# Applying the Akaike Information Criterion to Degree-Corrected Stochastic Block Models

Jacob Jensen with Cris Moore and Cosma Shalizi

The Akaike Information Criterion (AIC) is a powerful and highly principled method of model selection. When working with probabilistic models it is typical to fit them to the data by maximum likelihood, using Markov Chain Monte Carlo or Belief Propagation. The maximum likelihood value obtained in this way gives an indication of how good a fit the model is to the data. Unfortunately, models with more parameters tend to fit the data strictly better, simply because they have more degrees of freedom and not necessarily because they describe it better. The AIC counters this by taking the $-\ell(\hat{\theta})$ maximum likelihood value and adding a k = $|\theta|$. The best model according to the AIC is then the one that minimizes $-\ell(\hat{\theta}) + k$. This simple penalty has a powerful justification, but makes certain assumptions that pose major problems to models used to infer networks. The first of these assumptions, that all parameters in $\theta$ are continuous, can be remedied by a monte carlo prodcedure that marginalizes over these variables. The second, which assumes narrow peak around in the model's likelihood, is normally easy to deal with given enough data to accurately estimate parameters. In the case of the Degree-Corrected block model, which includes an independently parametrized poisson distribution for every, low-degree nodes violate this assumption. We hope to overcome this problem and find an extension of the AIC that can handle this case.