Additional simulation results with heavy-tailed innovations

In this section, we report on the simulation results with heavy-tailed innovations. That is, t-distribution with degrees of freedom 5 is used instead of standard Normal innovations to generate an autoregressive model of order 1 with parameter .5 in our simulation study. For fair comparison with the standard Normal innovations, we also standardized $t(5)$ distribution so that it has a unit variance. Model parameters are the same as in our simulation study. That is, DGP2 has no connectivity changes, so it is used to calculate empirical size and DGP4 has connectivity changes at $t = 150$ and $t = 300$. We do not observe any significant differences with the standard Normal case as can be seen from the following figures that should be compared to Figures 5, 7, 8 and 9 in the paper.

DGP2 with $t(5)$ innovations

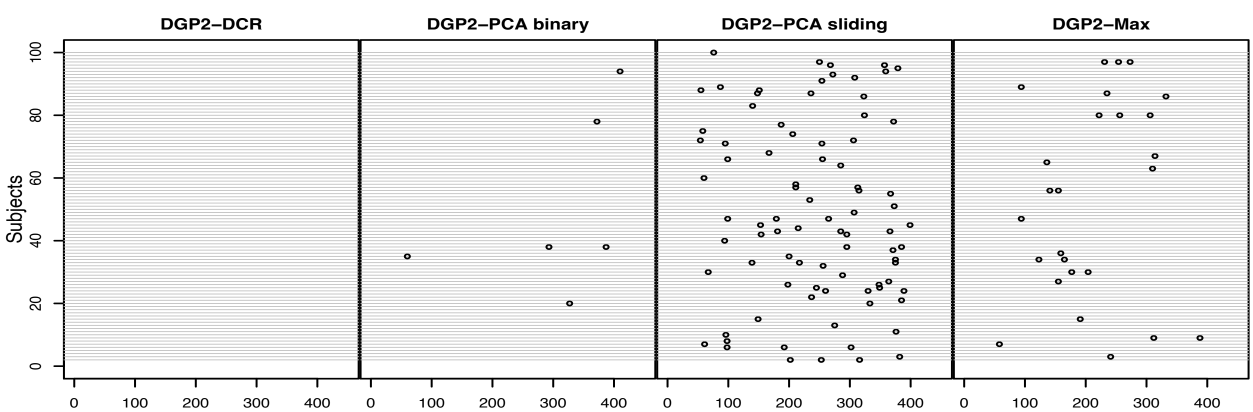

Figure 1.

The estimated change points in DGP2 with 20 nodes. No change point in DGP2 and a band matrix is used for covariance structure.

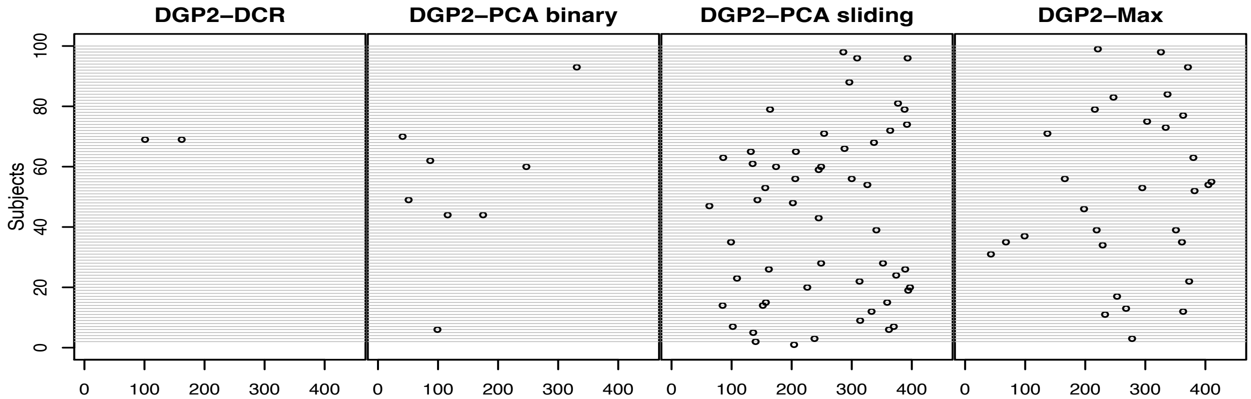

FIGURE 2.

The estimated change points in DGP2 with 250 nodes. DGP2 has no connectivity changes.

DGP4 with $t(5)$ innovations

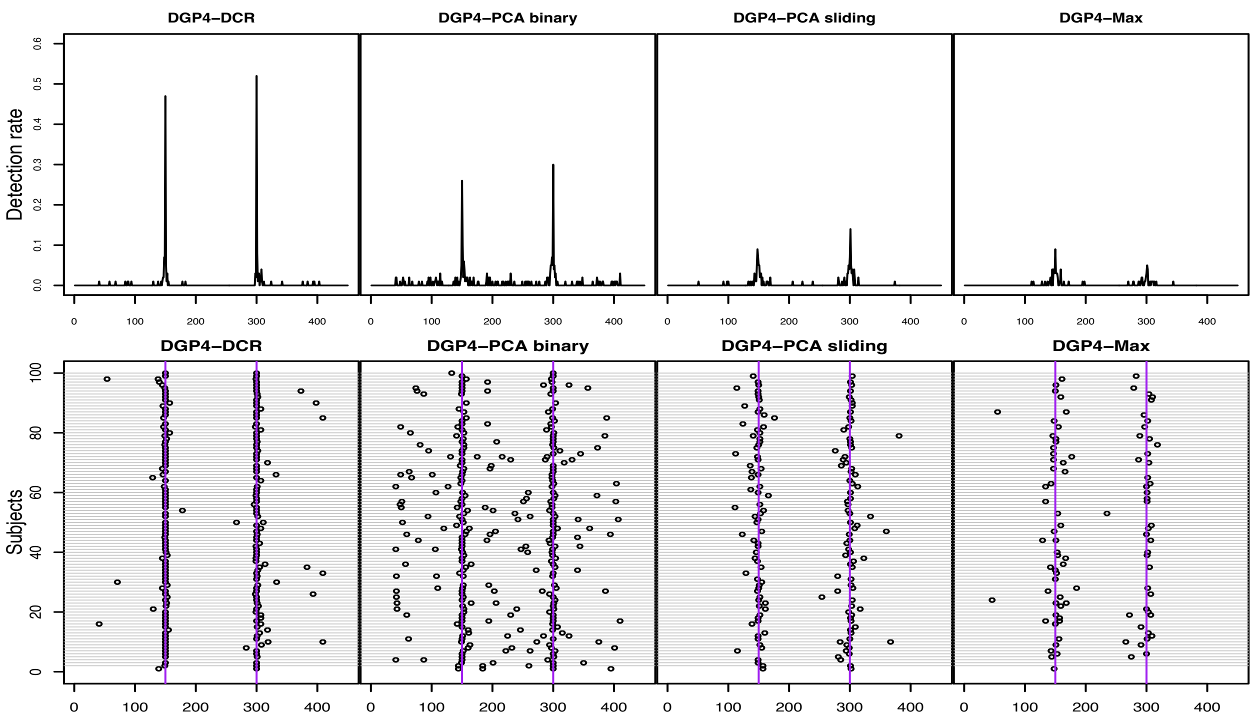

FIGURE 3.

The detection rates (top) and the estimated change points in DGP4 with 20 nodes. Connectivity changes at 150 and 300.

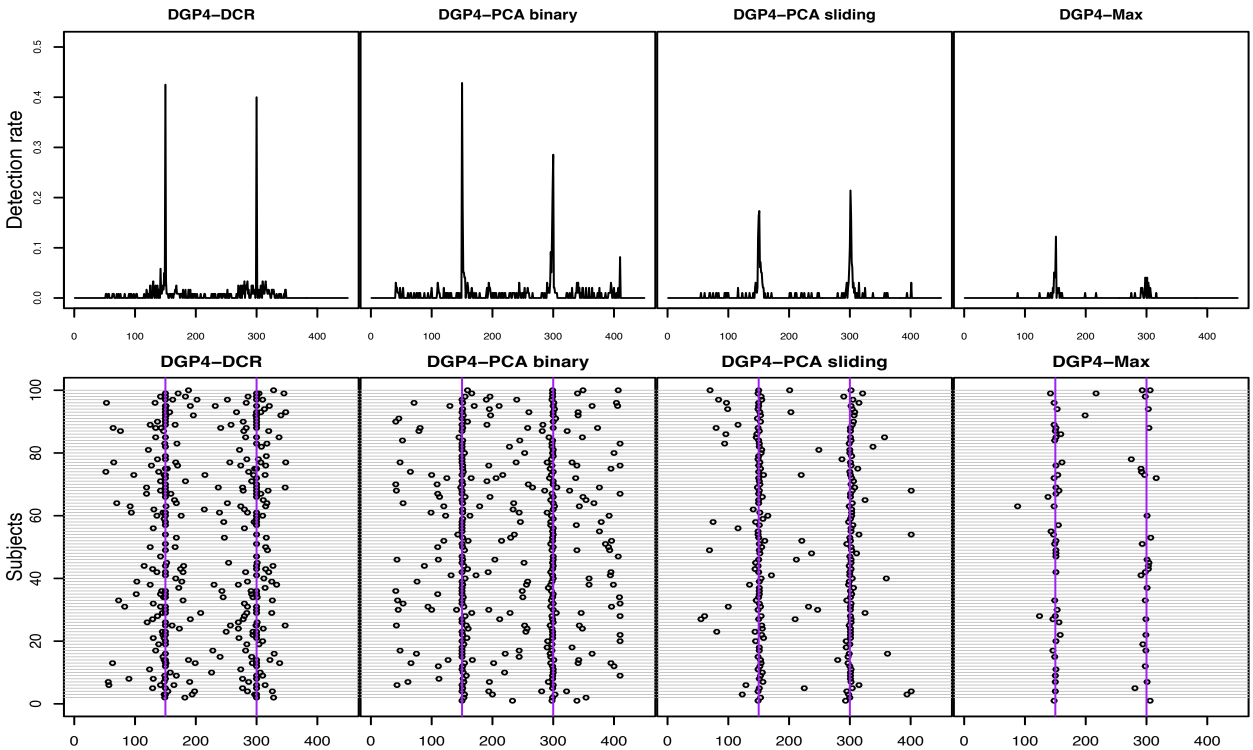

FIGURE 4.
The detection rates (top) and the estimated change points in DGP4 with 250 nodes. Connectivity changes at time points 150 and 300.

Modifications of DCR method

In this section, we detail our modifications of DCR method of Cribben et al. (2012) used in the paper. Two modifications were made in `detectR`: one in calculating BIC reduction and the other in finding change point. Note that Cribben et al. (2012) writes BIC based on the sample $\{X_s, X_{s+1}, \ldots, X_e\}$ as

$$\mathrm{BIC}(s:e) = \mathrm{tr}((e-s+1)\widehat{\Sigma}_{s,e}\widehat{\Omega}_{s,e}) - (e-s+1)\log|\widehat{\Omega}_{s,e}| + k_{s,e}\cdot\log(e-s+1) =: \mathrm{NLL}(s:e) + \mathrm{Pen}(s:e), \quad (1)$$

where s is the starting time and e is the end time, $\widehat{\Sigma}_{s,e} = (e-s+1)^{-1}\sum_{t=s}^{e} X_t X_t'$ and $\widehat{\Omega}_{s,e}$ is the graphical lasso estimator of the precision matrix based on the sample $\{X_s, X_{s+1}, \ldots, X_e\}$ with $k_{s,e}$ non-zero entries. BIC is the sum of the negative log likelihood and penalty terms, denoted $\mathrm{NLL}(s:e)$ and $\mathrm{Pen}(s:e)$ in (1). Suppose that we are considering one change point.

1. BIC reduction at point t_0:

   - (DCR) DCR calculates $\mathrm{BIC}(1:T) - \big(\mathrm{BIC}(1:t_0) + \mathrm{BIC}(t_0+1:T)\big)$. Observe that it is equal to the sum of the negative log likelihood reduction $\mathrm{NLL}(1:T) - \mathrm{NLL}(1:t_0) - \mathrm{NLL}(t_0+1:T)$, and the reduction of penalty term $\mathrm{Pen}(1:T) - \mathrm{Pen}(1:t_0) - \mathrm{Pen}(t_0+1:T)$, namely $k_{1,T}\cdot\log T - k_{1,t_0}\cdot\log t_0 - k_{t_0+1,T}\cdot\log(T-t_0)$.
   - (Modification) The `detectR` uses the same negative likelihood reduction, but the reduction of penalty term becomes $k_{1,T}\cdot\log T - (k_{1,t_0} + k_{t_0+1,T})\cdot\log T$. We modified the penalty term to reflect the fact that the same sample of size T is used in calculating BIC even with one change point at t_0.

2. Change point estimator:

   - (DCR) The point where the largest BIC reduction happens is the change point. It can be written as
   $$\underset{\Delta<t_0<T-\Delta}{\operatorname{argmin}} \{\mathrm{BIC}(1:T) - \big(\mathrm{BIC}(1:t_0) + \mathrm{BIC}(t_0+1:T)\big)\}.$$
   - (Modification) The `detectR` finds the change point without the penalty terms as
   $$\underset{\Delta<t_0<T-\Delta}{\operatorname{argmin}} \{\mathrm{NLL}(1:T) - \big(\mathrm{NLL}(1:t_0) + \mathrm{NLL}(t_0+1:T)\big)\},$$
   since the penalty terms introduce bias in finding change point.

References

Cribben, I., Haraldsdottir, R., Atlas, L. Y., Wager, T. D. and Lindquist, M. A. (2012), 'Dynamic connectivity regression: determining state-related changes in brain connectivity', *Neuroimage* **61**(4), 907–920.