 Device-independent certification of non-classical measurements via causal models

Ciarán M. Lee

1Department of Physics and Astronomy, University College London, Gower Street, London WC1E 6BT, UK

Quantum measurements are crucial for quantum technologies and give rise to some of the most classically counter-intuitive quantum phenomena. As such, the ability to certify the presence of genuinely non-classical measurements in a device-independent fashion is vital. However, previous work has either been non-device-independent, or has relied on post-selection. In the case of entanglement, the post-selection approach applies an entangled measurement to independent states and post-selects the outcome, inducing non-classical correlations between the states that can be device-independently certified using a Bell inequality. That is, it certifies measurement non-classicality not by what it is, but by what it does. This paper remedies this discrepancy by providing a novel notion of what measurement non-classicality is, which, in analogy with Bell’s theorem, corresponds to measurement statistics being incompatible with an underlying classical causal model. It is shown that this provides a more fine-grained notion of non-classicality than post-selection, as it certifies the presence of non-classicality that cannot be revealed by examining post-selected outcomes alone.

Quantum measurements are a key resource behind most quantum technologies [1] and, moreover, they reveal some of the most startling non-classical features of quantum theory [2,3]. Indeed, performing joint quantum measurements on composite systems is a key feature behind quantum teleportation, superdense coding, metrology, cryptography [3], quantum repeaters [1], and quantum networks more generally. Hence the ability to certify the non-classical nature of quantum measurements is vitally important for the functioning of quantum technology and additionally, for understanding some of the fundamental differences between quantum and classical physics. Moreover, as the manufacturers of quantum measurement devices may not always be trusted, such certifications should be device-independent. That is, they should rely only on output measurement statistics rather than any intrinsic quantum properties, such as knowledge of the underlying Hilbert space dimension.

Previous work on the certification of joint quantum measurements [24] falls into two categories. The first uses witnesses to certify the presence of non-classical measurements [5,7], but is manifestly not device-independent. The second is device-independent, but requires post-selection to certify the presence of a quantum measurement [8]. In the case of entanglement, such certification is accomplished by exploiting the fact that applying an entangled measurement to two initially independent entangled states and post-selecting the outcome induces entanglement between the states, which can then be certified device-independently by violating a Bell inequality. This method hence detects quantum measurements through their action on states. That is, it certifies an entangled measurement through what it does, not what it is. This is in stark contrast with entangled states, whose non-classicality is easily certified through the violation of a Bell inequality. Such violation implies a denial of (at least one of) the assumptions underlying Bell’s theorem. The modern treatment of which utilises the classical causal model framework to unify Bell’s original assumptions [9,11]. Composite states are thus said to be non-classical if the correlations generated by performing it on local preparations on each composite system are inconsistent with an underlying classical causal model. In the following section this classical causal model is introduced and a non-linear inequality on any distribution generated by it is derived. Violation of this inequality entails that the observed correlations are in conflict with the classical causal model. As the inequality depends only on observed output statistics, it is manifestly device-independent. Additionally, it will be demonstrated that this inequality provides a finer-grained notion of joint measurement non-classicality for general quantum measurements than the post-selection approach of Ref. [8], discussed above, as it certifies the presence of non-classicality that cannot be revealed by examining post-selecting the outcomes of measurement alone.

Certifying non-classical measurements.— Recently, tools and techniques from the classical causal models framework have begun to see myriad applications in quantum information [9–16,28]. In this framework, the inputs and outputs of agents measurement and preparation devices are represented by nodes in directed acyclic graphs (DAGs), with the arrows denoting the causal relationship between nodes. The structure of each DAG encodes conditional independence relations [29] among the nodes. For instance, the no-signalling conditions \( P(A|X,Y) = P(A|X) \) and \( P(B|X,Y) = P(B|Y) \), follow directly [9] from the structure of the DAG from in Fig. Furthermore, the structure of the DAG specifies all the conditional independences between the nodes [17,18]. In short, every relation between the inputs and outputs of

*Electronic address: ciaran.lee@ucl.ac.uk
Consider three agents Alice, Bob, and Charlie. Alice and Bob both have devices which prepare a quantum state from some ensemble of states, given a choice between different ensembles. Charlie has a measurement device which jointly measures the states prepared by Alice and Bob. The actions of these devices are represented in a black-box manner. Alice and Bob’s devices have a classical input \( x, y \) (the choice of different ensembles) respectively, and a classical output \( a, b \) (the state prepared) respectively. Here it is assumed that \( a, b, x, y \in \{0,1\} \).

Charlie has no classical input, as his device only performs a single measurement, but has a classical output \( C \) indexing the possible measurement outcomes. It is assumed in this section that \( C \) takes four values and hence is indexed by two bits, \( C = c_0c_1 \in \{00,01,10,11\} \). Preparing and measuring states in this manner gives rise to a conditional probability distribution \( P(a, b, c_0c_1|x, y) \).

In analogy with Bell’s theorem, a classical model for \( P(a, b, c_0c_1|x, y) \) is described by the DAG in Fig. 1 where \( \lambda_1, \lambda_2 \) are unobserved, independent random variables. If the correlations generated by performing Charlie’s measurement on Alice and Bob’s preparations are consistent with the DAG of Fig. 1 then they are said to be classical. That is, they are mediated by the hidden random variables \( \lambda_1, \lambda_2 \). One might wonder why there are two hidden variables, rather than one. This is due to the independence of Alice and Bob’s devices: \( P(a, b|x, y) = \sum_{c_0c_1} P(a, b, c_0c_1|x, y) = P(a|x)P(b|y) \). If the correlations between Alice, Bob, and Charlie were mediated by a single hidden variable, then Alice and Bob’s marginal distribution would not be independent.

**Result 1.** A distribution \( P(a, b, c_0c_1|x, y) \) generated by the DAG of Fig. 1 satisfies:

\[
\sqrt{|M|} + \sqrt{|N|} \leq 1,
\]

in which \( M = \frac{1}{4} \sum_{xy} (A_xB_yC^0) \) and \( N = \frac{1}{4} \sum_{xy} (-1)^{x+y}(A_xB_yC^1) \), and

\[
\langle A_xB_yC^i \rangle = \sum_{abc_0c_1} (-1)^{a+b+c_i} P(a, b, c_0c_1|x, y).
\]

Note that, in contrast to standard Bell inequalities, the inequality presented above is non-linear in the joint distribution \( P(a, b, c_0c_1|x, y) \). This is due to the independence of Alice and Bob’s preparations.

The proof of Result 1 is similar to the derivation of the bilocality inequality from Ref. [19], with a few key differences. First, in the case considered here, the hidden variables can a priori depend on the choice of preparation. That is, it does not follow from the DAG of Fig. 1 that \( P(\lambda_1|x) = P(\lambda_1) \). Lastly, Alice and Bob have preparation devices, rather than measurement devices.

**Proof.** Given the structure of the DAG from Fig 1 it follows that \( P(a, b, c_0c_1|x, y) \) decomposes as

\[
\int d\lambda_1 d\lambda_2 P(a|x)P(b|y)P(c_0c_1|\lambda_1\lambda_2)P(\lambda_1|x)P(\lambda_2|y).
\]

Define \( \langle A_x \rangle = \sum_{x} (-1)^a P(a|x) \), \( \langle B_y \rangle = \sum_{y} (-1)^b P(b|y) \), and \( \langle C^i \rangle_{\lambda_1\lambda_2} = \sum_{c_0c_1} (-1)^c P(c_0c_1|\lambda_1\lambda_2) \). It follows from the above decomposition that one can write \( \langle A_xB_yC^i \rangle \) as

\[
\int d\lambda_1 d\lambda_2 \langle A_x \rangle \langle B_y \rangle \langle C^i \rangle_{\lambda_1\lambda_2} P(\lambda_1|x)P(\lambda_2|y).
\]

This, together with \( \langle C^i \rangle_{\lambda_1\lambda_2} \leq 1 \), implies

\[
|M| \leq \left( \int d\lambda_1 \frac{|\langle A_0 \rangle| P(\lambda_1|0) + |\langle A_1 \rangle| P(\lambda_1|1)|}{2} \right) \cdot \left( \int d\lambda_2 \frac{|\langle B_0 \rangle| P(\lambda_2|0) + |\langle B_1 \rangle| P(\lambda_2|1)|}{2} \right).
\]

One can similarly bound \( |N| \), noting that \( (-1)^{x+y} \) introduces \(-1\)’s. For any real \( z, w, z', w' \geq 0 \), the inequality \( \sqrt{zw} + \sqrt{z'w'} \leq \sqrt{z + z'} \sqrt{w + w'} \) is valid [19]. Hence

\[
\sqrt{|M|} + \sqrt{|N|} \leq \int d\lambda_1 \left( \frac{|\langle A_0 \rangle| P(\lambda_1|0) + |\langle A_1 \rangle| P(\lambda_1|1)|}{2} + \frac{|\langle A_0 \rangle| P(\lambda_1|0) - |\langle A_1 \rangle| P(\lambda_1|1)|}{2} \right) \cdot \int d\lambda_2 \left( \frac{|\langle B_0 \rangle| P(\lambda_2|0) + |\langle B_1 \rangle| P(\lambda_2|1)|}{2} + \frac{|\langle B_0 \rangle| P(\lambda_2|0) - |\langle B_1 \rangle| P(\lambda_2|1)|}{2} \right)
\]

\[
\leq \int d\lambda_1 |\langle A_0 \rangle| P(\lambda_1|0)| \langle A_1 \rangle| P(\lambda_1|1)| \cdot \int d\lambda_2 |\langle B_0 \rangle| P(\lambda_2|0)| \langle B_1 \rangle| P(\lambda_2|1)| \leq 1 \quad \square
\]
The bound from Result 1 can be classically saturated. To see this, consider the following. Let \( x,y \) be independent, uniformly distributed random bits. Let Alice’s (Bob’s) device output \( a = x \oplus 1 \) (\( b = y \oplus 1 \)) with probability one, and let \( \lambda_1 (\lambda_2) \) equal \( a \oplus 1 \) (\( b \oplus 1 \)) with probability one. Let Charlie have two independent and identically distributed random bits \( \mu_0 \) and \( \mu_1 \). When both \( \mu_0 \) and \( \mu_1 \) equal zero, Charlie’s device outputs \((c_0, c_1) = (\lambda_1 \oplus \lambda_2, \nu)\) with probability one, where \( \nu \) is another random bit. When \( \mu_0 \) and \( \mu_1 \) equal one, Charlie’s device outputs \((c_0, c_1) = (\nu, \lambda_1 \oplus \lambda_2)\) with probability one. When \( \mu_0 \neq \mu_1 \) Charlie’s device outputs \((c_0, c_1) = (\lambda_1, \lambda_2)\) with probability one. When \( \mu_0 = \mu_1 = 0 \) it follows by a straightforward calculation that \( M = 1 \) and \( N = 0 \), and when \( \mu_0 = \mu_1 = 1 \), \( M = 0 \) and \( N = 1 \). In all remaining cases \( M = N = 0 \). As the probability that \( \mu_0 = \mu_1 = 0 \) is \( r^2 \) and the probability that \( \mu_0 = \mu_1 = 1 \) is \( (1 - r)^2 \), where \( r = P(\mu_0 = 0) = P(\mu_1 = 0) \), all points \((M,N) = (r^2, (1-r)^2)\) can be achieved. The boundary \( \sqrt{M} + \sqrt{N} = 1 \) is thus classically saturated.

Quantum violation.—Recall that Alice and Bob’s devices prepare a state from an ensemble of two states, given a choice between two possible ensembles. A simple quantum realisation of Alice’s (Bob’s) device is to prepare a maximally entangled \(|\psi^-\rangle\) state between Alice’s (Bob’s) system and an ancilla, and perform a measurement on the ancilla to prepare a state on Alice’s (Bob’s) system. Given two distinct measurements that can be performed on the ancilla, there are two distinct ensembles of states to which Alice’s (Bob’s) system can be steered. The specific measurement outcome prepares a fixed state from the ensemble associated with that measurement.

Now, consider the correlations generated by performing either \((\sigma_Z + \sigma_X)/2\) (for \( x = y = 0 \)) or \((\sigma_Z - \sigma_X)/2\) (for \( x = y = 1 \)) on Alice and Bob’s ancilla and performing a ‘noisy’ Bell state measurement on Charlie’s system \( \{E_{c_0,c_1}\} \), where \( E_{c_0,c_1} = p_0 |\psi_{c_0,c_1}\rangle \langle \psi_{c_0,c_1}| + (1 - p_0)I/4 \), where \( \{|\psi_{00}\rangle, |\psi_{01}\rangle, |\psi_{10}\rangle, |\psi_{11}\rangle\} \) is the Bell state measurement. As \( E_{c_0,c_1} \geq 0, \forall c_0,c_1 \), and \( \sum_{c_0,c_1} E_{c_0,c_1} = 1 \), it is a valid measurement. The correlations generated here are the same as those considered in Section III A of Ref. [19], namely:

\[
P(a,b,c_0|c_1|x,y) = \frac{1}{16} \left( 1 + p(-1)^a b \left\{ (-1)^{c_0} + (-1)^x y + c_1 \right\} \right).
\]

From this one obtains \( \sqrt{M} + \sqrt{N} = \sqrt{2p} \), providing a quantum violation for \( p > 1/2 \).

Post-selection.—Ref. [8] demonstrated that the presence of an entangled measurement can be certified in an device-independent fashion using post-selection. This was achieved by exploiting the fact that performing an entangled measurement on two initially independent entangled states and post-selecting the outcome induces entanglement between the states, which can then be certified device-independently by violating a Bell inequality. This method hence detects entangled measurements through their action on states, by showing that for each fixed measurement outcome the induced correlations are non-classical. In the current work a novel method has been introduced which certifies general measurement non-classicality not through what it does, but what it is. These two approaches coincide for entangled measurements [8], but do they coincide for general non-classical measurements? That is, if a measurement is non-classical in the sense that it violates the inequality from Result 1, are the correlations induced between Alice and Bob’s devices on post-selection of Charlie’s outcome always non-classical? It will now be shown that, surprisingly, the existence of a separate classical model for each post-selected measurement outcome does not imply the measurement is classical in the sense of Fig. [1].

Note that given the realisations of Alice and Bob’s devices involving steering using projective measurements on an ancilla, introduced in the previous section, it follows that non-classical correlations between Alice and Bob’s preparation devices are equivalent to non-classical correlations between projective measurements performed on their ancillas.

Now, consider the following. Allow Charlie to perform a noisy Bell state measurement with noise parameter \( p \) and post-select on an arbitrary fixed outcome. If Alice and Bob each have their own Bell state, then Charlie’s joint measurement on two of their systems induces a noisy Bell state—with the same noise parameter \( p \)—between Alice and Bob’s ancilla. For instance, if Charlie post-selects outcome \( E_{00} = p|\psi^-\rangle \langle \psi^-| + (1 - p)I/4 \), then Alice and Bob’s ancilla will be in the \( p|\psi^-\rangle \langle \psi^-| + (1 - p)I/4 \) state. Hence, classically simulating Charlie’s joint noisy Bell state measurement on Alice and Bob’s preparations is equivalent to classically simulating local projective measurements on Alice and Bob’s ancilla’s in the induced noisy Bell state. As shown in [20], such correlations can be classically simulated for \( p < 0.66 \). But, as shown in the previous section, the non-post-selected measurement is non-classical as long as \( p > 1/2 \). To summarise, the following has been shown:

**Result 2.** The existence of a separate classical model for each measurement outcome—adhering to the constraints imposed by Fig. [7]—does not imply the measurement is classical in the sense of Fig. [1].

An intuitive explanation of this result could be that, as Charlie’s measurements outcomes can overlap on certain states, classical models for each individual measurement outcome cannot always be combined consistently.

Generalisation to \( n \) systems and \( 2^n \) outcomes.—The inequality from Result 1 will now be generalised to allow for \( n \) systems, \( k \) choices for the each preparation device—each of which have two possible outcomes—and \( 2^k \) possible outcomes for Charlie’s joint measurement, indexed using \( k \) bits \( c_0 \cdots c_{k-1} \). Result 1 corresponds to the \( n = k = 2 \) case. As before, the classical causal model is depicted in Fig. [2].
implied by the structure of Fig. 2, it follows that

$$I_n = \sum_{i=0}^{k-1} P(a_i, \ldots, a_n|c_{k-1}|x_1, \ldots, x_n),$$

with $a_i, c_j \in \{0, 1\}$ and $x_i \in \{0, \ldots, k-1\}$, generated by the DAG of Fig. 2 satisfies the following inequality:

$$S := \sum_{i=0}^{k-1} |I_i|^{1/n} \leq k - 1,$$

where

$I_i = \frac{1}{2^n} \sum_{x_1, \ldots, x_n=1}^{i+1} \langle A_{x_1}^1 \cdots A_{x_n}^n, C_i \rangle$, for $i$ ranging from 0 to $k - 1$, with $A_{x_1}^1 = -A_{x_1}^1$ and $(A_{x_1}^1 \cdots A_{x_n}^n, B_y) = \sum (-1)^{y+i} P(a_1, \ldots, a_n|c_{k-1}|x_1, \ldots, x_n)$. 

Proof. Given the decomposition of the distribution over the agents preparations and Charlie’s measurement,

$$P(a_1, \ldots, a_n|c_{k-1}|x_1, \ldots, x_n),$$

implied by the structure of Fig. 2, it follows that

$$|I_i| \leq \prod_{j=1}^{n} \left( \frac{1}{2} \right) \int \sum_{x_j=1}^{n} (A_{x_j}^1 p(\lambda_j|x_j)|d\lambda_j),$$

where $\langle A_{x_j}^1 \rangle = \sum_{a_j} \left( -1 \right)^{a_j} P(a_j|x_j)$.

It was shown in Ref. [21] that, for $c_{i}^k \in \mathbb{R}_+$ and $m, n \in \mathbb{N}$, the following holds:

$$\sum_{k=1}^{m} \left( \prod_{i=1}^{n} c_{i}^k \right)^{1/n} \leq \prod_{i=1}^{n} \left( c_{i}^1 + c_{i}^2 + \cdots + x_{i}^n \right)^{1/n}. $$

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

$$S \leq \prod_{j=1}^{n} \left( \int \left( |A_{x}^1 p(\lambda_j|0) + A_{x}^2 p(\lambda_j|1) + \cdots + A_{x}^n p(\lambda_j|k-1) - A_{x}^0 p(\lambda_j|0)| \right) d\lambda_j \right)^{1/n}. $$

The following upper bound holds:

$$\frac{1}{2} \left( \int \left( |A_{x}^1 p(\lambda_j|0) + A_{x}^2 p(\lambda_j|1) + \cdots + A_{x}^n p(\lambda_j|k-1) - A_{x}^0 p(\lambda_j|0)| \right) d\lambda_j \right)^{1/n} \leq k - 1 \quad \square$$

Conclusion.— This paper has introduced a novel notion of non-classicality for joint quantum measurements. This notion took its cue from Bell’s theorem and the device-independent certification of entangled quantum states by stipulating a joint quantum measurement to be non-classical if the correlations generated by performing it on local preparations are inconsistent with an underlying classical causal model. A non-linear inequality was then derived as a witness for this inconsistency: a violation entails non-classicality. This inequality bounded the classically generated correlations achievable with this causal model. In future work it would be interesting to investigate the corresponding bounds for LOCC, unentangled, and entangled measurements, as was done in the semi-device independent case by Ref.’s [5] [6].

Moreover, this approach was shown to provide a more fine-grained notion of non-classicality than the post-selection method of Ref. [8]. That is, there exists quantum measurements which admit a classical hidden variable model for each post-selected measurement outcome, but which are nevertheless non-classical and violate the inequality from Result 1. It would be interesting to determine if a quantum protocol exhibiting an information-theoretic advantage due to this discrepancy existed. That is, can an agent with access to the entire collection of correlations generated by a quantum measurement gain an advantage over an agent who only has access to a post-selected subset of those correlations?

In future work, connections between the notion of non-classicality introduced here and that of contextuality discussed in Ref. [11] will be explored.

Acknowledgements.— CML thanks Dan Browne for helpful discussions and J. J. Barry for encouragement. This work was supported by the EPSRC through the UCL Doctoral Prize Fellowship.

[1] N. Sangouard, C. Simon, H. De Riedmatten, and N. Gisin, “Quantum repeaters based on atomic ensembles and linear optics,” Reviews of Modern Physics, vol. 83, no. 1, p. 33, 2011.

[2] C. Branciard, N. Gisin, and S. Pironio, “Characterizing the nonlocal correlations created via entanglement swapping,” Phys. Rev. Lett., vol. 104, no. 170401, 2010.

[3] M. D. Mazurek, M. F. Pusey, R. Kunjwal, K. J. Resch,
and R. W. Spekkens, “An experimental test of non-contextuality without unphysical idealizations,” Nature communications, vol. 7, p. ncomms11780, 2016.

[4] C. M. Lee and M. J. Hoban, “Towards device-independent information processing on general quantum networks,” Physical Review Letters, vol. 120, no. 2, p. 020504, 2018.

[5] T. Vértesi and M. Navascués, “Certifying entangled measurements in known hilbert spaces,” Physical Review A, vol. 83, no. 6, p. 062112, 2011.

[6] A. Bennet, T. Vértesi, D. J. Saunders, N. Brunner, and G. Pryde, “Experimental semi-device-independent certification of entangled measurements,” Physical Review Letters, vol. 113, no. 8, p. 080405, 2014.

[7] J. Bowles, N. Brunner, and M. Pawłowski, “Testing dimension and nonclassicality in communication networks,” Physical Review A, vol. 92, no. 2, p. 022351, 2015.

[8] 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.

[9] 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.

[10] 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.

[11] E. G. Cavalcanti, “Classical causal models for bell and kochen-specker inequality violations require fine-tuning,” Phys. Rev. X, vol. 8, p. 021018, Apr 2018.

[12] J.-M. A. Allen, J. Barrett, D. C. Horsman, C. M. Lee, and R. W. Spekkens, “Quantum common causes and quantum causal models,” arXiv preprint arXiv:1609.09187, 2016.

[13] R. Chaves, J. B. Brask, and N. Brunner, “Device-independent tests of entropy,” Phys. Rev. Lett., vol. 115, no. 110501, 2015.

[14] R. Chaves, D. Cavalcanti, and L. Aolita, “Causal hierarchy of multipartite bell nonlocality,” arXiv preprint arXiv:1607.07666, 2016.

[15] J. B. Brask and R. Chaves, “Bell scenarios with communication,” Journal of Physics A: Mathematical and Theoretical, 2017.

[16] C. M. Lee and R. W. Spekkens, “Causal inference via algebraic geometry: Feasibility tests for functional causal structures with two binary observed variables,” Journal of Causal Inference (DOI: 10.1515/jci-2016-0013), arXiv preprint arXiv:1506.03880, 2017.

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

[18] 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.

[19] 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.

[20] A. Acín, N. Gisin, and B. Toner, “Grothendieck’s constant and local models for noisy entangled quantum states,” Physical Review A, vol. 73, no. 6, p. 062105, 2006.

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

[22] J. Kaniewski, “Self-testing of binary observables based on commutation,” Physical Review A, vol. 95, no. 6, p. 062323, 2017.

[23] C. M. Lee and J. H. Selby, “Deriving grover’s lower bound from simple physical principles,” New Journal of Physics, vol. 18, no. 093047, 2016.

[24] C. M. Lee and J. H. Selby, “A no-go theorem for theories that decohere to quantum mechanics,” arXiv preprint arXiv:1701.07449, 2017.

[25] C. M. Lee and M. J. Hoban, “Bounds on the power of proofs and advice in general physical theories,” Proc. R. Soc. A, vol. 472, no. 2190, p. 20160076, 2016.

[26] C. M. Lee and M. J. Hoban, “The information content of systems in general physical theories,” arXiv preprint arXiv:1606.06801, EPTCS 214, 2016, pp. 22-28, 2016.

[27] Note that the work of Ref. [22] gave a method to device-independently certify the presence of quantum measurements acting on single systems. In this work, the case of quantum measurements acting jointly on multipartite systems is considered.

[28] For connections between related notions of causality and quantum information, see [24][25]

[29] Here the faithfulness condition is being assumed, see [10][11] for a discussion.