Supplementary material for "Towards device-independent information processing on general quantum networks"

Ciarán M. Lee[1, *] and Matty J. Hoban[2, 3]

[1]Department of Physics and Astronomy, University College London, Gower Street, London WC1E 6BT, UK
[2]University of Oxford, Department of Computer Science, Wolfson Building, Parks Road, OX1 3QD, UK.
[3]School of Informatics, University of Edinburgh, 10 Crichton Street, Edinburgh EH8 9AB, UK

Appendix A: Brief background on DAGs

The structure of each DAG encodes conditional independence relations [19] among the nodes. For instance, the no-signalling conditions in the Bell network, $P(A|X,Y) = P(A|X)$, and $P(B|X,Y) = P(B|Y)$, can be seen to directly follow [1] from the structure of the DAG depicted in Fig. 1(a). Indeed, the specification of the DAG subsumes and generalises the standard no-signalling relations [2–4]. Indeed, in the case of the general network from Fig. 2(a), the structure of the DAG ensures $P(A_i|x_i, A_j) = P(A_i|x_i)$ for all $j \notin \{i-1, i, i+1\}$ and $P(A_i|x_i, x_j) = P(A_i|x_i)$ for all $j \neq i$. Additionally, this structure ensures $P(A_i|A_j) = P(A_i)$, for $j \notin \{i-1, i, i+1\}$. This is due to the fact that non-neighbouring agents can only be correlated through knowledge of neighbouring agents outcomes. Moreover, the assumption that each agent has a secure laboratory is enforced by the lack of an arrow from an eavesdropper to each agent's outcome. In short, every constraint governing how the inputs and outputs of agents and eavesdroppers are related is specified by the DAG.

Appendix B: Proof of Result 1

Consider the following conditional distribution $P(A, B, E|X, Y, Z)$, where A, B are binary random variables and X, Y are k-valued, satisfying the "no-signalling" conditions:

$$\begin{aligned} P(A, B|X, Y, Z) &= P(A, B|X, Y) \\ P(A, E|X, Y, Z) &= P(A, E|X, Z) \\ P(B, E|X, Y, Z) &= P(B, E|Y, Z). \end{aligned} \tag{B1}$$

It was shown in Ref.'s [5–8] that

$$D\left(P(E|A, X, Z), P(E|Z)\right) \leq I_k\left(P(A, B|X, Y)\right), \tag{B2}$$

where I_k is the chained Bell inequality [9] on k measurement settings, defined as:

$$\begin{aligned} I_k\left(P(A, B|X, Y)\right) &:= P(A = B|X = 1, B = k) \\ &+ \sum_{|x-y|=1} P(A \neq B|X = x, Y = y) \end{aligned} \tag{B3}$$

Consider the left hand side of inequality (4), it will now be shown to decompose as

$$\begin{aligned} D\Big(&P(E_1 \cdots E_n|A_1, A_{n+1}x_1, x_{n+1}, z_1, \ldots, z_n), \\ &P(E_1|z_1) \cdots P(E_n|z_n)\Big) \leq D\left(P(E_1|A_1, x_1, z_1), P(E_1|z_1)\right) \\ &\qquad + D\left(P(E_n|A_n, x_n, z_n), P(E_n|z_n)\right), \end{aligned} \tag{B4}$$

where again, the systems held by the eavesdropper could be post-quantum (that is, non-signalling). Indeed, the structure of the DAG from Fig. 2(b) implies the following conditional independence relations:

$$\begin{aligned} &P(E_1 \cdots E_n|A_1, A_{n+1}, x_1, x_{n+1}, z_1, \ldots, z_n) = \\ &\qquad P(E_1|A_1, z_1) P(E_n|A_{n+1}, z_n) \prod_{i=2}^{n-1} P(E_i|z_i). \end{aligned} \tag{B5}$$

Combining with the definition of $D(\cdot, \cdot)$ yields Eq. (B4).

The DAG of Fig. 2(b) ensures the no-signalling relations of Eq. (B1) hold between A_i, E_i, x_i, z_i. The conjunction of this with Eq.'s (B2) and (B4) implies

$$\begin{aligned} D\Big(&P(E_1 \cdots E_n|A_1, A_{n+1}x_1, x_{n+1}, z_1, \ldots, z_n), \\ &P(E_1|z_1) \cdots P(E_n|z_n)\Big) \leq 2I_2 \end{aligned} \tag{B6}$$

The chained Bell inequality will now be connected to inequality (1). First let us look at the specific case of the measurements introduced in the repeater network section. After considering this example, the general case will be proved.

To this end, consider the following mapping:

$$\begin{aligned} &P(a_1, a_2^0 a_2^1, \ldots, a_{n+1}|x_1, x_{n+1}) \longrightarrow \\ &\qquad P(a_1, a_2, \ldots, a_n, a_{n+1}|x_1, x_2, \ldots, x_n, x_{n+1}) \\ &= \sum \delta_{a_2, a^{x_2}} \cdots \delta_{a_n, a^{x_n}} P(a_1, a_2^0 a_2^1, \ldots |x_1, x_{n+1}), \end{aligned} \tag{B7}$$

where the sum ranges over $\{a_i^0 a_i^1\}_i$. One can interpret the above mapping as follows: agents $i = 2$ to $i = n$ use their devices to simulate a two choice $x_i \in \{0, 1\}$, binary outcome measurement by outputting the bit $a_i^{x_i}$ from the pair output by their device $a_i^0 a_i^1$. It is clear that as each agent makes their choice locally, no correlations have been introduced between agents. Hence, both I, J from Eq. (2) and $\mathcal{N}$ from Eq. (1) are invariant under mapping (B7). For a more in-depth discussion on this point, see section III B from [10].

*Electronic address: ciaran.lee@ucl.ac.uk

Applying the above mapping to Eq. (3) from the main text yields

$$\frac{1+(-1)^{\sum_{i=2}^{n}a_i}\frac{\left(\prod_{i=2}^{n}\delta_{x_i,0}+(-1)^{x_1+x_{n+1}}\prod_{i=2}^{n}\delta_{x_i,01}\right)}{2}}{2^{n+1}} \tag{B8}$$

Any repeater network in which agents 2 to n have binary inputs and outputs that generates the above correlations has the same value for $\mathcal{N}$ as the one considered in the main text. Indeed, Eq. (B8) can be generated by agents 2 to n performing the separable measurements $A_i^y = A_{i,0}^y \otimes A_{i,1}^y \in \{\sigma_z\otimes\sigma_z, \sigma_x\otimes\sigma_x\}$, for $i=2,\dots,n$ [10]. Here, an agent measures both of their received systems in the same basis and outputs the parity of the individual measurement outcomes. Given these separable measurements for intermediate nodes, one has $\langle A_1^{x_1}A_2\cdots A_{n+1}^{x_{n+1}}\rangle = \langle A_1^{x_2}A_{2,0}^{x_2}\rangle\langle A_{2,1}^{x_2}A_{3,0}^{x_3}\rangle\cdots\langle A_{n,1}^{x_n}A_{n+1}^{x_{n+1}}\rangle$. Moreover, it follows that $\langle A_{i,1}^{x_i}A_{i+1,0}^{x_{i+1}}\rangle = 1$. Combining all this, it follows that

$$\begin{aligned} I &= \frac{1}{4}\left(\langle A_1^0A_{2,0}^0\rangle + \langle A_1^1A_{2,0}^0\rangle\right)\left(\langle A_{n+1}^0A_{n,1}^0\rangle + \langle A_{n+1}^1A_{n,1}^0\rangle\right)\\ J &= \frac{1}{4}\left(\langle A_1^0A_{2,0}^1\rangle - \langle A_1^1A_{2,0}^1\rangle\right)\left(\langle A_{n+1}^0A_{n,1}^1\rangle - \langle A_{n+1}^1A_{n,1}^1\rangle\right). \end{aligned} \tag{B9}$$

As agents A_1 and A_{n+1} and $A_{2,0}$ and $A_{n,1}$ respectively choose from the same set of measurements, one has

$$\begin{aligned} \mathcal{R} = \frac{1}{2}&|\langle A^0B^0\rangle + \langle A^1B^0\rangle| \\ &+\frac{1}{2}|\langle A^0B^1\rangle - \langle A^1B^1\rangle|, \end{aligned} \tag{B10}$$

where $\langle A^0B^0\rangle := \langle A_1^0A_{2,0}^0\rangle = \langle A_{n+1}^0A_{n,1}^0\rangle$ formalises the statement that A_1, A_{n+1} and $A_{2,0}, A_{n,1}$ choose from the same set of measurements, respectively.

The above is the $k=2$ instance of the CHSH inequality

$$\mathcal{C}_k = \sum_{i=0}^{k-1}\langle A_iB_i\rangle + \sum_{i=0}^{k-2}\langle A_{i+1}B_i\rangle - \langle A_0B_{k-1}\rangle. \tag{B11}$$

Using $\langle AB\rangle = 2P(A=B)-1$, it follows that

$$I_k = k - \frac{1}{2}\mathcal{C}_k. \tag{B12}$$

Combing this with Eq. (B6), Eq. (4) follows.

For the general case, consider the following. As in the above example, agents $i=2$ to $i=n$ use their devices to simulate a two choice $x_i\in\{0,1\}$, binary outcome measurement by outputting the bit $a_i^{x_i}$ from the pair output by their device $a_i^0a_i^1$. This does not change the value of the polynomial Bell inequality. Moreover, as this amounts to a classical post-processing of intermediate agents outcomes, it also does not affect any potential correlations between an eavesdropper and the first and last agent in the network.

It was shown in Theorem 1 of Ref. [11] that coarse-graining a two-qubit measurement with four outcomes in the manner discussed above results in a separable measurement. This is in fact always true, even for a two-qudit measurement. Indeed, as the dimension is not given a priori, without loss of generality one can always take this measurement to consist of four rank-1 projectors by appending ancillary systems. Coarse graining these as above results in operators which satisfy the conditions of Lemma 1 of [12], which states that these can be reduced to a direct sum of 2-qubit operators, as is done in Eq. (9) of [13] for instance. The problem has now been reduced to the case of 2-qubit measurements which were covered by [11], as discussed above.

Given intermediate separable measurements, one again has that $\langle A_1^{x_1}A_2\cdots A_{n+1}^{x_{n+1}}\rangle = \langle A_1^{x_2}A_{2,0}^{x_2}\rangle\langle A_{2,1}^{x_2}A_{3,0}^{x_3}\rangle\cdots\langle A_{n,1}^{x_n}A_{n+1}^{x_{n+1}}\rangle$. Following the reasoning of Eq. (B9), one can show the following holds (see Eq. (13) of Ref. [14] for more details):

$$\mathcal{R} \le \frac{1}{2}\sqrt{\mathcal{C}_k^{A_1A_{2,0}}\mathcal{C}_k^{A_{n,1}A_{n+1}}}$$

where $\mathcal{C}_k^{A_iA_j}$ is the CHSH inequality between agents i and j. Hence, using the fact that the arithmetic mean is larger than the geometric mean, one has

$$\mathcal{R} \le \frac{\mathcal{C}_k^{A_1A_{2,0}}}{4} + \frac{\mathcal{C}_k^{A_{n,1}A_{n+1}}}{4}.$$

Using Eq. (B12), one obtains

$$2I_2 = 4 - \frac{\mathcal{C}_k^{A_1A_{2,0}}}{2} - \frac{\mathcal{C}_k^{A_{n,1}A_{n+1}}}{2} \le 2\left(2-\mathcal{R}\right).$$

Inputting into Eq. (B6), provides the desired result.

Appendix C: Proof of result 2: classically simulating the quantum correlations of Eq. (3)

It will now be demonstrated that by correlating the $i=1$ and n sources, an eavesdropper can simulate the correlations of Eq. (3). Moreover, the sources only need to emit classical variables. To achieve this, the eavesdropper sends independent and uniformly distributed bits $\{\alpha,\lambda_i\}$ to agent A_i, for $i=1$ and $n+1$. Given these bits and the agents' input, the agent's device outputs $a_i = \lambda_i \oplus \alpha x_i$. The conditional probability distribution characterising the action of the device is $P(a_i|\alpha,\lambda_i,x_i) = \frac{1}{2}\left(1+(-1)^{a_i+\lambda_i+\alpha x_i}\right)$.

Agent A_i, for $i=2$ and n, is sent independent, uniformly distributed bits $\{\alpha,\nu_i,\lambda_{i-1},\lambda_i,\widetilde{\lambda}_i\}$, on receipt of which their device outputs

$$A_i = (a_i^0,a_i^1) = \begin{cases}(\lambda_{i-1}\oplus\lambda_i,\nu_i), & \text{if } \alpha=0,\\ (\nu_i,\lambda_{i-1}\oplus\widetilde{\lambda}_i), & \text{if } \alpha=1.\end{cases} \tag{C1}$$

Note that agents $1,2,n-1,$ and n receive a copy of the bit α. Hence, source 1 and n are now correlated. Recalling the transmitted bits are uniformly distributed, the conditional probability distribution characterising this device will now be derived:

$$\sum_{\nu} P\left(a_i^0 a_i^1|\alpha,\nu,\lambda_{i-1},\lambda_i,\widetilde{\lambda}_i\right)P(\nu)$$
$$=\sum_{\nu}\Big[P(a_i^0|\lambda_{i-1},\lambda_i)P(a_i^1|\nu)P(\nu)\delta_{\alpha,0}$$
$$+P(a_i^1|\lambda_{i-1},\widetilde{\lambda}_i)P(a_i^0|\nu)P(\nu)\delta_{\alpha,1}\Big]$$
$$=\frac{1}{4}\left(1+(-1)^{a_i^0+\lambda_{i-1}+\lambda_i}\delta_{\alpha,0}+(-1)^{a_i^1+\lambda_{i-1}+\widetilde{\lambda}_i}\delta_{\alpha,1}\right).$$

Finally, all remaining agents $A_i = a_i^0 a_i^1$, $i = 3,\dots,n-1$, are sent uniformly distributed bits $\{\lambda_{i-1},\widetilde{\lambda}_{i-1},\lambda_i,\widetilde{\lambda}_i\}$. On receipt of which their devices output $a_i^0 = \lambda_{i-1}\oplus\lambda_i$ and $a_i^1 = \widetilde{\lambda}_{i-1}\oplus\widetilde{\lambda}_i$. The conditional probability distribution is $\frac{1}{4}\left(1+(-1)^{a_i^0+\lambda_{i-1}+\lambda_i}\right)\left(1+(-1)^{a_i^1+\widetilde{\lambda}_{i-1}+\widetilde{\lambda}_i}\right)$. Combining all of these conditional probability distributions yields the following:

$$P(a_1,a_2^0a_2^1,\dots,a_n^0a_n^1,a_{n+1}|x_1,x_{n+1})=$$
$$\sum_{\alpha,\nu,\lambda_1,\dots,\lambda_n}P(a_1|\alpha,\lambda_1)P(a_2^0a_2^1|\alpha,\nu,\lambda_1,\lambda_2,\widetilde{\lambda}_2)P(a_3^0a_3^1|\lambda_{i-1},\widetilde{\lambda}_{i-1},\lambda_i,\widetilde{\lambda}_i)\cdots P(a_1|\alpha,\lambda_1)P(\alpha)P(\nu)P(\lambda_1)\cdots P(\lambda_n)$$
$$=\frac{1}{2^{2n}}\sum_{\alpha,\lambda_1,\dots,\lambda_n}\left(1+(-1)^{a_i+\lambda_i+\alpha x_i}\right)\left(1+(-1)^{a_i^0+\lambda_{i-1}+\lambda_i}\delta_{\alpha,0}+(-1)^{a_i^1+\lambda_{i-1}+\widetilde{\lambda}_i}\delta_{\alpha,1}\right)\cdots P(\lambda_n)P(\alpha)$$
$$=\frac{1}{2^{2n}}\left(1+(-1)^{a_1+a_{n+1}}\sum_{\alpha}\left((-1)^{\sum_{i=2}^n a_i^0}\delta_{\alpha,0}P(\alpha)+(-1)^{\sum_{i=2}^n a_i^1+x_1+x_{n+1}}\delta_{\alpha,1}P(\alpha)\right)\right).$$

Performing the sum over α and recalling that $P(\alpha=0)=1/2=P(\alpha=1)$, results in the quantum distribution of Eq. (3). Hence the eavesdropper can perfectly simulate quantum correlations by correlating sources thought to be independent. The last line of the above equation follows by noting that, as one multiplies out each conditional distribution, terms of the form $\sum_\lambda(-1)^\lambda$ vanish.

Appendix D: Proof of Result 4

The proof will follow the strategy of [15, Proof of Theorem 1] and [16, Proof of Eq. (20)]. Consider a classical model for Fig. 3(a), where all the λ_i are random variables. Writing

$$\langle A^i_{x_i}\rangle_{\lambda_i}=\sum_{a_i}(-1)^{a_i}P(A_i=a_i|x_i,\lambda_i)$$
$$\langle B_y\rangle_\lambda=\sum_b(-1)^bP(B=b|y,\lambda),\tag{D1}$$

where λ is shorthand for $\lambda_1\cdots\lambda_n$, one has

$$I_i=\frac{1}{2^n}\sum_{x_1,\dots,x_n=i}^{i+1}\int\Big(\prod_{j=1}^n q_j(\lambda_j)\langle A^j_{x_j}\rangle_{\lambda_j}\Big)\langle B_i\rangle_\lambda d\lambda_j\tag{D2}$$

for $i=0,\dots,k-1$, where $q_j(\lambda_j)$ is the distribution over the λ_i's. Taking the absolute value yields

$$|I_i|\le\prod_{j=1}^n\Big(\frac{1}{2}\int q_j(\lambda_j)\Big|\sum_{x_j=1}^n\langle A^j_{x_j}\rangle_{\lambda_j}\Big|d\lambda_j\Big),\tag{D3}$$

as $\left|\langle B_i\rangle_\lambda\right|\le 1$.

It was shown in Ref. [15] that, for $c_i^k\in\mathbb{R}_+$ and $m,n\in\mathbb{N}$, the following holds:

$$\sum_{k=1}^m\Big(\prod_{i=1}^n c_i^k\Big)^{1/n}\le\prod_{i=1}^{i+1}\left(c_i^1+c_i^2+\cdots+x_i^m\right)^{1/n}.\tag{D4}$$

Applying this result to $\mathcal{S}=\sum_{i=0}^{k-1}|I_i|^{1/n}$ yields

$$\mathcal{S}\le\Big[\prod_{j=1}^n\frac{1}{2}\int q_j(\lambda_j)\Big(\Big|\langle A_0^j\rangle_{\lambda_j}+\langle A_1^j\rangle_{\lambda_j}\Big|+\Big|\langle A_1^j\rangle_{\lambda_j}+\langle A_2^j\rangle_{\lambda_j}\Big|+\cdots+\Big|\langle A_{k-1}^j\rangle_{\lambda_j}-\langle A_0^j\rangle_{\lambda_j}\Big|\Big)d\lambda_j\Big]^{1/n}.\tag{D5}$$

The following upper bound holds:

$$\frac{1}{2}\Big(\Big|\langle A_0^j\rangle_{\lambda_j}+\langle A_1^j\rangle_{\lambda_j}\Big|+\Big|\langle A_1^j\rangle_{\lambda_j}+\langle A_2^j\rangle_{\lambda_j}\Big|+\cdots+\Big|\langle A_{k-1}^j\rangle_{\lambda_j}-\langle A_0^j\rangle_{\lambda_j}\Big|\Big)\le k-1.\tag{D6}$$

Hence, one has

$$\mathcal{S}\le\Big(\prod_{j=1}^n\int q_j(\lambda_j)\,(k-1)^n\,d\lambda_j\Big)^{1/n}=k-1\tag{D7}$$

finishing the derivation of Eq. (5).

Appendix E: Proof of result 5

The structure of the DAG from Fig. 3(a) yields the following conditional independence relation:

$$P(E_1 \cdots E_n | A_1, \ldots A_n, x_1, \ldots x_{n+1}, z_1, \ldots, z_n) = P(E_1|A_1, z_1) P(E_2|A_2, z_2) \cdots P(E_n|A_n, z_n). \tag{E1}$$

From this it follows that

$$D\Big(P(E_1 \cdots E_n | A_1, \ldots, A_n, x_1, \ldots, x_n, z_1, \ldots, z_n), P(E_1|z_1) \cdots P(E_n|z_n)\Big) \leq \sum_i D\left(P(E_i|A_i, x_i, z_i), P(E_i|z_i)\right).$$

As stated in the main text, it is assumed that the central agents device is implementing separable measurements, hence one has $B_y = B_y^1 \otimes \cdots \otimes B_y^n$. Given this separable measurements, one can show that the following holds (again, see Eq. (13) of Ref. [14] for more details)

$$\mathcal{S} \leq \frac{1}{2} \left(\prod_i C_k^{A^i B^i} \right)^{\frac{1}{n}}. \tag{E2}$$

Following the same analysis as the end of Appendix B, the conjunction of Eq. (E2) with Eq. (B2) yields Eq. (6).

The upper bound of Eq. (E2) can in fact be reached using the measurements introduced in the main paper. As all external agents choose from the same set of measurements, one has $C_k^{A^i B^i} = C_k^{A^j B^j}, \forall i, j$. This implies the maximal quantum value of Inequality (5) is the maximal quantum value of C_k^{AB}, which has been shown by [17] to be

$$k \cos\left(\frac{\pi}{2k}\right).$$

Appendix F: Proof of result 6

Combing Eq. (E1) with the fact that p_l from Eq. (7) satisfies $|p_l| \leq 1 \ \forall i$, Eq. (8) follows.

[1] C. J. Wood and R. W. Spekkens, "The lesson of causal discovery algorithms for quantum correlations: Causal explanations of bell-inequality violations require fine-tuning," *New Journal of Physics*, vol. 17, no. 033002, 2015.

[2] J. Henson, R. Lal, and M. F. Pusey, "Theory-independent limits on correlations from generalized bayesian networks," *New Journal of Physics*, vol. 16, no. 113043, 2014.

[3] T. Fritz, "Beyond bell's theorem: correlation scenarios," *New Journal of Physics*, vol. 14, no. 103001, 2012.

[4] R. Chaves, R. Kueng, J. B. Brask, and D. Gross, "Unifying framework for relaxations of the causal assumptions in bell's theorem," *Phys. Rev. Lett.*, vol. 114, no. 140403, 2015.

[5] J. Barrett, A. Kent, and S. Pironio, "Maximally nonlocal and monogamous quantum correlations," *Phys. Rev. Lett.*, vol. 97, no. 170409, 2006.

[6] R. Colbeck and R. Renner, "Hidden variable models for quantum theory cannot have any local part," *Phys. Rev. Lett.*, vol. 101, no. 050403, 2008.

[7] R. Colbeck and R. Renner, "No extension of quantum theory can have improved predictive power," *arXiv preprint arXiv:1005.5173*, 2010.

[8] J. Barrett, R. Colbeck, and A. Kent, "Unconditionally secure device-independent quantum key distribution with only two devices," *Physical Review A*, vol. 86, no. 062326, 2012.

[9] S. L. Braunstein and C. M. Caves, "Wringing out better bell inequalities," *Annals of Physics*, vol. 202, no. 1, pp. 22–56, 1990.

[10] K. Mukherjee, B. Paul, and D. Sarkar, "Correlations in n-local scenario," *Quantum Information Processing*, vol. 14, pp. 2025–2042, 2015.

[11] F. Andreoli, G. Carvacho, L. Santodonato, R. Chaves, and F. Sciarrino, "Maximal violation of n-locality inequalities in a star-shaped quantum network," *arXiv preprint arXiv:1702.08316*, 2017.

[12] L. Masanes, "Asymptotic violation of bell inequalities and distillability," *Physical Review Letters*, vol. 97, no. 5, p. 050503, 2006.

[13] R. Rabelo, M. Ho, D. Cavalcanti, N. Brunner, and V. Scarani, "Device-independent certification of entangled measurements," *Physical Review Letters*, vol. 107, no. 5, p. 050502, 2011.

[14] A. Tavakoli, M. O. Renou, N. Gisin, and N. Brunner, "Correlations in star networks: from bell inequalities to network inequalities," *arXiv preprint arXiv:1702.03866*, 2017.

[15] A. Tavakoli, P. Skrzypczyk, D. Cavalcanti, and A. Acín, "Nonlocal correlations in the star-network configuration," *Physical Review A*, vol. 90, no. 062109, 2014.

[16] C. Branciard, D. Rosset, N. Gisin, and S. Pironio, "Bilocal versus nonbilocal correlations in entanglement-swapping experiments," *Physical Review A*, vol. 85, no. 3, p. 032119, 2012.

[17] S. Wehner, "Tsirelson bounds for generalized clauser-horne-shimony-holt inequalities," *Physical Review A*, vol. 73, no. 022110, 2006.

[18] J. Pearl, *Causality: Models, Reasoning, and Inference*. Cambridge: Cambridge University Press, 2nd ed., 2009.

[19] Here the *faithfulness* condition is being assumed, see [1, 18] for a discussion.