Hierarchial Exponential Random Graph Models

Manasa Bollavaram

June 9, 2017

1 Introduction

In large networks, exponential random graph models (ergms) with global dependency often lead to degeneracy issues. These degeneracy issues can be solved by introducing local dependency in networks leading to networks local in nature.

Hierarchial Random Graph Models (Hergms) come into picture to solve the degeneracy issues induced by global dependence. Fig 1. shows the classification of Hergms. If dyadic independence is assumed to model a network, then based on whether the memberships of the network are observed or unobserved, we can employ a multi-level ergm or stochastic block model. Accordingly, if we assume dyadic dependence and we would like to estimate the memberships, then it's next-generation model.

2 Next-generation Model

Next-generation models are employed to estimate the unobserved neighborhood structure. These models are Hergms with local-dependence. Ergms with global dependence can be thought of as a special case of next-generation models with one neighborhood. One drawback of next-generation models compared to ergms with global dependence can be greater computation time.

3 Local vs Global Dependence

Modeling networks with local dependency is preferred to global dependency based on the following reasons

Figure 1: Classification of Hergms

	Dyadic Independence	Dyadic dependence
Observed Memberships	Multi-level Ergm	Multi-level Ergm
Unobserved Memberships	Stochastic Block Model	Next-Generation Model

- Scientific appeal: Many studies have shown that networks are local in nature
- Probabilistic advantages: Modeling networks locally induces weak dependence and solves the degeneracy issues as long as the sizes of the neighborhoods are bounded
- Statistical advantages: Hergms with local dependency have shown to be consistent with estimates compared to ergms with global dependence

4 Local Dependence

Ergms with global dependence are defined by the Equation (1).

$$P_\theta(X = x) = exp(< \theta - s(x) > -\psi(\theta)), x \in X \tag{1}$$

Local Dependence is defined as follows

If we can partition a set of nodes A into $K \geq 2$ non-empty subsets of nodes $A_1, A_2, A_3, ..., A_k$, called neighborhoods, then there exists local dependency in the network.

Model with the local dependency is defined by the Equation (2).

$$P_\theta(X = x) = \Pi_{k=1}^{K} P_\theta(X_{A_k,A_k} = x_{A_k,A_k}) \Pi_{l=k+1}^{K} \Pi_{i \in A_k < j \in A_l} P_\theta(X_{i,j} = x_{i,j}, X_{j,i} = x_{j,i}) \tag{2}$$

where X_{A_k,A_k} denotes within-neighborhood subgraph induced by neighborhood A_k.

Dependence induced by the model defined in Equation (2) is restricted to within-neighborhood subgraphs.

5 Models to estimate the neighborhood structure

Hierarchial Exponential Random Graph Model

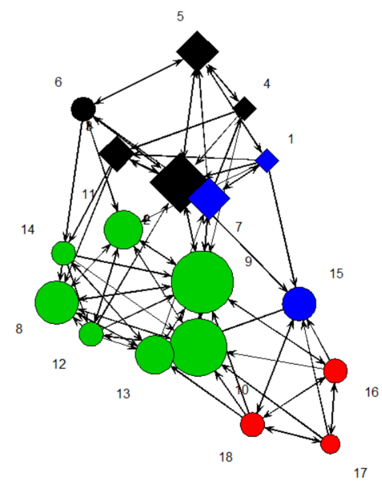
$$P_\theta(X = x, Z = z) = P(Z = z) \Pi_{k=1}^{K} P_\theta(X_{A_k,A_k} = x_{A_k,A_k} | Z = z) . \Pi_{l=k+1}^{K} \Pi_{i \in A_k < j \in A_l} P_\theta(X_{i,j} = x_{i,j}, X_{j,i} = x_{j,i} | Z \tag{3}$$

where $Z = (Z_1, Z_2, ..., Z_n)$ are the neighborhood memberships of nodes $1, 2, ..., n$ and $Z_i = k$ implies node i is a membership of neighborhood $A_k, k = 1, 2, ...K$

Conditional distributions of within- and between- neighborhood subgraphs given $Z = z$ are considered from exponential families of distributions

$$P_\theta(X = x | Z = z) = exp(< \eta(\theta, z), s(x, z) > -\psi(\theta, z)) \tag{4}$$

Figure 2: Sampson's network - colored by group, sizes based on in-degree, shape based on whether the monks attended cloisterville



Hergm with reciprocity and transitivity is given by

$$P_{\theta}(X=x|Z=z) \propto exp(\sum_{i,j}^{n} \eta_{1,i,j}(\theta,z)x_{i,j} + \sum_{i<j}^{n} \eta_{2,i,j}(\theta,z)x_{i,j}x_{j,i} + \sum_{i,j,k}^{n} \eta_{3,i,j,k}(\theta,z)x_{i,j}x_{j,k}x_{i,k}) \tag{5}$$

Stochastic Block Model

$P_{\theta}(X=x|Z=z) \propto exp(\sum_{i<j}^{n} \eta_{i,j}(\theta,z)x_{i,j})$

Stochastic Block Models are useful to estimate the neighborhood structure in the network data, assuming dyadic independence within neighborhoods.

6 Sampson's Dataset

Sampson collected directed network data for 18 monks in a monastery, where an edge from node i to node$j$, represents that monk i likes monk j. The network has a total of 88 edges. Node attributes are Names, Group, Cloisterville (cloisterville is a seminary which either a monk attended or not before coming to the monastery). Sampson divided the monks into four groups - loyals, turks, outcasts and waverers. Loyals (nodes 2,3,4,5,6) are the people who have been following the orthodox monastery rules. Turks (nodes 8,9,10,11,12,13,14) are the monks who joined the monastery later and question the orthodox rules followed by loyals. Waverers (nodes 1,7,15) are the monks who did not take sides in the debate. Outcasts (nodes 16,17,18) are the monks not accepted in the group. A plot of the network data is represented in figure 2.

7 Various Models fitted to the data

Figures 3-19 provide summary, goodness-of-fit and sample simulated network of the below models fitted to the data.

1. Unobserved memberships, allowing maximum blocks (18)

   - model5 ← hergm(samplike ~edges_ij + mutual_ij + ttriple_ijk,max_number = 18, relabel = 3, sample_size = 2e+5)
   - Estimated the neighborhood structure by allowing for the maximum blocks. The model identifies only three neighborhoods as can be seen in Fig. 5(b). It is hard to see from Fig. 5(b) the neighborhood memberships of the nodes but from the simulated network it was easy to identify the node memberships.
   - Estimated neighborhood memberships:
     - Loyals - 1,2,3,4,5,6,7
     - Turks - 8,9,10,11,12,13,14
     - Outcasts - 15,16,17,18
   - Fig. 5(b) shows the neighborhoods but not the network. Since, a large number of blocks, 18, are allowed in the model, the option relabel=3 has been used to reduce the computation time. Relabel=3 solves the popular label switching problem in network data analysis.

2. Unobserved memberships, restricting to 4 blocks

   - model4← hergm(samplike ~edges_ij + mutual_ij + ttriple_ijk,max_number = 4, sample_size = 2e+5)
   - Tried forcing the blocks to 4, to see whether the model could identify the four groups. But still the model estimates only three groups. In most of the models specified below, only three groups are recovered. By specifying additional vertex information, we might get the models to recover all the four groups.
   - From Figure 3(a), we can deduce that there is a weak tendency to form ties between groups which is highly encouraged by reciprocity
   - From Figure 3(a), a general tendency to form ties is greater, again enhanced by reciprocity and transitivity. In the following models, we see a different tendency.

3. Observed loyal group

   - indicator3 ← c(rep.int(NA, 1), rep.int(2, 5), rep.int(NA, 12))
     model3 ← hergm(samplike ~edges_ij + mutual_ij + ttriple_ijk,max_number = 4,indicator=indicator3, sample_size = 2e+5)
   - It would be interesting to check if the goodness of fit would improve if one group is observed. The model still recovers only three groups. The goodness of fit is almost the same as unobserved memberships. Details on the change in goodness of fit is given in the Section 8.

4. Observed Turk group

   - indicator6 ← c(rep.int(NA, 7), rep.int(1, 7), rep.int(NA, 4)) model6 ← hergm(samplike ~edges_ij + mutual_ij + ttriple_ijk,max_number = 4,indicator=indicator6, sample_size = 2e+5)
   - It would be interesting to note the changes in goodness-of-fit if the observed group is different. Model fits slightly better if the observed group is Turk compared to Loyals as can be seen from Figure 20. Model fit is only slightly better or can be stated as almost the same.

5. Multi-level Ergm

   - indicator2 ← c(rep.int(1, 1), rep.int(2, 5), rep.int(1, 1), rep.int(3,7), rep.int(1,1), rep.int(4,3)) model2 ← hergm(samplike ~edges_ij + mutual_ij + ttriple_ijk,max_number = 4,indicator=indicator2, sample_size = 2e+5)
   - All the previous models recovered only three groups, it would be interesting to specify all the groups and note the differences in model fits. It performs better than one observed membership but poorly to unobserved memberships.
   - From Figure 11(a), it can be deduced that it is a weak tendency to form a tie between groups but which again is greatly enhanced by reciprocity
   - From Figure 11(a), in a generalized statement, within groups, tendency to form ties is strong and which is enhanced by reciprocity and either discouraged or a little tendency for transitivity

6. Ergm with global dependence

   - indicator ← rep.int(1, 18)
     model1 ← hergm(samplike ~edges_ij + mutual_ij + ttriple_ijk,max_number = 1,indicator=indicator, sample_size = 2e+5)
   - Comparing the model fits to one membership would assess local dependency fit vs global dependency fit.
   - From figure 14, it is clear that edges, 2-out stars and transitive triples are modeled better by local dependence models compared to global dependence
   - From figure 13, summary of parameters, it indicates a weak tendency to form a tie and the log-odds is greatly increased in the case of reciprocity and only slight increase in log-odds in the case of transitivity

7. Ergm with homophily terms

   - indicator ← rep.int(1, 18)
     model11 ← hergm(samplike ~edges_ij + mutual_ij + ttriple_ijk +gwesp(decay=.5,fixed=TRUE)+gwidegree(d
     = 1,indicator=indicator,sample_size = 2e+5)

- It would be interesting to see if adding the nodematch terms on cloisterville, gwesp terms would improve the fit of edges
- From figure 15, it can be seen that the fit definitely improves compared to the previous model but still performs poorly compared to local dependence models
- Model performs better in terms of geodesic distances and out-degrees compared to all the models with local dependency
- Including homophily terms etc. in local dependency terms might also improve the fit of geo-desics and out-degrees
- nodematch on group has not been added in the model because we are trying to model the data in a case where we do not know the memberships of nodes, a more often case in large networks

8. Hierarchial Bernoulli Model

   - indicator7 ← c(rep.int(1, 1), rep.int(2, 5), rep.int(1, 1), rep.int(3,7), rep.int(1,1), rep.int(4,3)) model7 ← hergm(samplike ~edges_ij,max_number = 4,indicator=indicator7, sample_size = 2e+5)
   - With the observed memberships, bernoulli models assumes dyadic independence within neighborhoods
   - The goodness-of-fit is almost the same as previous local dependency models, excluding the markov dependency does not seem to have an affect on this network data

9. Stochastic Block Model

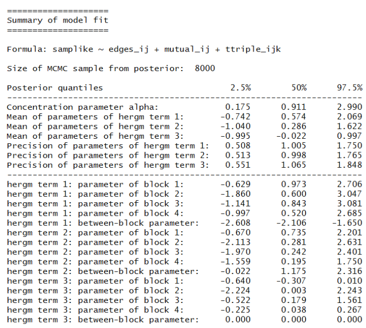
   - model8 ← hergm(samplike ~edges_ij,max_number = 18, relabel = 3, sample_size = 2e+5)
   - Stochastic block models are quicker to estimate unobserved neighborhood structure compared to next-generation models
   - Stochastic Block Model could also estimate only three groups
   - In this network data, Stochastic block models perform equal to next-generation models
   - In large networks, the fits might vary dramatically

8 Goodness-of-fits of Models

Figure 20 provides a comparison of goodness-of-fit of models described in Section 8. Columns of Figure 20 indicate root mean squared deviation of posterior predicted number of statistic to observed statistic.

Clearly ergms with global dependence do not capture the network structure but rest all ergms with local dependence fit perfect. There is a minimum deviation in goodness-of-fit with the rest of local dependent ergms, probably because of the small network and we may observe prominent deviations in large networks. The interesting point to note, is that even for small networks, local dependence triumphs over global dependence and in large networks local dependence would do wonders in fitting. In that direction, hergms are the models to pick for modeling large networks. However, local dependency models take a lot of time compared to global dependence. Decision of picking models would depend on time vs accuracy of fit.

Figure 3: Next-generation model - unobserved memberships, restricting to 4 blocks



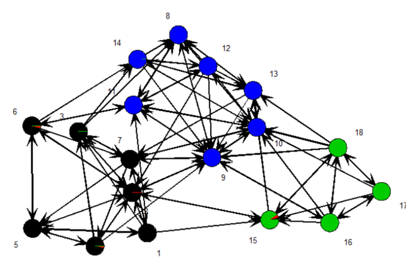
```
===================
Summary of model fit
===================

Formula: samplike ~ edges_ij + mutual_ij + ttriple_ijk

Size of MCMC sample from posterior:  8000

Posterior quantiles                            2.5%      50%     97.5%
----------------------------------------------------------------------
Concentration parameter alpha:                0.175     0.911     2.990
Mean of parameters of hergm term 1:          -0.742     0.574     2.069
Mean of parameters of hergm term 2:          -1.040     0.286     1.622
Mean of parameters of hergm term 3:          -0.995    -0.022     0.997
Precision of parameters of hergm term 1:      0.508     1.005     1.750
Precision of parameters of hergm term 2:      0.513     0.998     1.765
Precision of parameters of hergm term 3:      0.551     1.065     1.848
----------------------------------------------------------------------
hergm term 1: parameter of block 1:          -0.629     0.973     2.706
hergm term 1: parameter of block 2:          -1.860     0.600     3.047
hergm term 1: parameter of block 3:          -1.141     0.843     3.081
hergm term 1: parameter of block 4:          -0.997     0.520     2.685
hergm term 1: between-block parameter:       -2.608    -2.106    -1.650
hergm term 2: parameter of block 1:          -0.670     0.735     2.201
hergm term 2: parameter of block 2:          -2.113     0.281     2.631
hergm term 2: parameter of block 3:          -1.970     0.242     2.401
hergm term 2: parameter of block 4:          -1.559     0.195     1.750
hergm term 2: between-block parameter:       -0.022     1.175     2.316
hergm term 3: parameter of block 1:          -0.640    -0.307     0.010
hergm term 3: parameter of block 2:          -2.224     0.003     2.243
hergm term 3: parameter of block 3:          -0.522     0.179     1.561
hergm term 3: parameter of block 4:          -0.225     0.038     0.267
hergm term 3: between-block parameter:        0.000     0.000     0.000
----------------------------------------------------------------------
```

(a) Summary



(b) Simulated network

9 References

1. Schweinberger, Michael, and Pamela Luna. "HERGM: Hierarchical exponential-family random graph models." Journal of Statistical Software (2015).

2. Hunter, David R., Steven M. Goodreau, and Mark S. Handcock. "Goodness of fit of social network models." Journal of the American Statistical Association 103.481 (2008): 248-258.

3. Holland, Paul W., Kathryn Blackmond Laskey, and Samuel Leinhardt. "Stochastic blockmodels: First steps." Social networks 5.2 (1983): 109-137.

4. Krivitsky, Pavel N., et al. "Representing degree distributions, clustering, and homophily in social networks with latent cluster random effects models." Social networks 31.3 (2009): 204-213.

Figure 4: Goodness-of-fit of next-generation model - unobserved memberships, restricting to 4 blocks

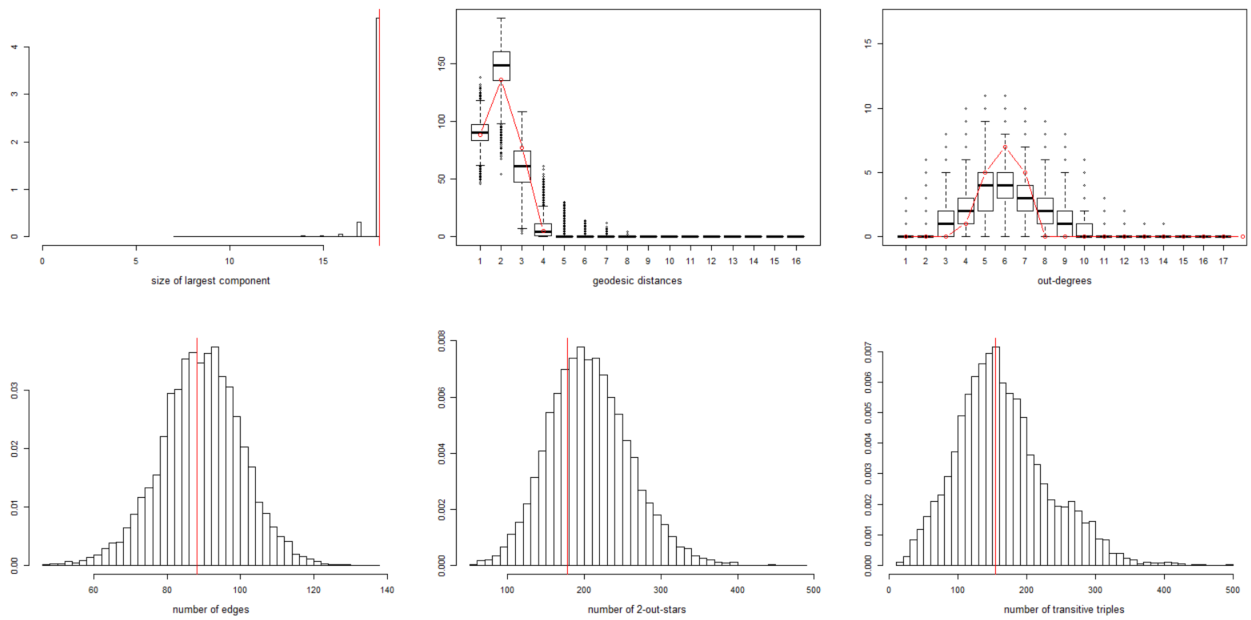


Figure 5: Next-generation model - unobserved memberships, allowing 18 blocks

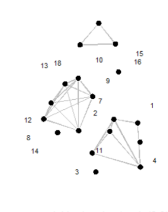
```
====================
Summary of model fit
====================

Formula: samplike ~ edges_ij + mutual_ij + ttriple_ijk

Size of MCMC sample from posterior:  8000

Posterior quantiles                                2.5%      50%     97.5%
---------------------------------------------------------------------------
Concentration parameter alpha:                    0.410    1.220     3.613
Mean of parameters of hergm term 1:              -0.803    0.593     2.095
Mean of parameters of hergm term 2:              -1.138    0.260     1.648
Mean of parameters of hergm term 3:              -1.038    0.005     1.034
Precision of parameters of hergm term 1:          0.569    1.010     1.703
Precision of parameters of hergm term 2:          0.569    1.012     1.686
Precision of parameters of hergm term 3:          0.584    1.042     1.732
---------------------------------------------------------------------------
hergm term 1: parameter of block 1:              -1.170    0.869     3.073
hergm term 1: parameter of block 2:              -1.552    0.657     2.980
hergm term 1: parameter of block 3:              -1.161    0.695     2.839
hergm term 1: parameter of block 4:              -1.551    0.712     3.013
hergm term 1: parameter of block 5:              -1.739    0.610     3.071
hergm term 1: parameter of block 6:              -1.628    0.639     3.003
hergm term 1: parameter of block 7:              -1.807    0.590     3.052
hergm term 1: parameter of block 8:              -1.755    0.609     3.077
hergm term 1: parameter of block 9:              -1.839    0.606     2.986
hergm term 1: parameter of block 10:              -1.794    0.616     3.060
hergm term 1: parameter of block 11:              -1.803    0.589     3.067
hergm term 1: parameter of block 12:              -1.826    0.605     3.036
hergm term 1: parameter of block 13:              -1.850    0.620     3.090
hergm term 1: parameter of block 14:              -1.784    0.608     3.107
hergm term 1: parameter of block 15:              -1.740    0.628     3.092
hergm term 1: parameter of block 16:              -1.767    0.610     3.092
hergm term 1: parameter of block 17:              -1.781    0.595     3.056
hergm term 1: parameter of block 18:              -1.795    0.606     3.022
hergm term 1: between-block parameter:           -2.560   -2.053    -1.546
hergm term 2: parameter of block 1:              -1.731    0.477     2.401
hergm term 2: parameter of block 2:              -1.972    0.387     2.554
hergm term 2: parameter of block 3:              -1.617    0.261     2.075
hergm term 2: parameter of block 4:              -2.014    0.339     2.494
```

(a) Summary



(b) Neighborhood structure

Figure 6: Goodness-of-fit of next-generation model - unobserved memberships, allowing 18 blocks

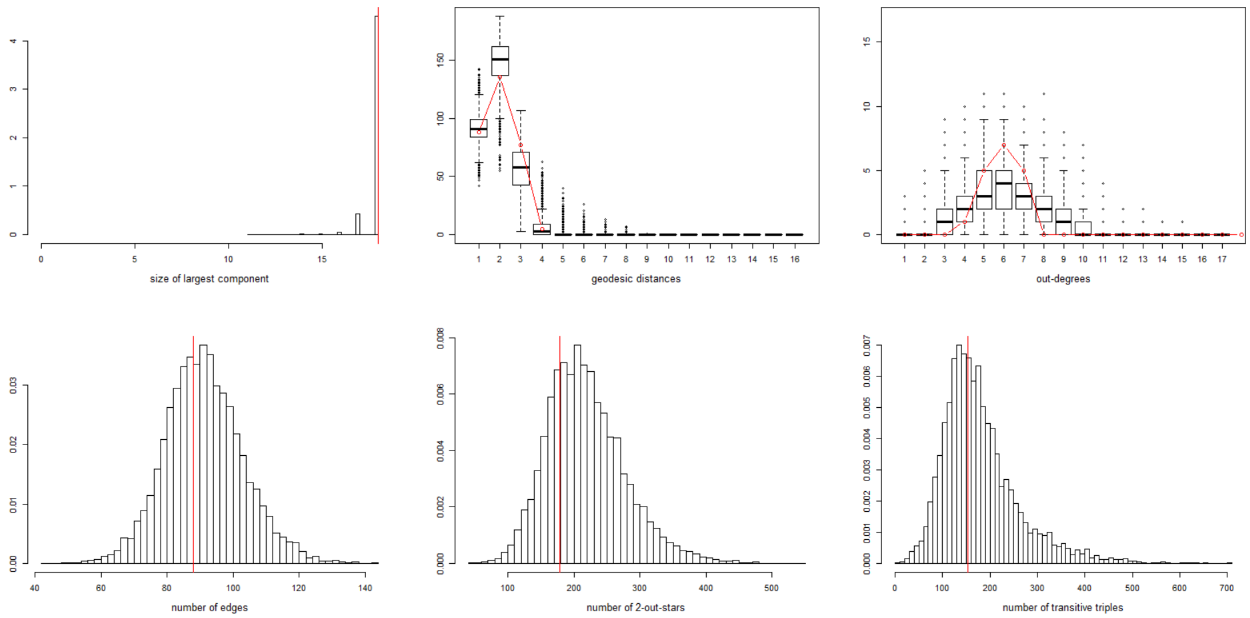


Figure 7: next-generation model - observed loyal group

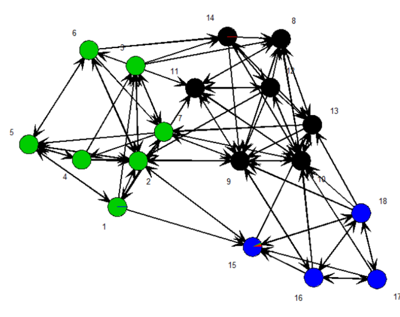
```
====================
Summary of model fit
====================

Formula: samplike ~ edges_ij + mutual_ij + ttriple_ijk

Size of MCMC sample from posterior:  8000

Posterior quantiles                             2.5%       50%     97.5%
-------------------------------------------------------------------------
Concentration parameter alpha:                 0.184     0.893     3.012
Mean of parameters of hergm term 1:           -0.763     0.629     2.050
Mean of parameters of hergm term 2:           -1.037     0.289     1.658
Mean of parameters of hergm term 3:           -1.039    -0.033     1.046
Precision of parameters of hergm term 1:       0.521     0.998     1.718
Precision of parameters of hergm term 2:       0.519     1.000     1.737
Precision of parameters of hergm term 3:       0.541     1.062     1.820
-------------------------------------------------------------------------
hergm term 1: parameter of block 1:           -0.966     0.538     2.573
hergm term 1: parameter of block 2:           -1.786     0.623     3.036
hergm term 1: parameter of block 3:           -0.553     1.002     2.808
hergm term 1: parameter of block 4:           -1.146     0.914     3.071
hergm term 1: between-block parameter:        -2.610    -2.121    -1.674
hergm term 2: parameter of block 1:           -1.425     0.183     1.715
hergm term 2: parameter of block 2:           -2.065     0.293     2.718
hergm term 2: parameter of block 3:           -0.668     0.747     2.198
hergm term 2: parameter of block 4:           -1.899     0.258     2.403
hergm term 2: between-block parameter:         0.000     1.200     2.331
hergm term 3: parameter of block 1:           -0.211     0.033     0.243
hergm term 3: parameter of block 2:           -2.228    -0.051     2.140
hergm term 3: parameter of block 3:           -0.632    -0.316     0.002
hergm term 3: parameter of block 4:           -0.525     0.167     1.248
hergm term 3: between-block parameter:         0.000     0.000     0.000
-------------------------------------------------------------------------
```

(a) Summary



(b) Simulated network

Figure 8: Goodness-of-fit of next-generation model - observed turk group

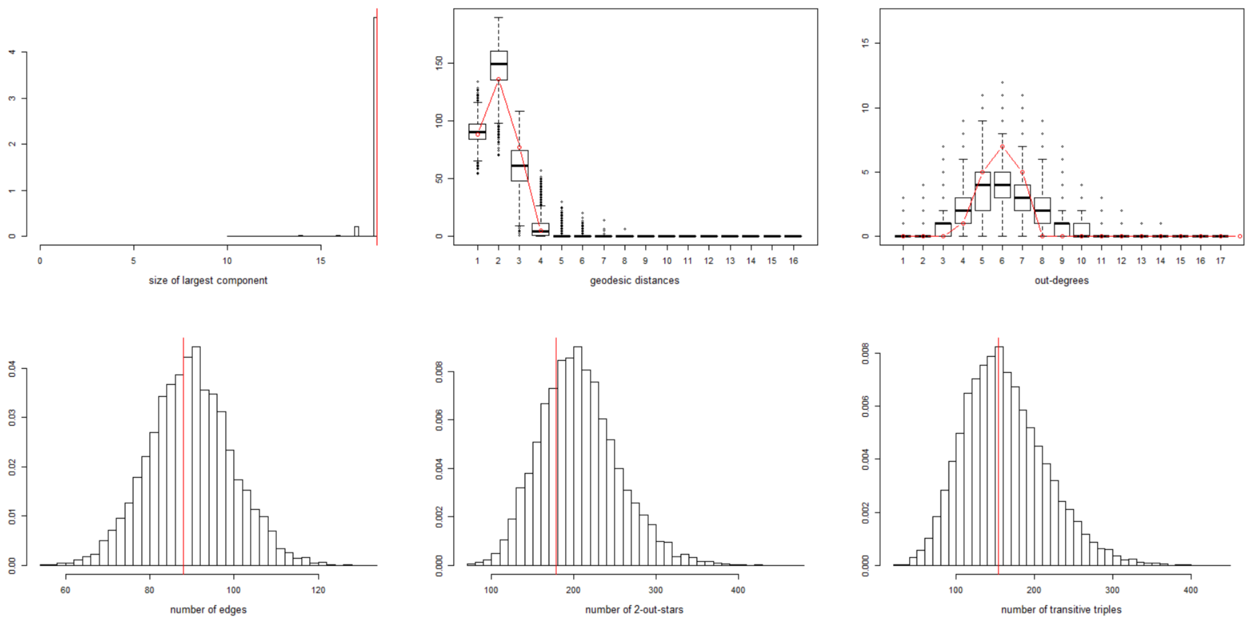


Figure 9: Next-generation model - observed turk group

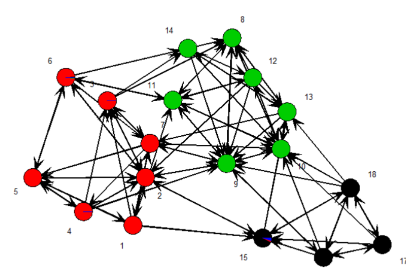
```
====================
Summary of model fit
====================

Formula: samplike ~ edges_ij + mutual_ij + ttriple_ijk

Size of MCMC sample from posterior:  8000

Posterior quantiles                              2.5%       50%     97.5%
-------------------------------------------------------------------------
Concentration parameter alpha:                  0.228     1.297     4.012
Mean of parameters of hergm term 1:            -0.791     0.579     1.970
Mean of parameters of hergm term 2:            -1.090     0.289     1.630
Mean of parameters of hergm term 3:            -1.034    -0.015     0.986
Precision of parameters of hergm term 1:        0.510     1.003     1.742
Precision of parameters of hergm term 2:        0.518     1.010     1.743
Precision of parameters of hergm term 3:        0.546     1.064     1.841
-------------------------------------------------------------------------
hergm term 1: parameter of block 1:            -1.085     0.833     3.016
hergm term 1: parameter of block 2:            -0.585     0.951     2.704
hergm term 1: parameter of block 3:            -1.025     0.445     2.320
hergm term 1: parameter of block 4:            -1.786     0.602     2.920
hergm term 1: between-block parameter:         -2.588    -2.117    -1.658
hergm term 2: parameter of block 1:            -1.939     0.225     2.365
hergm term 2: parameter of block 2:            -0.639     0.733     2.149
hergm term 2: parameter of block 3:            -1.370     0.233     1.681
hergm term 2: parameter of block 4:            -2.136     0.322     2.679
hergm term 2: between-block parameter:         -0.029     1.186     2.339
hergm term 3: parameter of block 1:            -0.504     0.183     1.303
hergm term 3: parameter of block 2:            -0.643    -0.304     0.008
hergm term 3: parameter of block 3:            -0.198     0.042     0.255
hergm term 3: parameter of block 4:            -2.218    -0.034     2.136
hergm term 3: between-block parameter:          0.000     0.000     0.000
-------------------------------------------------------------------------
```

(a) Summary



(b) Simulated network

Figure 10: Goodness-of-fit of next-generation model - observed turk group

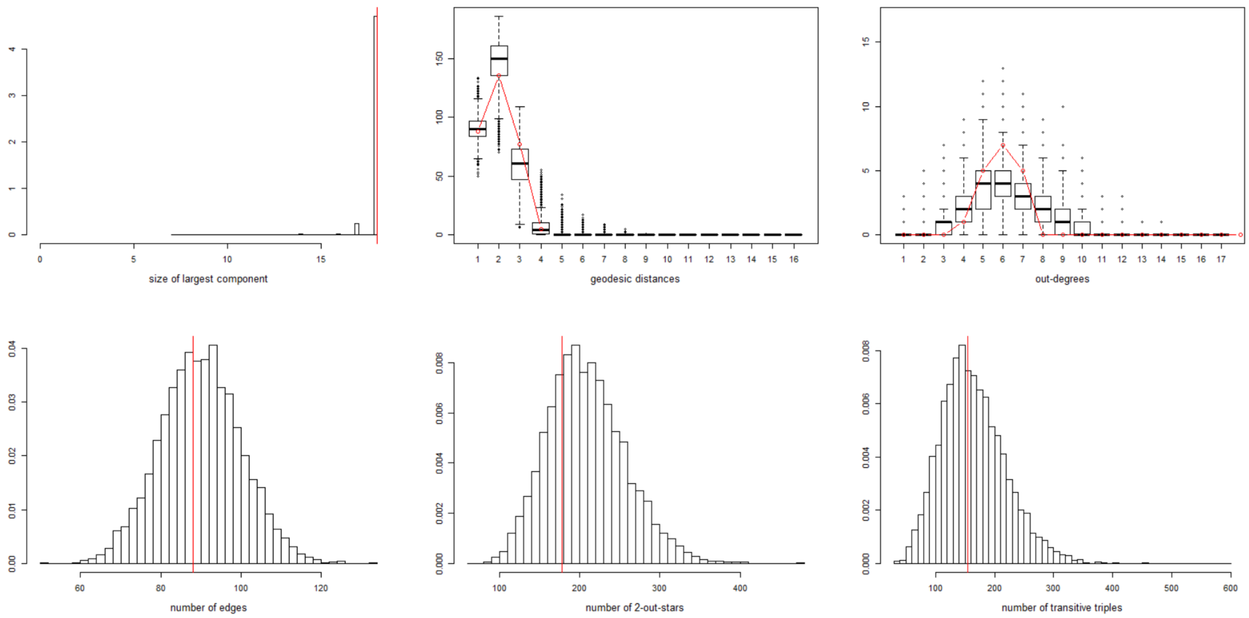


Figure 11: Multi-level ergm - observed memberships

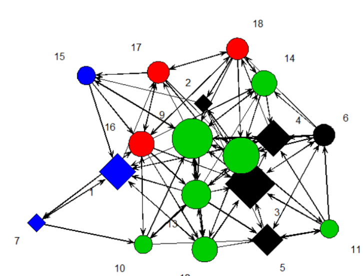
```
====================
Summary of model fit
====================

Formula: samplike ~ edges_ij + mutual_ij + ttriple_ijk

Size of MCMC sample from posterior:  8000

Posterior quantiles                             2.5%        50%      97.5%
----------------------------------------------------------------------------
Concentration parameter alpha:                 0.380      1.340      3.470
Mean of parameters of hergm term 1:           -0.950      0.345      1.716
Mean of parameters of hergm term 2:           -0.727      0.644      2.035
Mean of parameters of hergm term 3:           -1.045     -0.016      0.991
Precision of parameters of hergm term 1:       0.515      1.019      1.748
Precision of parameters of hergm term 2:       0.498      0.989      1.734
Precision of parameters of hergm term 3:       0.505      1.002      1.763
----------------------------------------------------------------------------
hergm term 1: parameter of block 1:           -1.539      0.165      2.064
hergm term 1: parameter of block 2:           -1.183      0.413      2.238
hergm term 1: parameter of block 3:           -1.088      0.360      2.225
hergm term 1: parameter of block 4:           -1.373      0.707      3.021
hergm term 1: between-block parameter:        -2.359     -1.927     -1.499
hergm term 2: parameter of block 1:           -1.529      0.594      2.763
hergm term 2: parameter of block 2:           -0.534      1.226      3.061
hergm term 2: parameter of block 3:           -1.224      0.425      2.078
hergm term 2: parameter of block 4:           -1.364      0.991      3.416
hergm term 2: between-block parameter:         0.306      1.363      2.371
hergm term 3: parameter of block 1:           -2.597     -0.729      0.754
hergm term 3: parameter of block 2:           -0.761     -0.251      0.193
hergm term 3: parameter of block 3:           -0.202      0.033      0.235
hergm term 3: parameter of block 4:           -0.506      0.887      2.700
hergm term 3: between-block parameter:         0.000      0.000      0.000
----------------------------------------------------------------------------
```

(a) Summary



(b) Simulated network

Figure 12: Goodness-of-fit of multi-level ergm - observed memberships

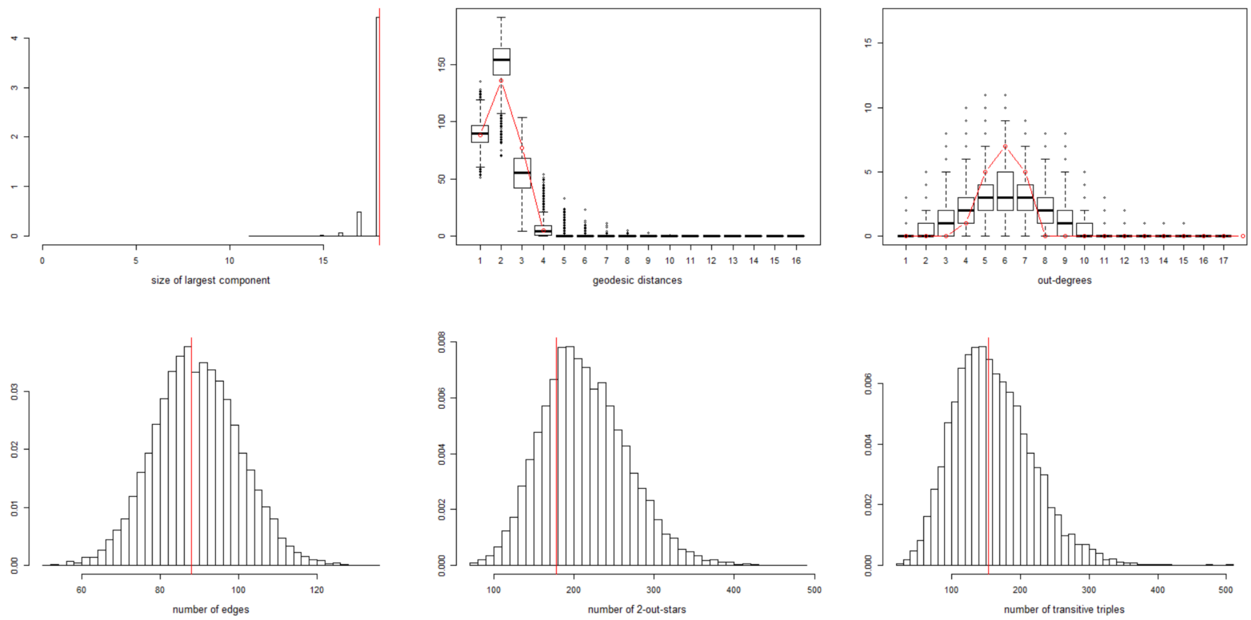


Figure 13: Summary of ergm with global dependence

```
Summary of model fit
====================

Formula: samplike ~ edges_ij + mutual_ij + ttriple_ijk

Size of MCMC sample from posterior:  8000

Posterior quantiles                             2.5%        50%      97.5%
-------------------------------------------------------------------------
Concentration parameter alpha:                 0.029      0.676      3.548
Mean of parameters of hergm term 1:           -2.286     -0.897      0.612
Mean of parameters of hergm term 2:           -0.530      1.004      2.474
Mean of parameters of hergm term 3:           -1.350      0.037      1.388
Precision of parameters of hergm term 1:       0.467      0.949      1.700
Precision of parameters of hergm term 2:       0.453      0.940      1.666
Precision of parameters of hergm term 3:       0.503      0.991      1.736
-------------------------------------------------------------------------
hergm term 1: parameter of block 1:           -2.329     -1.837     -1.325
hergm term 2: parameter of block 1:            1.286      2.042      2.919
hergm term 3: parameter of block 1:           -0.041      0.046      0.109
-------------------------------------------------------------------------
```

Figure 14: Goodness-of-fit of ergm with global dependence

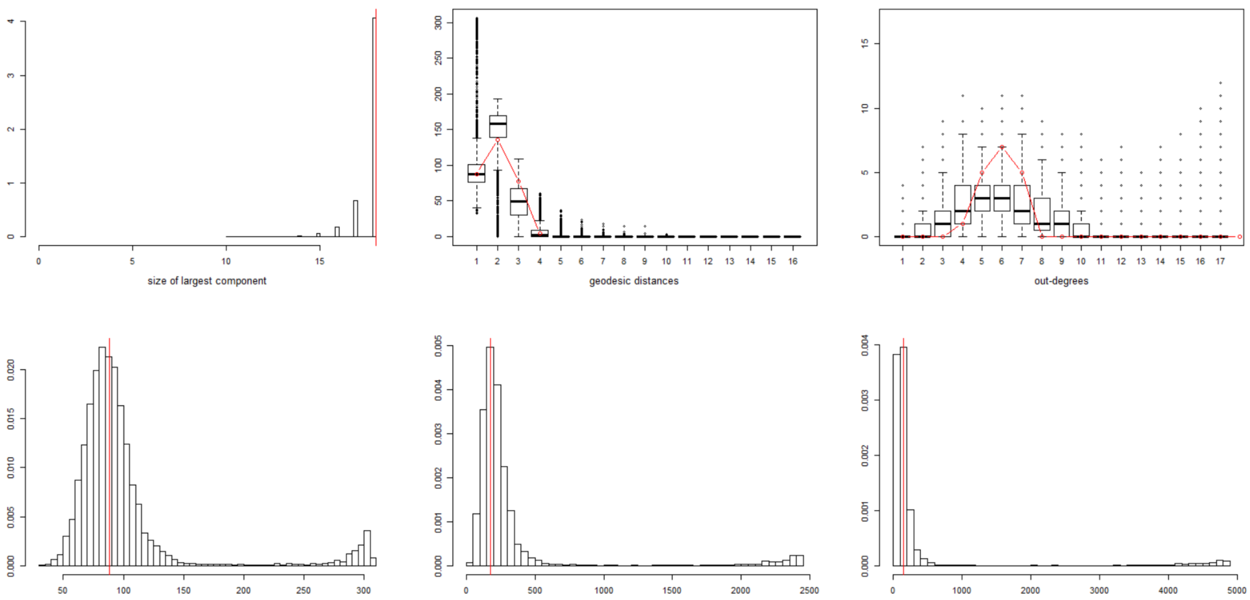


Figure 15: Goodness-of-fit of ergm with homophily terms

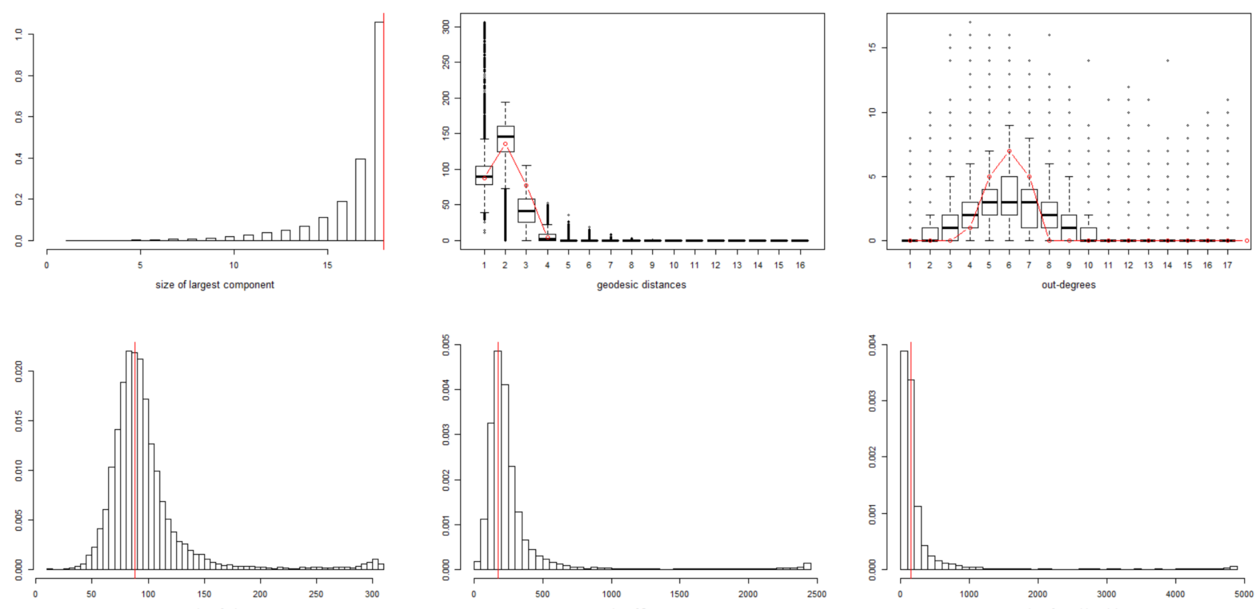


Figure 16: Hierarchial bernoulli model

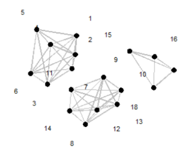
```
====================
Summary of model fit
====================

Formula: samplike ~ edges_ij

Size of MCMC sample from posterior:  8000

Posterior quantiles                             2.5%       50%     97.5%
------------------------------------------------------------------------
Concentration parameter alpha:                 0.211     1.138     3.589
Mean of parameters of hergm term 1:           -0.456     0.638     1.727
Precision of parameters of hergm term 1:       0.530     1.018     1.772
------------------------------------------------------------------------
hergm term 1: parameter of block 1:           -0.761     0.507     2.509
hergm term 1: parameter of block 2:           -0.916     0.627     2.646
hergm term 1: parameter of block 3:            0.266     0.922     1.774
hergm term 1: parameter of block 4:           -1.239     0.812     2.755
hergm term 1: between-block parameter:        -2.300    -1.852    -1.456
------------------------------------------------------------------------
```

(a) Summary



(b) Network neighborhood structure

Figure 17: Goodness-of-fit of hierarchial bernoulli model

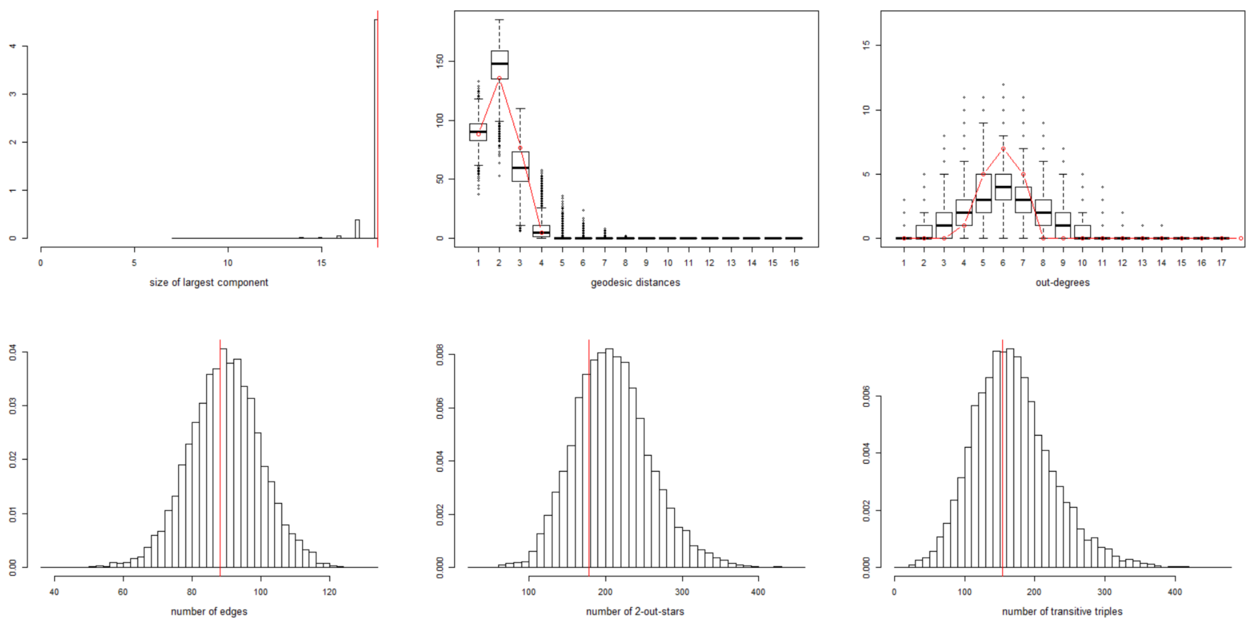


Figure 18: Stochastic block model

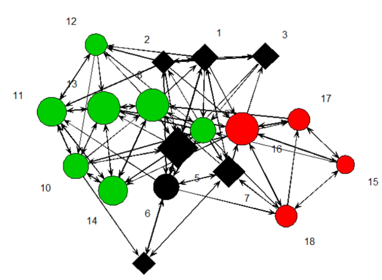
```
====================
Summary of model fit
====================

Formula: samplike ~ edges_ij

Size of MCMC sample from posterior:  8000

Posterior quantiles                             2.5%       50%     97.5%
------------------------------------------------------------------------
Concentration parameter alpha:                 0.376     0.910     2.668
Mean of parameters of hergm term 1:           -0.430     0.641     1.721
Precision of parameters of hergm term 1:       0.577     1.024     1.714
------------------------------------------------------------------------
hergm term 1: parameter of block 1:           -0.431     0.382     2.358
hergm term 1: parameter of block 2:            0.256     0.925     1.767
hergm term 1: parameter of block 3:           -0.735     0.982     2.782
hergm term 1: parameter of block 4:           -1.444     0.711     2.817
hergm term 1: parameter of block 5:           -1.655     0.649     2.812
hergm term 1: parameter of block 6:           -1.520     0.638     2.895
hergm term 1: parameter of block 7:           -1.609     0.662     2.830
hergm term 1: parameter of block 8:           -1.570     0.659     2.863
hergm term 1: parameter of block 9:           -1.597     0.633     2.900
hergm term 1: parameter of block 10:           -1.545     0.620     2.899
hergm term 1: parameter of block 11:           -1.603     0.641     2.870
hergm term 1: parameter of block 12:           -1.584     0.628     2.876
hergm term 1: parameter of block 13:           -1.587     0.649     2.892
hergm term 1: parameter of block 14:           -1.553     0.621     2.811
hergm term 1: parameter of block 15:           -1.586     0.654     2.916
hergm term 1: parameter of block 16:           -1.581     0.628     2.898
hergm term 1: parameter of block 17:           -1.578     0.638     2.924
hergm term 1: parameter of block 18:           -1.520     0.639     2.900
hergm term 1: between-block parameter:        -2.285    -1.843    -1.394
------------------------------------------------------------------------
```

(a) Summary



(b) Simulated network

Figure 19: Goodness-of-fit of Stochastic Block model

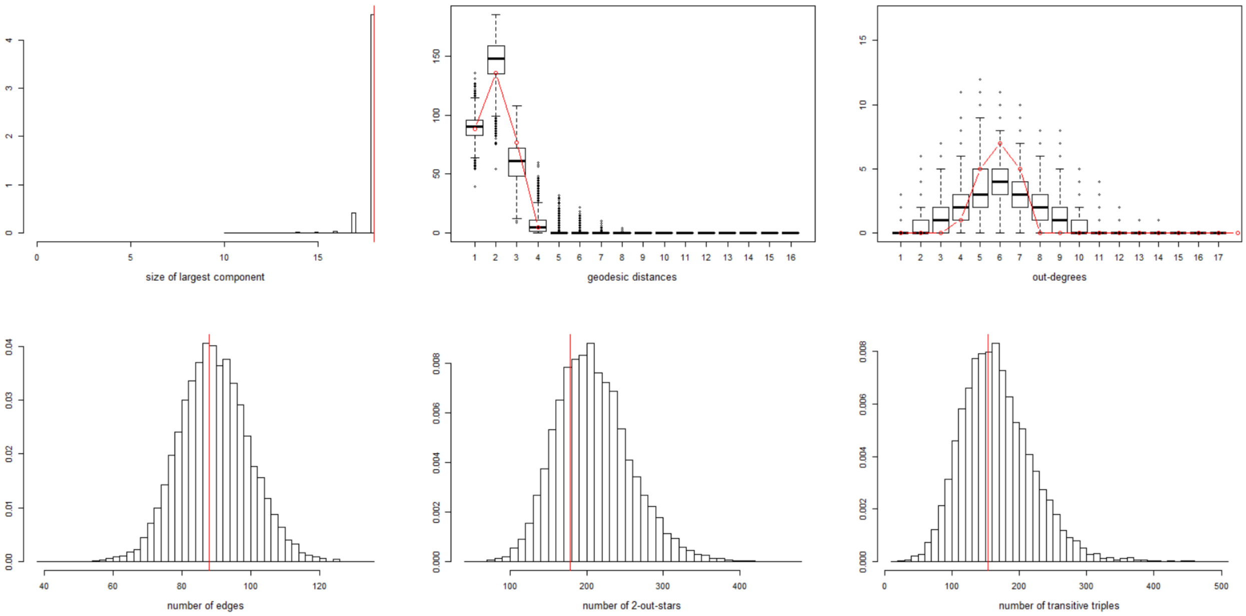


Figure 20: Goodness-of-fits comparison

Model	RMS Deviation - Edges	RMS Deviation – Transitive triples	RMS Deviation, 2-out stars
ERGM with global dependence	53.7	1036.9	517.9
Ergm with gwesp terms etc., global dependence	40.3	680.5	355.98
Next-generation model, 18 blocks	12.5	82.7	72.3
Next-generation model, 4 blocks	11.3	72.63	60.9
Multilevel ergm, observed memberships	10.9	65.4	63.7
Multilevel ergm, dyadic independence, Observed memberships	10.6	N.A.	58.5
Observed Turk Group, next-generation model	10.4	61.8	58.2
Stochastic Block Model, Unobserved memberships	10.3	N.A.	57.5
Observed loyal group, next-generation model	10.1	60.6	56.2