We thank the reviewers for their thorough and inspiring comments. The overall feedback is positive, with the main suggestions for improvement being i) an application to real data, ii) further tests and discussions (non-time series case, high cardinality, discrete variables), and iii) a comparison to residualization approaches. We will follow these suggestions by including further experiments and discussions in the camera-ready version. Below, our brief answers.

Real data: We agree that, despite the difficulty of basing the evaluation of causal discovery methods on real data, a real data example would be an asset to the paper. We will therefore include an application of LPCMCI to a river discharge dataset. A first analysis shows encouraging results (given our understanding of the causal mechanisms). **Non-time series case:** The idea of increasing effect sizes by default conditioning on parents in principle also applies to the non-time series case. We speculate, however, that the gain is most significant in the presence of strong autocorrelation. Moreover, in the non-time series case LPCMCI finds only few default conditions because i) PAGs tend to be more unoriented and ii) parentships can only be found after having oriented some colliders first (to find parents one first needs some heads '>', which come for free in the time series case). We already cover the non-autocorrelated case that still has time order, see the $a = 0$ point in Fig. 2B as well as all plots for $a = 0$ in the Supplement, which shows comparable performance of SVAR-FCI and LPCMCI. We additionally ran experiments in the true non-time series case and got similar results. **Higher cardinality:** As we state, there is a tradeoff between the positive effect of conditioning on parents and the negative effect of higher cardinality. For the setting that LPCMCI is designed for, autocorrelated time series, our experiments show a significant performance gain (excluding the discrete case for now). In other settings, e.g. the non-time series case, the relative effect is less clear. LPCMCI's conditioning sets consist of two parts: The standard PC-like set plus the default conditions (known parents). The cardinality constraint mentioned in L275-277 i) only restricts the former part, ii) is used only in the last phase of LPCMCI (pseudocode line 6), iii) applies to SVAR-FCI too, and iv) is used to limit excessive runtime (mostly needed for SVAR-FCI). In the continuous case, loosing $\mathcal{O}(1)$ degrees of freedom by default conditions is negligible to, e.g., $\mathcal{O}(100)$ sample sizes. While we did not implement a constraint on the number of default conditions, this would indeed be a good idea as it would allow to analyze the effect of higher cardinality and might be relevant for the discrete case. **Discrete variables:** Fair point. LPCMCI in principle also works with discrete variables as it can utilize any CI test, but evaluation is needed. While preliminary experiments did not show significant differences between the methods, we will run more experiments and accordingly extend the camera-ready version. The range of applicability of LPCMCI will remain broad in any case. In climate science applications, e.g., there usually are only few discrete variables, if any. **Residualization:** The question is whether instead of conditioning on parents one might use a residualization procedure in data preprocessing. We ran two tests. 1) Fit independent AR(1) models and run SVAR-FCI on the residuals. 2) Instead of AR(1) use GP regression as proposed in Flaxman et al., 2016 (using sklearn with RBF kernel and $\alpha = 1$). In both cases adjacency TPR and orientation recall increase but are still lower than for LPCMCI, whereas adjacency FPR increases and orientation precision drops. Among the two, AR(1) performed better. Generally, we are not sure what the ground truth MAG / PAG should be after residualization. Perhaps they should not contain auto-links. This seems to require a substantially different theory.

Other questions and comments will all be addressed by further explanations in the camera-ready version, here our brief answers. **Do we compare to genuine FCI or SVAR-FCI?** To SVAR-FCI, as stated in L103f. **Relation of Theorem 1 to higher recall:** For a single CI test with null $I(X,Y|Z) = 0$ and alternative $I(X,Y|Z) > 0$ the effect size is the value of $I(X,Y|Z)$ in the true (unknown) distribution. For X and Y adjacent, effect size $I(X,Y|Z) > 0$. The larger this true value, the higher the probability of its sample value lying in the test's rejection region and hence of correctly retaining the edge (thus higher recall). Recall is influenced both by effect size and by the cardinality of Z, with the details depending on the particular test statistic. **How are parents determined?** LPCMCI alternates between performing CI tests and applying orientation rules, the latter of which may identify some parentships that are then used as default conditions in the next iteration of CI tests. See also L230-236. **Relation to PCMCI:** Our work borrows, and by means of Theorem 1 formalizes, PCMCI's intuition that effect size increases by default conditioning on parents. PCMCI does use default conditions, but it tries to limit their number. In the causally insufficient setting of LPCMCI, bidirected edges can point into the past. To ensure that no m-separations are destroyed, all default conditions must be ancestors of X or Y (though only parents are used to not make cardinality unnecessarily large). This requires orienting edges before having found a final skeleton, which in turn requires our new graphical theory in Secs. 3.2 and 3.3. **Not assuming orientation-faithfulness:** LPCMCI orients colliders not with the potentially overly restrictive 'conservative rule' but with a variant of the 'majority rule' (Colombo and Maathuis, 2014). It also marks conflicts when contradicting orientations are proposed. We assume full faithfulness to prove soundness and do not attempt to discover violations of orientation-faithfulness (we are not aware of such work in the causally insufficient case). **Use of known parents in Lee and Honavar 2017:** We will add a citation. **Small drop of precision drop from $T = 500$ to $T = 1000$:** We found a consistent slight decrease in precision only for contemporaneous links and strong autocorrelation, whereas for lagged links precision sometimes even slightly increases (see Fig. 12 bottom right). Since cardinality increases for both type of links, we do not see an easy explanation. **'Stationarity is enforced':** Whenever an edge is removed (oriented), all equivalent time shifted edges are removed too (oriented in the same way). **Taking into account background knowledge about parentship:** Yes, exactly! We plan to implement this in a future version of LPCMCI.