Supplementary information

Discovering cause-effect relationships in spatial systems with a known direction based on observational data

Konrad Mielke K.MIELKE@SCIENCE.RU.NL
Tom Claassen TOMC@CS.RU.NL
Mark A.J. Huijbregts M.HUIJBREGTS@SCIENCE.RU.NL
Aafke Schipper A.SCHIPPER@SCIENCE.RU.NL
Tom Heskes T.HESKES@SCIENCE.RU.NL
Faculty of Science, Radboud University, Postbus 9010, 6500 GL Nijmegen, The Netherlands

Proof of Theorem 2

We split the theorem into three parts and proof each part separately. We presume that the fast causal inference is sound (Spirtes et al. (2000)).

Lemma 1 *The inferred graph over variables $\boldsymbol{I}$ in which each variable i is replaced by its partner variable u is the same as the inferred graph over variables $\boldsymbol{U}$.*

Proof Suppose that the inferred graph was not the same. That means that at least one conditional independence test, w.l.o.g. $i_1 \perp\!\!\!\perp i_2 | i_3$, gave a different test result than the corresponding conditional independence test, $u_1 \perp\!\!\!\perp u_2 | u_3$. That, however, is against our assumption of spatial invariance which dictates that the causal structure of the system does not depend on the location. ■

Lemma 2 *Removing edges between variables $\boldsymbol{U}$ does not introduce erroneous orientations in the orientation phase.*

Proof Assume that removing the edge between vertices $X \in \boldsymbol{U}$ and $Y \in \boldsymbol{U}$ introduces an unshielded triple between X, $Z \in \boldsymbol{I}$ and Y. But then, we would have oriented the edges $X - Z$ and $Y - Z$ into Z already before the orientation phase which means that the collider rule would not trigger at all.

Next, assume that removing the edge between $X \in \boldsymbol{U}$ and $Y \in \boldsymbol{U}$ introduces a discriminating path between X and Y for Z. That would mean that there exist edges $Z_k \rightarrow Y$ between variables on the path Z_k and Y. However, that can never happen because causal links from variables in $\boldsymbol{I}$ to variables in $\boldsymbol{U}$ are not allowed. ■

Lemma 3 *Orienting edges from $\boldsymbol{U}$ and $\boldsymbol{O}$ into $\boldsymbol{I}$ and $\boldsymbol{R}$ before the orientation phase does not lead to the propagation of erroneous orientations in the orientation phase.*

Proof Here, we only sketch the proof for the orientation of colliders and the first orientation rule. W.l.o.g., we focus on variables from $\boldsymbol{O}$ and $\boldsymbol{I}$. Let X, Z and Y be an unshielded triple. Then either, the collider has already been oriented due to background knowledge which we presume to be correct, or only one arrowhead has been inserted yet. But then, the collider orientation is not affected and thus no wrong orientations are introduced.

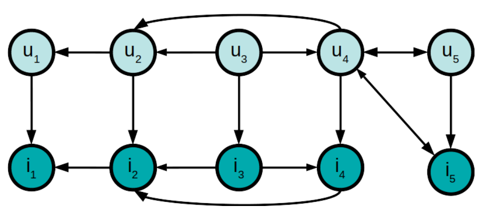
Next, consider another triple X, Z and Y of which the edge from X into Z has been oriented as an arrowhead due to background knowledge, but Z is not a collider. Then rule 1 orients the edge $Z \rightarrow Y$. Assume that this was not the orientation in the true underlying graph. Then either, Z would have to be a collider, which we already ruled out. Or our background knowledge was faulty, which we presume cannot happen.

... ■

Now, we can put everything together.

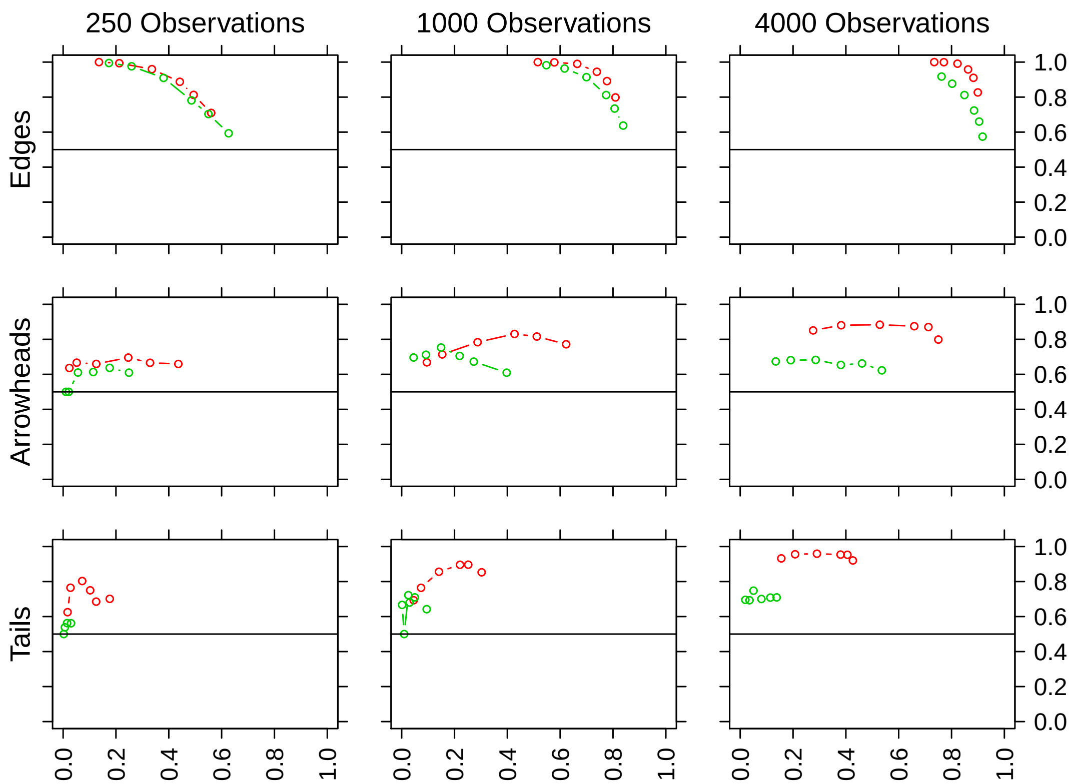
Proof Together, Lemma 1, Lemma 2 and Lemma 3 proof that not erroneous edges or orientations are introduced by the currentFCI algorithm. ■

Figure SI 1



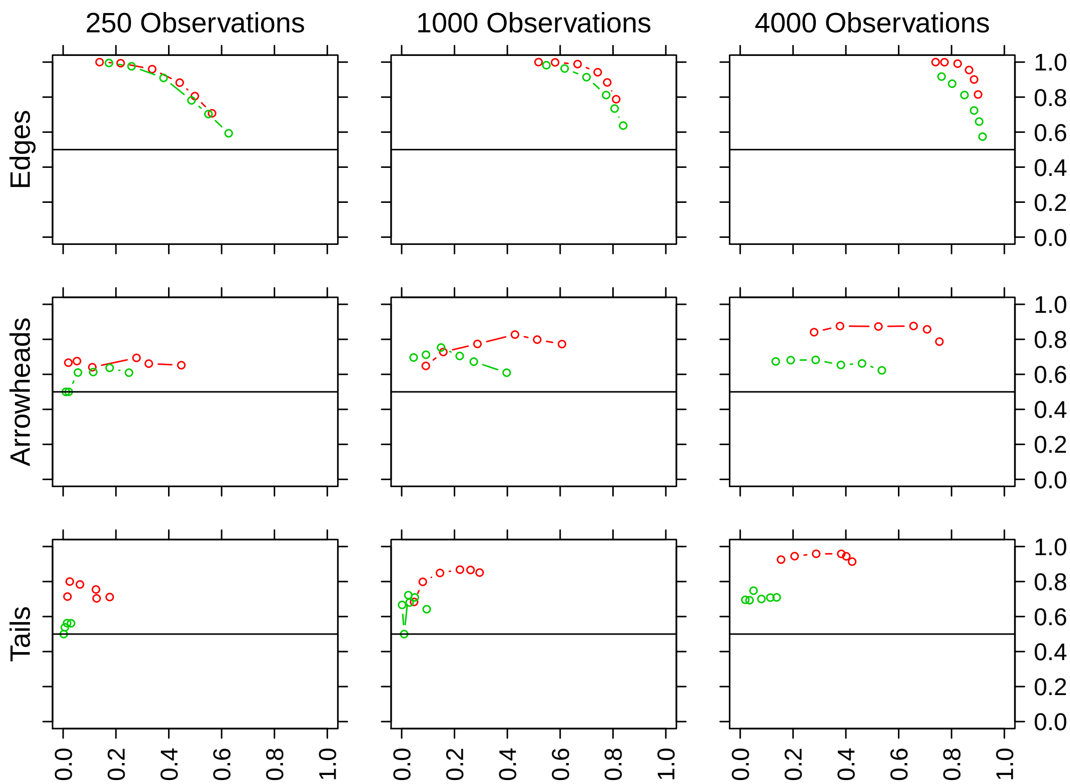
Schematic graph describing a system with $I = 5$ variables. Here, we explain why it is suboptimal to cut all edges between vertices of $\boldsymbol{U}$ or to keep all edges between vertices of $\boldsymbol{U}$. Assume that we cut all edges between vertices of $\boldsymbol{U}$. To separate the vertices u_2 and i_5, we need to condition on u_3 and u_4. In the skeleton phase, we will never perform a conditional independence test of u_2 and i_5 given u_3 and u_4, because at the time we would do the test, u_3 is neither the neighbor of u_2 nor of i_5. This means that we can only find the independence in the possible d-separation phase. Assume on the other hand that we keep all edges between vertices of $\boldsymbol{U}$. Then we will do unnecessary tests for the independence of u_4 and i_5, conditioning on sets that are not the neighbor of either of the two in the true underlying graph. Both strategies lead to superfluous conditional independence tests.

Figure SI 2



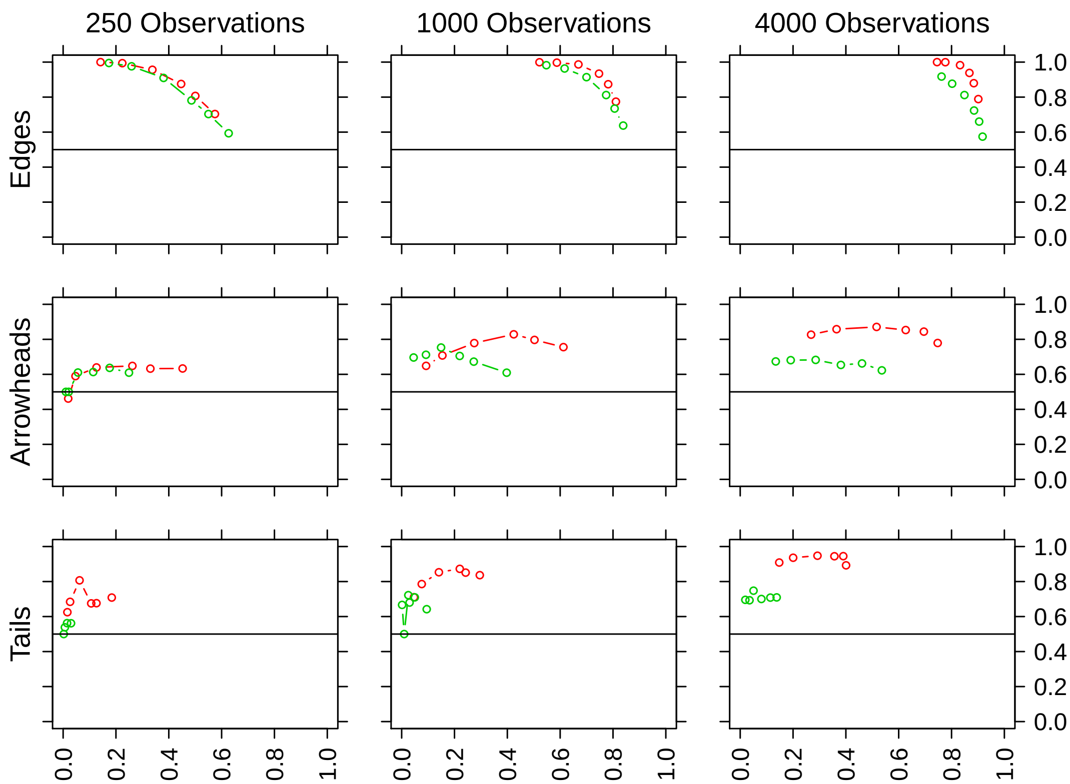
Results of the simulation study with additional noise with variance $\sigma^2 = 0.1$. Recall (x-axis) against precision (y-axis) obtained for the standard FCI (green) and the currentFCI (red) for different numbers of observations (columns) and edges (top) and edge marks (center and bottom), respectively.

Figure SI 3



Results of the simulation study with additional noise with variance $\sigma^2 = 0.3$. Recall (x-axis) against precision (y-axis) obtained for the standard FCI (green) and the currentFCI (red) for different numbers of observations (columns) and edges (top) and edge marks (center and bottom), respectively.

Figure SI 4



Results of the simulation study with additional noise with variance $\sigma^2 = 0.5$. Recall (x-axis) against precision (y-axis) obtained for the standard FCI (green) and the currentFCI (red) for different numbers of observations (columns) and edges (top) and edge marks (center and bottom), respectively.

References

P. Spirtes, C. N. Glymour, R. Scheines, and D. Heckerman. *Causation, prediction, and search*. MIT press, 2000.