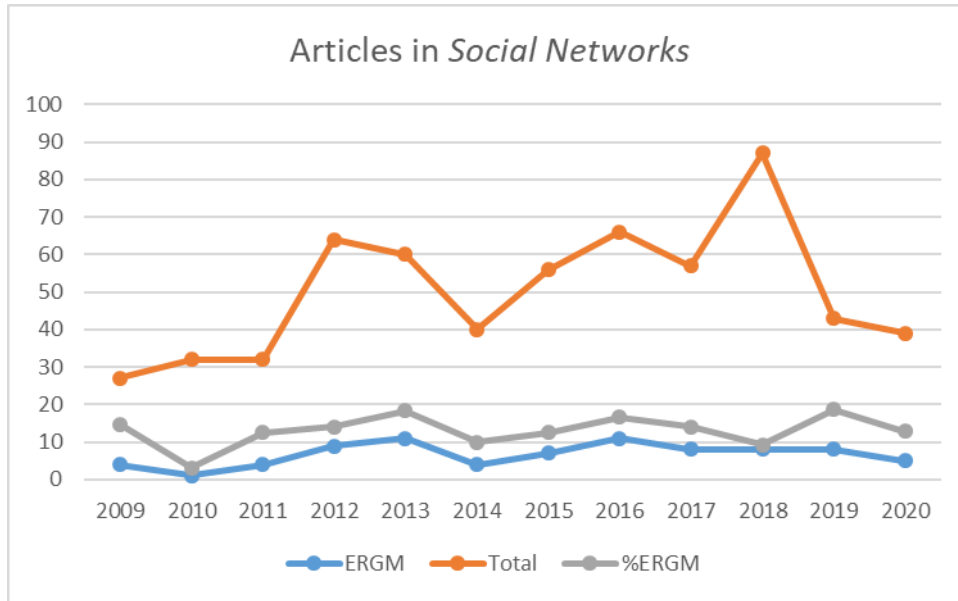


Figure 1. Total number of articles, articles with topic ERGM, and the percentage of such articles, published in Social Networks 2009–2020. Total is the query “PUBLICATION NAME: (SOCIAL NETWORKS)” for the specified year. ERGM adds “AND TOPIC: (“exponential random graph” OR *ergm*)”. The queries were run on 23 July 2020, so the 2020 data are incomplete.

In his Comment, Martin also states that it “is far too early to attempt to push for a monoculture”, with which I can only agree (and would suggest, as I am sure Martin also believes, that there would never be a time when pushing for such a methodological monoculture would be appropriate anyway, no matter how well developed a method). In a footnote to this, he mentions the short life of “best practices” in ERGM estimation methods (maximum pseudo-likelihood, varieties of MCMC methods, etc.). Certainly this fact indicates that it is too early to push for ERGM to become the sole acceptable method (even if this were for some reason desirable) for analyzing network data to which it is applicable. But ongoing methodological development is not a reason *not* to use ERGM, or indeed *dk*-series, where appropriate, either: as Abbott (1988) states “... one cannot require that alternative methods should not be considered until fully developed. The GLM did not emerge fully developed ... it became a full paradigm through a long process of development, criticism, and growth. To ask that alternatives achieve that development instantaneously is to deny the possibility of alternatives” (p. 184).

In closing, Martin suggests that we should be more enthusiastic about pursuing invariants or general laws (like the Ideal Gas Law) for social science, rather than tending to restrict ourselves to finding well-fitting models (“Actual Gas Fits”) and attempting to interpret statistically significant parameters. While I certainly cannot claim that such laws could never be found, an example from network science suggests that neglecting Actual Gas Fits can also potentially lead us astray. Consider the enthusiasm for “scale-free” or power law degree distributions where actual fit is perhaps neglected in the enthusiasm for simplicity and universality¹⁵ (Stumpf and Porter, 2012). However, it is possible that such networks are not even common, let alone universal (Broido and Clauset, 2019). Preferential attachment (Barabási and Albert, 1999)

¹⁵ Note that scale-free networks were suggested to be “more universal” than just social networks, including networks from other domains, such as biological and technological networks, as well.

generated much enthusiasm as an explanation for this supposedly universal feature of networks, but no such single process explains, for example, sexual contact network formation, in which behavioral heterogeneity is important (Handcock and Jones, 2004).

Universal laws are unlikely to be found if nobody looks for them, and perhaps this example is another illustration of how, as Martin suggests in the case of ERGM, we might learn more from failure than from “success”.

Appendix A: Details of failed models not included in the manuscript

Patricia 1990: all included.

Patricia 1992: all included.

Patricia 1993: a model that included a homophily (nodematch) term for the “Behind” component did not converge, because the “Behind” nodes are a separate connected component, so nodes with this nodal covariate connect only to each other. As noted by Martin’s Reply, including this heteroplanar component complicates things, and in any case, just as for the Sphere, it is a mistake to use a nodal covariate for this structural feature. And hence despite the “superior” fit of Models 3 and 4, Model 2 is a better model. Model 2 does not fit geodesic cycles of length 3, 4, and 5 as well as Model 3 (Figure 7) – but as discussed perhaps it is through the lack of fit on these that we learn something, and this model is “overfit” due to the (mis)use of structural features as nodal covariates.

Grey’s Anatomy: a model including the triangle term did not converge. As discussed in my paper (and the citations for this data therein) this is because there are in fact no odd-numbered cycles at all (and hence no triangles) in this network. So the “out-of-the-box” model which includes a GWESP or some other triangle (closure) term cannot be used. The models included are based on those cited in the Table B2 caption.

Dolphins: a model where the decay parameter is estimated (curved ERGM) rather than fixed for GWDEGREE did not converge. Using the default (MCMCMLE) estimation algorithm, the models did not converge, so as noted the “stepping” algorithm (Hummel et al., 2012) was used.

Law firm friendship: a model including a nodefactor term for OFFICE (differential activity depending on office location) showed poor convergence diagnostics and was not included. Another model using separate degree terms for degrees 0 through 10 rather than GWDEGREE (in an attempt to improve the goodness-of-fit of the degree distribution) did not converge and was not included.

Zachary Karate club: all models included.

Kapferer tailor shop: a model with the decay parameter for GWESP and GWDSP set to 0.5 (the “out of the box” or default value I used in this work for other networks) rather than 0.25 did not converge. As noted in the Table B6 caption, the final models merely replicate those in Hummel et al. (2012), and therefore I used that algorithm and not the default MCMCMLE.

High school friendship: as well as already noted in the paper (Table B7 caption) that models without class homophily do not converge, a model with a nodefactor term on class (separate activity for each class) but no sex homophily, showed poor goodness-of-fit and was not included.

References

- Abbott, A. 1988. “Transcending General Linear Reality”. *Sociological Theory* 6(2):169-86.
- Barabási, A.L., and R. Albert 1999. “Emergence of Scaling in Random Networks”. *Science* 286(5439):509-12.
- Bearman, P.S., J. Moody, and K. Stovel. 2004. “Chains of Affection: The Structure of Adolescent Romantic and Sexual Networks”. *American Journal of Sociology* 110(1):44-91.
- Block, P., J. Hollway, C. Stadtfeld, J. Koskinen, and T.A.B. Snijders. 2019. ““Predicting” after Peeking into the Future: Correcting a Fundamental Flaw in the SAOM – TERGM Comparison of Leifeld and Cranmer (2019)”. *arXiv preprint arXiv:1911.01385v1*.
- Block, P., J. Koskinen, J. Hollway, C. Steglich, and C. Stadtfeld. 2018. “Change We Can Believe in: Comparing Longitudinal Network Models on Consistency, Interpretability and Predictive Power”. *Social Networks* 52:180-91.
- Block, P., C. Stadtfeld, and T.A.B. Snijders. 2019. “Forms of Dependence: Comparing SAOMs and ERGMs from Basic Principles”. *Sociological Methods & Research* 48(1):202-39.
- Broido, A.D., and A. Clauset. 2019. “Scale-Free Networks are Rare”. *Nature Communications* 10(1):1017.
- Butts, C.T. 2017. “Comment: Actor Orientation and Relational Event Models”. *Sociological Methodology* 47(1):47-56.
- Colomer-de-Simón, P., and M. Boguñá. 2014. “Double Percolation Phase Transition in Clustered Complex Networks”. *Physical Review X* 4(4):041020.
- Colomer-de-Simón, P., M.A. Serrano, M.G. Beiró, J.I. Alvarez-Hamelin, and M. Boguñá. 2013. “Deciphering the Global Organization of Clustering in Real Complex Networks”. *Scientific Reports* 3(1):2517.
- Fritz, C., M. Lebacher, and G. Kauermann. 2020. “Tempus Volat, Hora Fugit: A Survey of Tie-oriented Dynamic Network Models in Discrete and Continuous Time”. *Statistica Neerlandica* 74(3):275-99.
- Gondal, N., and P.D. McLean. 2013. “What Makes a Network Go Round? Exploring the Structure of a Strong Component with Exponential Random Graph Models”. *Social Networks* 35(4):499-513.
- Handcock, M.S., and J.H. Jones. 2004. “Likelihood-based Inference for Stochastic Models of Sexual Network Formation”. *Theoretical Population Biology* 65(4):413-22.
- Hummel, R.M., D.R. Hunter, and M.S. Handcock. 2012. “Improving Simulation-based Algorithms for Fitting ERGMs”. *Journal of Computational and Graphical Statistics* 21(4):920-39.
- Hunter, D.R. 2007. “Curved Exponential Family Models for Social Networks”. *Social Networks* 29(2):216-30.
- Leifeld, P., and S.J. Cranmer. 2019. “A Theoretical and Empirical Comparison of the Temporal Exponential Random Graph Model and the Stochastic Actor-oriented Model”. *Network Science* 7(1):20-51.
- Levy, M.A. 2016. “gwdegree: Improving Interpretation of Geometrically-Weighted Degree Estimates in Exponential Random Graph Models”. *Journal of Open Source Software* 1(3):36.
- Lusher, D., and G. Robins. 2013. “Formation of Social Network Structure”. In *Exponential Random Graph Models for Social Networks*, edited by D. Lusher, J. Koskinen and G. Robins, 16-28. New York: Cambridge University Press.
- Mahadevan, P., D. Krioukov, K. Fall, and A. Vahdat. 2006. “Systematic Topology Analysis and Generation Using Degree Correlations”. *ACM SIGCOMM Computer Communication Review* 36(4):135-46.

- Martin, J. L. 2017. "The Structure of Node and Edge Generation in a Delusional Social Network". *Journal of Social Structure* 18(1):1-21.
- Orsini, C., M.M. Dankulov, P. Colomer-de-Simón, A. Jamakovic, P. Mahadevan, A. Vahdat, ... and S. Fortunato. 2015. "Quantifying Randomness in Real Networks". *Nature Communications* 6(1):8627.
- Pattison, P., and T.A.B. Snijders. 2013. "Modeling Social Networks: Next Steps". In *Exponential Random Graph Models for Social Networks*, edited by D. Lusher, J. Koskinen and G. Robins, 287-301. New York: Cambridge University Press.
- Robins, G., and G. Daraganova. 2013. "Social Selection, Dyadic Covariates, and Geospatial Effects". In *Exponential Random Graph Models for Social Networks*, edited by D. Lusher, J. Koskinen and G. Robins, 91-101. New York: Cambridge University Press.
- Robins, G., P. Elliott, and P. Pattison. 2001. "Network Models for Social Selection Processes". *Social Networks*, 23(1):1-30.
- Rolls, D.A., P. Wang, E. McBryde, P. Pattison, and G. Robins. 2015. "A Simulation Study Comparing Epidemic Dynamics on Exponential Random Graph and Edge-Triangle Configuration Type Contact Network Models". *PLoS ONE* 10(11):e0142181.
- Snijders, T.A.B. 2017. "Comment: Modeling of Coordination, Rate Functions, and Missing Ordering Information". *Sociological Methodology* 47(1):41-7.
- Snijders, T.A.B., P.E. Pattison, G.L. Robins, and M.S. Handcock. 2006. "New Specifications for Exponential Random Graph Models". *Sociological Methodology* 36(1):99-153.
- Stadtfeld, C., and P. Block. 2017. "Interactions, Actors, and Time: Dynamic Network Actor Models for Relational Events". *Sociological Science* 4:318-52.
- Stadtfeld, C., J. Hollway, and P. Block. 2017a. "Dynamic Network Actor Models: Investigating Coordination Ties Through Time". *Sociological Methodology* 47(1):1-40.
- Stadtfeld, C., J. Hollway, and P. Block. 2017b. "Rejoinder: DyNAMs and the Grounds for Actor-Oriented Network Event Models". *Sociological Methodology* 47(1):56-67.
- Stumpf, M.P., and M.A. Porter. 2012. "Critical Truths about Power Laws". *Science*, 335(6069):665-6.