Bayesian Analysis (2014) **9**, Number 3, pp. 561–590

Contributed Discussion on Article by Finegold and Drton

Comment by Guido Consonni[1] and Luca La Rocca[2]

This is a very interesting paper providing both theoretical and computational results for robust structure estimation in decomposable graphical models. Finegold & Drton (F&D hereafter) do a splendid job in motivating and illustrating the various ramifications of this attractive research path. We will comment on prior specification, hoping to add further insights to a paper already rich in content. Notice that model choice results strongly depend on prior specification; see, e.g., O'Hagan and Forster (2004, ch. 7).

Priors on graphs Formula (3) of F&D specifies a product of Bernoulli priors with fixed edge inclusion probability d. As F&D mention in their Discussion, one could place a prior on d. We suggest exploring this avenue in real terms, because recent results suggest that substantial improvements can be obtained by placing, say, a beta prior on d; see for instance Scott and Berger (2010) and Castillo and van der Vaart (2012).

Priors on matrices The Hyper Inverse Wishart (HIW) prior on Ψ, or Σ in the Gaussian case, requires the hyperparameters δ and Φ. F&D choose $\delta = 1$ and $\Phi = cI_p$, referring to Armstrong et al. (2009) for alternative choices of Φ. A related option would be using the Fractional Bayes Factor (FBF) to implement model choice based on objective improper priors: a fraction of the likelihood would be used to make the prior proper, then its complementary fraction would be used for inference (avoiding double use of data); see O'Hagan and Forster (2004, ch. 7).

In the Gaussian case the FBF turns a default improper HIW prior on Σ into a proper HIW prior with Φ proportional to the sample covariance matrix (Carvalho and Scott 2009) and this results in a markedly improved performance with respect to the standard choice $\Phi = cI_p$ (applied to the whole likelihood). The problem is that vague priors assign too much probability to parameter values not supported by the data, and this alters the evidence conveyed by the marginal likelihoods. We note that Consonni and La Rocca (2012) extend the results of Carvalho and Scott (2009) to the larger class of Directed Acyclic Graphs (DAGs).

Implementing the FBF in the setup of F&D should be feasible, because, conditionally on τ, the results for the Gaussian case extend in a straightforward way. Since the

[1]Dipartimento di Scienze Statistiche, Università Cattolica del Sacro Cuore, Milano, Italy, guido.consonni@unicatt.it

[2]Dipartimento di Scienze Fisiche, Informatiche e Matematiche, Università di Modena e Reggio Emilia, Modena, Italy, luca.larocca@unimore.it

 DOI:10.1214/13-BA856D