Suppose that experiment E_1 showed that x_2 and x_1 were independent, such that the ordering between x_1 and x_2 remains underdetermined.

The only experiment that establishes whether $x_1 \rightarrow x_2$ are experiments E_2 with x_1 in J_2 and x_2 not in J_2.

Experiments E_1 and E_2 resolve the order between x_1 and x_2, suppose without loss of generality that it is $x_1 \rightarrow x_2$. In the worst case this required two experiments.

Now for the remainder: the only experiment that establishes whether $x_1 \rightarrow x_3$ are experiments E_3 with x_1 and x_2 in J_3 and x_3 not in J_3. Note that none of the previous experiments could have been an E_3.

The only experiment that establishes whether $x_1 \rightarrow x_4$ are experiments E_4 with x_1, x_2, x_3 in J_4 and x_4 not in J_4. None of the previous experiments could have been an E_4.

. . .

The only experiment that establishes whether $x_1 \rightarrow x_n$ is an experiment E_n with $x_1, \ldots, x_{n-1}$ in J_n and x_n not in J_n. None of the previous experiments could have been an E_n. It follows that n experiments are in the worst case necessary to discover the causal structure. QED

The above proof shows that in the worst case a sequence of n experiments is necessary that have intervention sets that intervene on at least $n - i$ variables simultaneously for each integer i in $1 < i < n$.

Appendix B

Parameterization PM1 for structure 1 in figure 2 (all variables are binary):

$p(u = 1) = .5$

$p(v = 1) = .5$

$p(x = 1|u = 1) = .8$
$p(x = 1|u = 0) = .2$

$p(y = 1|v = 1, x = 1) = .8$
$p(y = 1|v = 1, x = 0) = .8$
$p(y = 1|v = 0, x = 1) = .8$
$p(y = 1|v = 0, x = 0) = .2$

$p(z = 1|u = 1, v = 1, x = 1, y = 1) = .8$
$p(z = 1|u = 1, v = 1, x = 1, y = 0) = .8$
$p(z = 1|u = 1, v = 1, x = 0, y = 1) = .84$
$p(z = 1|u = 1, v = 1, x = 0, y = 0) = .8$
$p(z = 1|u = 1, v = 0, x = 1, y = 1) = .8$
$p(z = 1|u = 1, v = 0, x = 1, y = 0) = .8$
$p(z = 1|u = 1, v = 0, x = 0, y = 1) = .64$
$p(z = 1|u = 1, v = 0, x = 0, y = 0) = .8$
$p(z = 1|u = 0, v = 1, x = 1, y = 1) = .8$
$p(z = 1|u = 0, v = 1, x = 1, y = 0) = .8$
$p(z = 1|u = 0, v = 1, x = 0, y = 1) = .79$
$p(z = 1|u = 0, v = 1, x = 0, y = 0) = .8$
$p(z = 1|u = 0, v = 0, x = 1, y = 1) = .8$
$p(z = 1|u = 0, v = 0, x = 1, y = 0) = .2$
$p(z = 1|u = 0, v = 0, x = 0, y = 1) = .84$
$p(z = 1|u = 0, v = 0, x = 0, y = 0) = .2$

Parameterization PM2 for structure 2 in figure 2:

$p(u = 1) = .5$
$p(v = 1) = .5$

$p(x = 1|u = 1) = .8$
$p(x = 1|u = 0) = .2$

$p(y = 1|v = 1, x = 1) = .8$
$p(y = 1|v = 1, x = 0) = .8$
$p(y = 1|v = 0, x = 1) = .8$
$p(y = 1|v = 0, x = 0) = .2$

$p(z = 1|u = 1, v = 1, y = 1) = .8$
$p(z = 1|u = 1, v = 1, y = 0) = .8$
$p(z = 1|u = 1, v = 0, y = 1) = .8$
$p(z = 1|u = 1, v = 0, y = 0) = .8$
$p(z = 1|u = 0, v = 1, y = 1) = .8$
$p(z = 1|u = 0, v = 1, y = 0) = .8$
$p(z = 1|u = 0, v = 0, y = 1) = .8$
$p(z = 1|u = 0, v = 0, y = 0) = .2$

Passive observational distribution:

$$\text{PM1: } P(X, Y, Z) = \sum_{uv} P(U)P(V)P(X|U)P(Y|V, X)P(Z|U, V, X, Y)$$

$$\text{PM2: } P(X, Y, Z) = \sum_{uv} P(U)P(V)P(X|U)P(Y|V, X)P(Z|U, V, Y)$$

Experimental distribution when x is subject to an intervention (I write $P(A|B||B)$ to mean the conditional probability of A given B in an experiment where B has been subject to a surgical intervention):

$$\text{PM1: } P(Y, Z|X||X) = \sum_{uv} P(U)P(V)P(Y|V, X)P(Z|U, V, X, Y)$$

$$\text{PM2: } P(Y, Z|X||X) = \sum_{uv} P(U)P(V)P(Y|V, X)P(Z|U, V, Y)$$

Experimental distribution when y is subject to an intervention:

$$\text{PM1: } P(X, Z|Y||Y) = \sum_{uv} P(U)P(V)P(X|U)P(Z|U, V, X, Y)$$

$$\text{PM2: } P(X, Z|Y||Y) = \sum_{uv} P(U)P(V)P(X|U)P(Z|U, V, Y)$$

Experimental distribution when z is subject to an intervention:

$$\text{PM1: } P(X, Y|Z||Z) = \sum_{uv} P(U)P(V)P(X|U)P(Y|V, X)$$

$$\text{PM2: } P(X, Y|Z||Z) = \sum_{uv} P(U)P(V)P(X|U)P(Y|V, X)$$

By substituting the terms of PM1 and PM2 in the above equations it can be verified that PM1 and PM2 have identical passive observational and single-intervention distributions but that they differ for the following double-intervention distribution on x and y.

Experimental distribution when x and y are subject to an intervention:

$$\text{PM1: } P(Z|X,Y||X,Y) = \sum_{uv} P(U)P(V)P(Z|U,V,X,Y)$$

$$\text{PM2: } P(Z|X,Y||X,Y) = \sum_{uv} P(U)P(V)P(Z|U,V,Y)$$

PM1 and PM2 (unsurprisingly) have identical distributions for the other two double-intervention distributions, since the $x \rightarrow z$ edge is broken and the remaining parameters are identical in the parameterizations:

Experimental distribution when x and z are subject to an intervention:

$$\text{PM1: } P(Y|X,Z||X,Z) = \sum_{v} P(V)P(Y|V,X)$$

$$\text{PM2: } P(Y|X,Z||X,Z) = \sum_{v} P(V)P(Y|V,X)$$

Experimental distribution when y and z are subject to an intervention:

$$\text{PM1: } P(X|Y,Z||Y,Z) = \sum_{u} P(U)P(X|U)$$

$$\text{PM2: } P(X|Y,Z||Y,Z) = \sum_{u} P(U)P(X|U)$$

Appendix C

Parameterization PM3 for structure 1 in figure 2:

$p(u = 1) = .5$

$p(v = 1) = .5$

$p(x = 1|u = 1) = .8$
$p(x = 1|u = 0) = .2$

$p(y = 1|v = 1, x = 1) = .8$
$p(y = 1|v = 1, x = 0) = .8$
$p(y = 1|v = 0, x = 1) = .8$
$p(y = 1|v = 0, x = 0) = .2$

$\mathbf{p(z = 1|u = 1, v = 1, x = 1, y = 1) = .825}$
$p(z = 1|u = 1, v = 1, x = 1, y = 0) = .8$
$\mathbf{p(z = 1|u = 1, v = 1, x = 0, y = 1) = .8}$
$p(z = 1|u = 1, v = 1, x = 0, y = 0) = .8$
$p(z = 1|u = 1, v = 0, x = 1, y = 1) = .775$
$p(z = 1|u = 1, v = 0, x = 1, y = 0) = .8$
$p(z = 1|u = 1, v = 0, x = 0, y = 1) = .8$
$p(z = 1|u = 1, v = 0, x = 0, y = 0) = .8$
$p(z = 1|u = 0, v = 1, x = 1, y = 1) = .7$
$p(z = 1|u = 0, v = 1, x = 1, y = 0) = .8$
$p(z = 1|u = 0, v = 1, x = 0, y = 1) = .8$
$p(z = 1|u = 0, v = 1, x = 0, y = 0) = .8$
$p(z = 1|u = 0, v = 0, x = 1, y = 1) = .9$
$p(z = 1|u = 0, v = 0, x = 1, y = 0) = .2$
$p(z = 1|u = 0, v = 0, x = 0, y = 1) = .8$
$p(z = 1|u = 0, v = 0, x = 0, y = 0) = .2$

Substituting the parameters of PM3 in the equations for the passive observational or any experimental distributions of PM1 in appendix 2, it can be