Erratum to “Stability of causal inference”

Leonard J. Schulman* Piyush Srivastava†

In this erratum, we fix an error in the proof of Theorem 1.2 in the above paper from UAI 2016. The theorem presents an example on which the **ID** algorithm has a high condition number. With the fix, the theorem stands as stated in the paper. However, we remark that a comment about a feature of the constructed example with high condition number, made towards the beginning of Section 2, is incorrect after the proposed fix. We discuss this in more detail below.

Modification of the description of the hidden edges in the gadget

The paragraph describing the hidden edges (this is the paragraph beginning “Our final task...” in Section 2.1) needs to be modified as follows. The sentences beginning “In addition, we have further....” should read as below:

> We also have an unnamed hidden variable for each adjacent pair of the X_i, and a final unnamed hidden variable incident on X_2 and S_1. In addition, we have further (named) hidden variables $\{U_i \mid 2 \le i \le k\}$, such that U_i has as children the node X_i and all Y nodes at the “level” just below i. Formally, the hidden edges incident on these named hidden nodes are:
>
> $$\begin{aligned} H := &\{(U_i, X_i) \mid 2 \le i \le k\} \\ &\cup \left\{(U_i, Y_{s,i-1}) \mid 2 \le i \le k, 1 \le s \le n-1\right\}. \end{aligned}$$

The description of fig. 2 in the paragraph just below this will also have to change accordingly as follows. The sentences starting “Instead, the hyperedges. . . ” are modified as follows:

> Instead, the different shaded regions represent the C-components in $\mathcal{G}_6^4$ and its sub-graphs: the lowest region depicts the unique C-component in $\mathcal{G}_6^2$, the next higher region depicts the unique C-component in $\mathcal{G}_6^3$, and the topmost region depicts the unique C-component in $\mathcal{G}_6^4$.

The rest of the proof remains unchanged.

Modification of comment regarding the features of the ill-conditioned example

Near the beginning of Section 2, the strategy of the proof of the Theorem 1.2 is described. In particular, the following sentence appears (emphasis not in the original).

> The crux of our proof is a construction of two probability distributions: the first of these, Q, will be a distribution on the states of the nodes in $U \cup V$ <u>which respects $\mathcal{G}_n^k$</u>. The second, $\tilde{Q}$, will be a distribution only on the states of V, such that it is ϵ-close to the marginal of Q on V.

With the modification above, the property of Q claimed in the underlined part is not true. The condition number lower bound of Theorem 1.2 is therefore obtained at an input point Q which is not on the manifold of probability distributions which respect $\mathcal{G}_n^k$. To contrast this with the condition number upper bound result, Proposition 1.3, of the paper, we note that the upper bound obtained there also does not require the input observed distribution P to respect

*California Institute of Technology. Email: schulman@cms.caltech.edu.
†Tata Institute of Fundamental Research, Mumbai. Email: piyush.srivastava@tifr.res.in.

the given graph: the **ID** formula for graphs in the graph family in Proposition 1.3 remains well-conditioned even for input observed distributions P that do not respect the underlying graph. Understanding the condition number of **ID** for inputs P that lie on the manifold of probability distributions respecting the given graph G remains an open problem.