Report on “The locally Gaussian partial correlation”

The authors propose a new conditional dependence measure (called the local Gaussian partial correlation coefficient, LGPC). They then use it to construct an alternative test for conditional independence and test Granger causality in time series. I enjoy reading the paper and only have a few specific comments/suggestions that matter for the exposition of the paper. And I recommend a revision.

In the following, I list the expositional issues that I would like to point out and discuss with the authors.

Comment 1: The whole paper discussed the LGPC coefficient as the local version of the partial correlation coefficient between X_1 and X_2 given $\boldsymbol{X}^{(2)}$, with $\boldsymbol{X}^{(1)} = (X_1, X_2)$ and $\boldsymbol{X}^{(2)} = (X_3, \ldots, X_p)$. While it is easy to understand, it seems that the discussed dependence measure is only limited to bivariate case conditional on a third vector. Many of the existing conditional independence tests mentioned in Section 4 are, however, designed to work for general dimensions; for example, when testing $\boldsymbol{X} \perp \boldsymbol{Y}|\boldsymbol{Z}$ with $(\boldsymbol{X}, \boldsymbol{Y}, \boldsymbol{Z}) \in \mathbb{R}^{d_x+d_y+d_z}$. Of course, “curse of dimensionality” would always be a serious issue depending on the specific testing methodology adopted; among them, the empirical process based tests would be less affected by the dimension of $\boldsymbol{Z}$ in theory. Perhaps more importantly, the aforementioned conditional independence tests do not place restrictions on the dimensions of $\boldsymbol{X}$ and $\boldsymbol{Y}$, and theoretically speaking, they should work for very general settings in which both components X_1 and X_2 in $\boldsymbol{X}^{(1)}$ can be multivariate. How can the current LGPC measure adapt to the multivariate X_1 and X_2 case? Could you provide some insights or remarks towards this direction?

Comment 2: In Equation (6), isn’t it an approximation of the density $f_{\boldsymbol{Z}}$ of $\boldsymbol{Z}$ in the neighborhood of $\boldsymbol{z}$? The way it writes now seems to indicate that the expression is exactly the density of $f_{\boldsymbol{Z}}$ of $\boldsymbol{Z}$, but $f_{\boldsymbol{Z}}(\boldsymbol{z})$ can certainly differ from $\psi(\boldsymbol{z}, R(\boldsymbol{z}))$ in the general case.

Comment 3: In the definition of $F_{j,n}(x)$ of page 5, X_i should be X_{ji}. Also the empirical analogue for each Z_j can only be written as $Z_{ji,n} = \Phi^{-1}(F_{j,n}(X_{ji}))$ for $j = 1, \ldots, p$ and $i = 1, \ldots, n$.

Comment 4: Above Equation (10), for the conditional density, use notation $f_{Z_1|\boldsymbol{Z}_3}(z_1|\boldsymbol{z}_3)$

to be consistent with subsequent notations.

Comment 5: In Equation (15), change $\hat{\boldsymbol{z}}$ to $\boldsymbol{z}$. And isn't the definition of kernel function $K_{\boldsymbol{b}}(\boldsymbol{x}) = |\boldsymbol{b}|^{-1}K(\boldsymbol{b}^{-1}\boldsymbol{x})$?

Comment 6: In Equation (17), change $\hat{\boldsymbol{z}}$ to $\boldsymbol{z}$, $(\hat{z}_j, \hat{z}_k)$ to (z_j, z_k), and there is a ψ missing in log.

Comment 7: In Equation (19), I suspect the expression should be $\boldsymbol{M}_{\boldsymbol{b}}^{-1/2}\boldsymbol{J}_{\boldsymbol{b}}$ instead of $\boldsymbol{J}_{\boldsymbol{b}}\boldsymbol{M}_{\boldsymbol{b}}^{-1/2}$. Pls double check. Although the the exact expression of the asymptotic covariance matrix $\boldsymbol{J}_{\boldsymbol{b}}^{-1}\boldsymbol{M}_{\boldsymbol{b}}\boldsymbol{J}_{\boldsymbol{b}}^{-1}$ is not available, could you state the orders of $\boldsymbol{M}_{\boldsymbol{b}}$ and $\boldsymbol{J}_{\boldsymbol{b}}$? It is hard to understand the convergence rate without knowing the respective orders of $\boldsymbol{M}_{\boldsymbol{b}}$ and $\boldsymbol{J}_{\boldsymbol{b}}$.

Comment 8: In page 18, line 10 from below, starting from "It is straightforward to see that". The sentence is very confusing here as it holds only when $\rho = 0$ in the context of Figure 3.

Comment 9: In page 25, line 13 from below, starting from "Departures from many types of conditional independence". This sentence is also confusing as conditional independence is a restriction to the underlying populations under the null hypothesis. Are there many other different types of conditional independence?

Comment 10: In Table 1, correct the typos in DGPs 3 (change X_{t-1}^2) and 10 (change Y_t^2).

Comment 11: In Table 2, what are the tests SCM and SKS? I couldn't find these definitions anywhere.

Comment 12: In the proof of Theorem 3.2, to show that

$$\frac{1}{n}\sum_{i=1}^{n}(F_n(X_i) - F(X_i)) \to_p 0,$$

a conventional U-statistic argument (e.g. Hoeffding decomposition) for dependent data (e.g.

strong mixing) is perhaps much easier and preferable, by noting that

$$
\begin{aligned}
&\frac{1}{n}\sum_{i=1}^{n}(F_n(X_i)-F(X_i)) \\
=&\frac{1}{n^2}\sum_{i=1}^{n}\sum_{j\neq i}^{n}\{I(X_j\leq X_i)-F(X_i)\}+\frac{1}{n^2}\sum_{i=1}^{n}\{1-F(X_i)\} \\
=&\frac{1}{n}\sum_{i=1}^{n}\{0.5-F(X_i)\}+O_p(n^{-1}) \\
=&O_p(n^{-1/2})+O_p(n^{-1}) \\
=&O_p(n^{-1/2}).
\end{aligned}
$$

This also implies that the order of $\frac{1}{n}\sum_{i=1}^{n}(F_n(X_i)-F(X_i))$ is sharper than the obtained order $n^{-1/2}(2\log\log n)^{1/2}$.