Expressible sets and expressible assignments

RWS

I. EXPRESSIBLE SETS

Example 1 Incompatibility of Pienaar distribution with DAG #16

Consider the DAG of Fig. 1. Henson, Lal and Pusey showed that this DAG is a candidate for being 'interesting', that is, the compatible distributions satisfy constraints over and above the conditional independence relations that follow from d-separation relations in the DAG.

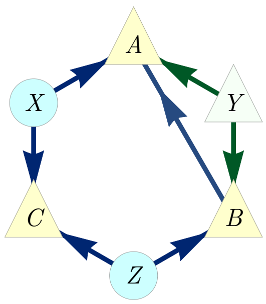

FIG. 1. DAG #16 in Ref. [**?**].

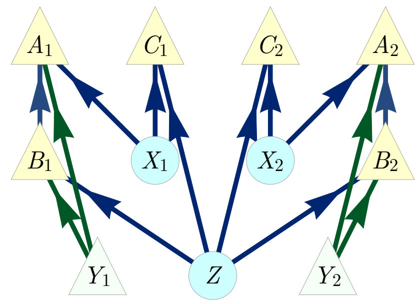

FIG. 2. The Rocket inflation of **??**.

?] identified a distribution which satisfies the CI relations among the observed variables in DAG #16, namely, $Y \perp\!\!\!\perp C$ and $A \perp\!\!\!\perp B|Y$ [**?**], but is nonetheless incompatible with it:

$$P^{\text{Pien}}_{ABCY} := \frac{[0000]+[0110]+[0001]+[1011]}{4}, \quad \text{i.e.,} \quad P^{\text{Pien}}_{YABC}(yabc) := \begin{cases} \frac{1}{4} & \text{if } y\cdot c = a \text{ and } (y\oplus 1)\cdot c = b, \\ 0 & \text{otherwise.} \end{cases} \tag{1}$$

In words, if $Y=0$, then B and C are in a maximally correlated state, while A takes the value 0, while if $Y=1$, then A and C are maximally correlated, while B takes the value 0.

Here, we will establish this incompatibility using the inflation technique. To do so, we use the Parachute inflation of the Modified Triangle scrnario, depicted in Fig. 2. The injectable sets include $\{A_1B_1C_1Y_1\}$ and $\{A_2B_1C_2Y_1\}$. They share the same image on the original DAG, namely, the set of all observed variables, $\{ABCY\}$. It follows that

$$P_{A_1B_1C_1Y_1} = P_{A_2B_1C_2Y_1} = P^{\text{Pien}}_{ABCY}. \tag{2}$$

If P^{Pien}_{ABCY} is compatible with the Modified Triangle Scenario, then by Lemma ?, $P_{A_1B_1C_1Y_1}$ and $P_{A_2B_1C_2Y_1}$ are compatible with the Parachute inflation of the Modified Triangle Scenario. To show that P^{Pien}_{ABCY} is *incompatible* with the Modified Triangle Scenario, we assume do so, we assume that $P_{A_1B_1C_1Y_1}$ and $P_{A_2B_1C_2Y_1}$ are compatible with the Parachute inflation of the Modified Triangle Scenario and derive a contradiction.

Note first that we can rewrite Eq. 1 as

$$P_{ABCY} = \frac{1}{2}([00]_{BC}+[11]_{BC})[0]_A[0]_Y + \frac{1}{2}([00]_{AC}+[11]_{AC})[0]_B[1]_Y \tag{3}$$

It follows that

$$P_{A_1C_1|Y_1=1} = \frac{1}{2}([00]_{A_1C_1}+[11]_{A_1C_1}), \tag{4}$$

$$P_{A_2C_2|Y_1=1} = \frac{1}{2}([00]_{A_2C_2}+[11]_{A_2C_2}). \tag{5}$$

and that

$$P_{B_1C_1|Y_1=0} = \frac{1}{2}([00]_{B_1C_1}+[11]_{B_1C_1}), \tag{6}$$

$$P_{B_1C_2|Y_1=0} = \frac{1}{2}([00]_{B_1C_2}+[11]_{B_1C_2}). \tag{7}$$

Any distribution $P_{B_1C_1C_2Y_1}$ which has marginals $P_{B_1C_1Y_1}$ and $P_{B_1C_2Y_1}$ that reproduce the conditional distributions of Eq. (6) and Eq. (7) respectively, must be such that

$$P_{B_1C_1C_2|Y_1=0} = \frac{1}{2}([000]_{B_1C_1C_2} + [111]_{B_1C_1C_2}). \tag{8}$$

The reason is that if B_1 and C_1 are perfectly correlated, as in Eq. (6), and B_1 and C_2 are perfectly correlated, as in Eq. (7), then C_1 and C_2 must be perfectly correlated as well. Indeed, marginalizing Eq. (8) over B_1, we obtain

$$P_{C_1C_2|Y_1=0} = \frac{1}{2}([00]_{C_1C_2} + [11]_{C_1C_2}). \tag{9}$$

But the Parachute inflation of the Modified Triangle Scenario is such that C_1C_2 and Y_1 are ancestrally independent, so that

$$P_{C_1C_2|Y_1} = P_{C_1C_2}, \tag{10}$$

and therefore $P_{C_1C_2|Y_1=0} = P_{C_1C_2|Y_1=1}$, so that we can infer from Eq. (9) that

$$P_{C_1C_2|Y_1=1} = \frac{1}{2}([00]_{C_1C_2} + [11]_{C_1C_2}). \tag{11}$$

Finally, we note that any distribution $P_{A_1A_2C_1C_2Y_1}$ which has marginals $P_{A_1C_1Y_1}$, $P_{A_2C_2Y_1}$, and $P_{C_1C_2Y_1}$ that reproduce the conditional distributions of Eq. (4), Eq. (5) and Eq. (11) respectively must be such that

$$P_{A_1A_2C_1C_2|Y_1=1} = \frac{1}{2}([0000]_{A_1A_2C_1C_2} + [1111]_{A_1A_2C_1C_2}). \tag{12}$$

The reason is that if A_1 and C_1 are perfectly correlated, as in Eq. (4), and A_2 and C_2 are perfectly correlated, as in Eq. (5), and C_1 and C_2 are perfectly correlated, as in Eq. (11), then A_1 and A_2 must be perfectly correlated as well.

Marginalizing Eq. (12) over C_1C_2, we obtain

$$P_{A_1A_2|Y_1=1} = \frac{1}{2}([00]_{A_1A_2} + [11]_{A_1A_2}). \tag{13}$$

Finally, we note that in the Parachute inflation of the Modified Triangle Scenario, A_1 is d-separated from A_2 given Y_1, which implies that $P_{A_1A_2|Y_1} = P_{A_1|Y_1}P_{A_2|Y_1}$. This is inconsistent with Eq. (13), so we have derived a contradiction.