Report on “The locally Gaussian partial correlation”

Generally speaking, I am quite satisfied with the revision that the authors have made and I appreciate the efforts very much. I also agree that the current version stress the main contribution as the introduction of the new concept of local partial Gaussian correlation (LGPC) rather than proposing a new conditional independence test (although the test performs very competitively).

After a careful reading of everything including the long appendix (I apologize for being slow), I still found the current revision contains some parts that may require the authors to make some straightforward modifications, remarks and refinements. But these are mostly cosmetic issues in nature. In particular, I would like to suggest that the authors further change/clarify the following minor issues.

Comment 1: It seems to me that it is better to replace v in Equation (4) by x instead. Otherwise, when f_x itself is a Gaussian density, it is hard to simply reduce $\psi(x, v)$ to $f_x(x)$. This is also in line with the expression in Equation (6).

Comment 2: The authors need to change $\hat{Z}_j$ in page 6 to be consistent with the one in Equation (12) in page 11, because the estimated version can only be defined as $\Phi^{-1}(\hat{F}_j(X_{jt}))$ for each component $j = 1, \ldots, p$. Also please make the notation consistent, using either $X_{j,t}$ or X_{jt}. There are other notation issues across the main text and the supplemental material, with some listed below.

Comment 3: Make notations like $f_{z_1|z_3}$ and $f_{Z_1|Z_3}$ consistent in page 9. In Equation (10), the definition of $\mu_{1|3}$ depends on both z_1 and z_3, so it is better defined as $\mu_{1|3}(z_1, z_3)$. Similarly, $\sigma_{1|3}(z_1, z_3)$ should be used and $\Sigma_{Z_1|Z_3}$ should be replaced by $\Sigma_{Z_1|Z_3}(z_1, z_3)$ in Equation (10). In addition, I do not see the product of $f_{z_1|z_3}$ and $f_{z_2|z_3}$ (with their expressions given by Equation (10) and alike) leads directly to the expression below Equation (10). Pls double-check.

Comment 4: Make Equation (13) consistent with Equation (7).

Comment 5: Is it beneficial to exploit the symmetric structure in Equation (17) in Section 3.1.2, say considering only the cases of $j < k$?

Comment 6: I’m afraid that in page 13 replacing Equation (18) by Equation (A.13) (defined in page 42) helps understand better the optimization problem for the pairwise case than described in Equation (18).

Comment 7: I would suggest the authors to unify the notations in Equation (19) and below, because $u(\cdot, \rho_b)$ and $u(\cdot, R_b)$ (or $\psi(\cdot, \rho_b)$ and $\psi(\cdot, R_b)$) are used as equivalent expressions. Although I understand that they could serve the purpose, $u(\cdot, R_b)$ and $\psi(\cdot, R_b)$ should be reserved for the case of general dimensions.

Comment 8: Make Equation (20) consistent with Equations (21) and (22); i.e, deciding to add the argument z or not.

Comment 9: The authors stress one desirable property of LGPC being able to distinguish between positive and negative conditional dependence. I cannot find too much discussion in Section 4 about this property, which could be viewed as a nice empirical device in contrast to the conventional Granger causality testing procedure that fails to say anything about the potentially asymmetric pattern of the underlying problem.

Comment 10: In Section 5.3, what is the exact form of function $h(x)$ as given in Equation (28), say the standard $h(x) = x^2$?

Comment 11: The online supplement contains a detailed analysis and is very useful. I appreciate your efforts very much. Equation (A.14) is defined for the minimization problem, but Equation (A.13) is define for the maximization problem. Pls change Equation (A.14) accordingly as in the subsequent theoretical analysis including the assumptions it is about maximization problem.

Comment 12: In the second parts of both Theorem F.2 and Theorem F.3, they do not involve any bias terms that typically diverge to infinity as $n \to \infty$ and $b \to 0$; i.e., T_b can be simply replaced by T under the condition of $n \to \infty$ and $b \to 0$. This may seem to be a bit surprising to me given the popular L_2 form of the proposed test statistic in Equation (28). Is it because of the particular bandwidth condition imposed in Assumption 3 in Theorem 3.1? Is it somewhat undersmoothing to take care of the potential bias terms?