Bayesian Analysis (2014) **9**, Number 3, pp. 591–596

Rejoinder

Michael Finegold * and Mathias Drton †

In a nutshell, our paper treats the problem of how to extract information from partially corrupted observations in multivariate statistical problems—specifically, we focused on the problem of graphical model selection. Our approach considers different versions of multivariate distributions with t-marginals that are obtained by using, to varying extent, distinct Gamma-divisors for the different coordinates of a random vector. As described in the contributions to this discussion, there are a lot of ways one could modify or generalize the models we used, and related ideas have appeared or might be useful in contexts other than graphical modeling. There is also a vast literature on the general themes our paper touches upon, and the discussants have provided many additional references giving a far more comprehensive description of the existing literature than our own paper.

In this rejoinder we try to summarize and comment on the ideas we see emerge from the discussion.

Directions of multivariate tails. Figure 2 in our paper contrasts different versions of t-distributions in a bivariate setting. The figure shows a 'spherical' case, that is, the dispersion matrix $\mathbf{\Psi}$ is the identity. **Anthony O'Hagan**'s Figure 1 shows a case where $\mathbf{\Psi}$ exhibits strong correlation. His figure illustrates nicely that the alternative t-distribution has heavy tails along the coordinate directions, which also underlies the bounds on correlations we mention in our Section 4.2. O'Hagan's Figure 2 shows an example of a different type of t-distribution with heavy tails following principal component directions. We primarily thought of applications in which the latent dependence pattern captured by $\mathbf{\Psi}$ and the pattern of outliers are not tied together, the former being of say biological nature and the latter a matter of experimental technology. In this case focusing on the coordinate directions seems natural to us, but principal component directions or another coordinate system could be of interest in other applications.

Inference on the degrees of freedom. Our numerical work was based on the default choice of $\nu = 3$ degrees of freedom. **Abel Rodriguez** raised the point of inference on the degrees of freedom. Sticking with the precise setup of our paper, we expect that the message regarding the relative merits of the different t-distributions remains the same under inference on ν. However, one would certainly be able to decrease the gap that is visible in the ROC curve for normal data in Figure 3 of our paper. The work of Besag and Higdon (1999) is one example of Bayesian inference on the degrees of freedom. The two used a finite set of values for ν and suggested marginalization of the Gamma-divisors for a block Gibbs step, which poses no problems in their setup of independence ($\mathbf{\Psi}$ diagonal, in our setting). For non-diagonal $\mathbf{\Psi}$, their blocking strategy might prove useful for what we called the classical t-distribution but it seems more

*Department of Statistics, Carnegie Mellon University, Pittsburgh, U.S.A. mfinegol@andrew.cmu.edu

†Department of Statistics, University of Washington, Seattle, U.S.A. md5@uw.edu

 DOI:10.1214/14-BA856