ALGEBRAIC PARADOXES IN ADAPTIVE QUANTUM COMPUTATION

SAMSON ABRAMSKY, RUI SOARES BARBOSA, CARMEN CONSTANTIN, AND MARTTI KARVONEN

Abstract. Measurement-based quantum computation (MBQC) is a universal model of quantum computation whose full power requires adaptivity. Contextuality is known to power quantum advantage in MBQC, yet it has resisted algebraic analysis in the adaptive setting.

We show that if an adaptive $\mathbb{Z}_2$-linear measurement-based quantum computing protocol deterministically computes a non-affine Boolean function, then the underlying quantum resource satisfies an inconsistent set of linear equations. This witnesses an algebraic form of strong contextuality generalising Mermin's All-versus-Nothing arguments. Such algebraic contextuality can be detected cohomologically, resolving an open question posed by Raussendorf, who had established cohomological witnesses of contextuality for non-adaptive protocols, but left the adaptive case open. We prove this result constructively: we model adaptive measurement protocols as ordinary measurements on a larger scenario of tree-like measurements, and explicitly build the inconsistent equations inductively.

Contents

1. Introduction

If quantum computers are to outperform their classical counterparts, their additional power must stem from observable behaviour that lies beyond the reach of classical systems. Contextuality – the impossibility of reconciling quantum measurement statistics with pre-existing outcome values – stands out as such a provably nonclassical phenomenon [20, 41] and has been shown to enable quantum advantage across multiple models of computation and information-processing tasks [11, 22, 35, 40, 52]. Closer to the metal, contextuality is being explored for benchmarking quantum hardware [31], as the constraints it imposes enable error detection beyond standard randomised benchmarking. At the heart of contextuality lies a logical tension – a "paradox": each empirically-accessible set of compatible measurements (*context*) provides a consistent classical snapshot, yet attempting to combine these into a single global description leads to a contradiction. This local-to-global obstruction is neatly captured in the language of sheaf theory [12].

Particularly elegant are purely algebraic witnesses of contextuality – known as All-versus-Nothing (AvN) arguments – in which a system of linear equations establishes such a logical contradiction: quantum measurement outcomes satisfy each equation locally, yet no global classical value assignment can satisfy them all. Classic examples of such algebraic contextuality include the GHZ paradox [32] and the Peres–Mermin magic square [47, 49], closely related to the parity-based games studied by Arkhipov [15] and more generally quantum solutions to linear constraint systems in non-local games [26, 59]. Their rigid algebraic form makes such arguments particularly tractable: detecting inconsistency reduces to linear algebra rather than general constraint satisfaction, an advantage that becomes substantial at scale, and all the more relevant given growing interest in many-body contextuality [34, 62].

Measurement-based quantum computation (MBQC) [23,55] is a standard universal model of quantum computation, equivalent in power to the circuit model and actively pursued in optical and ion-trap architectures [42, 51, 61].[1] In this paradigm, a classical control computer orchestrates the computation by performing successive local measurements on a multi-qubit resource state and processing the observed outcomes, treating the quantum system as a black box that responds to measurement queries. This black-box perspective makes MBQC a particularly natural framework for analysing quantum resources through an operational lens, and it is indeed standard in the study of contextuality in MBQC [14, 52].

Crucially in MBQC, the choice of later measurement settings may depend on prior outcomes, a feature known as *adaptivity*. This feed-forward capability is essential for quantum universality. Yet many foundational results on MBQC have been confined to the non-adaptive setting. The limitation runs deep: key algebraic and topological tools for analysing contextuality – including AvN arguments and cohomological methods – have hitherto only been applied in non-adaptive, "flat" measurement scenarios. Raussendorf [53, 54] articulated an instance of this gap as an explicit open question: can cohomological witnesses for contextuality in MBQC be extended to adaptive protocols? This matters beyond MBQC: quoting [56], 'a web of cohomological facts relates quantum error correction, measurement-based quantum computation, symmetry protected topological order, and contextuality'. Understanding how adaptivity fits into this web is therefore of broad foundational importance.

We address a related but stronger question: can the contextuality required for adaptive MBQC always be witnessed by an explicit algebraic contradiction – an AvN argument? The answer is yes.

[1] See the recent comparative survey [43], which discusses growing interest in this paradigm for hardware implementations.

To state our contribution precisely, we consider $\mathbb{Z}_2$-linear MBQC, where the classical control computer, including pre- and post-processing and adaptive control, is restricted to linear operations modulo two. Raussendorf [52] showed that if such a (possibly adaptive) $\mathbb{Z}_2$-MBQC protocol computes a nonaffine Boolean function deterministically, then the protocol must exploit strong contextuality in its resource state. Our main contribution strengthens Raussendorf's result by showing that the requisite strong contextuality is algebraic in nature: it can always be witnessed via an AvN argument. Our proof is constructive: given an adaptive protocol, we explicitly build the inconsistent system of equations from its tree structure, rather than inferring its existence from properties of the resource state. This also implies the contextuality is cohomologically nontrivial, thus resolving Raussendorf's open question [53, 54]: cohomological witnesses for contextuality do extend to adaptive protocols. We establish this in both the group-cohomological and Čech-cohomological frameworks for contextuality.

A key technical obstacle is that existing tools, including the very definition of AvN arguments [8, 10, 30] and cohomological witnesses for contextuality [10, 13, 53, 54], are formulated only for "flat" measurement scenarios, making it unclear how to even state, let alone prove, adaptive analogues. Rather than developing an entirely new theory for adaptive protocols, we employ a systematic "flattening" construction that models adaptive protocols as non-adaptive protocols on appropriately enlarged measurement scenarios. Drawing on techniques from [9], where adaptivity is captured via a (co)Kleisli construction, this perspective allows us to apply the full arsenal of existing tools directly in the adaptive setting, while making the inductive structure of the AvN arguments explicit; see also [57, 58] for similar flattening constructions in related settings. We emphasise that this "flattening" does not eliminate adaptivity: measurements in the enlarged scenario are themselves adaptive operations on the original scenario, and would be implemented as such in practice. The point is not to reduce adaptive to non-adaptive computation, but to bring mathematical tools to bear on the adaptive case. Beyond the immediate result, our flattening construction provides a general pathway for extending results from the non-adaptive to the adaptive setting, enabling further study of adaptivity in MBQC.

An AvN argument: the GHZ paradox. To illustrate AvN arguments and their relationship to measurement-based computation, we start by presenting a canonical example.

Consider a scenario in which a physical system is distributed over three sites. At each of these, one may perform a local measurement and record the observed outcome. There are two measurement settings available at each site, corresponding to two measurement procedures one may perform locally on the system. We label these x_0, x_1 for the first site, y_0, y_1 for the second, and z_0, z_1 for the third. Each of these measurements is dichotomic, yielding an outcome in $\mathbb{Z}_2$. Importantly, we view $\mathbb{Z}_2$ as a ring, not merely as the bare set $\{0, 1\}$; this algebraic structure is essential for expressing the kind of argument that follows.

In each run of the experiment, a measurement is performed at each site and the corresponding outcome is recorded. Over many runs with varying choices of measurement settings, one collects measurement statistics that characterise the observable behaviour. These data constitute an *empirical model*: for each choice of measurement settings at the three sites – a *context* – it specifies a probability distribution over the joint outcomes. Crucially, we treat the system as a black box that responds to measurement queries, without presupposing any particular physical theory by which outcomes are produced: the model captures only what is observable. Table 1 shows one particular empirical model for this scenario: four contexts are shown; for the remaining four contexts, the joint outcomes are uniformly

TABLE 1. The GHZ empirical model on B_3

	000	001	010	011	100	101	110	111
$x_0\, y_0\, z_0$	1/4	0	0	1/4	0	1/4	1/4	0
$x_0\, y_1\, z_1$	0	1/4	1/4	0	1/4	0	0	1/4
$x_1\, y_0\, z_1$	0	1/4	1/4	0	1/4	0	0	1/4
$x_1\, y_1\, z_0$	0	1/4	1/4	0	1/4	0	0	1/4

distributed. Observe that the probability distributions over different contexts agree on their marginals: the choice of measurement setting at one site does not affect the probability distribution over outcomes at the other sites. This property is known as *no-signalling* or *no-disturbance*.

Classically, one thinks of measurements as revealing pre-existing properties of the physical system under observation. More precisely, on each fully specified state (even if empirically inaccessible, i.e. a hidden variable), we expect a function g that assigns a definite outcome to each measurement, independently of what other measurements are performed at the other sites; the observable behaviour then arises from a probability distribution over such deterministic states. The no-signalling property lends plausibility to this view.

However, examining the empirical model reveals a striking pattern. In each of the four contexts shown, the observed joint outcomes always satisfy a system of linear equations over $\mathbb{Z}_2$:

$$\begin{aligned} x_0 + y_0 + z_0 &= 0 \qquad & x_0 + y_1 + z_1 &= 1 \\ x_1 + y_0 + z_1 &= 1 \qquad & x_1 + y_1 + z_0 &= 1 \end{aligned} \tag{1}$$

This system of equations is readily seen to be inconsistent: adding all four equations, each measurement appears twice on the left-hand side and thus cancels, yielding $0 = 1$, a contradiction. Therefore, no assignment $g : \{x_0, x_1, y_0, y_1, z_0, z_1\} \to \mathbb{Z}_2$ of definite values to the six measurements is consistent with the observed outcomes across the four contexts. This establishes that the empirical model is *strongly contextual* (see Section 2 for a formal definition), and algebraic contradictions of this form are known as *All-versus-Nothing* (AvN) arguments [10].

Remarkably, quantum systems exhibit precisely such counterintuitive behaviour. The empirical model in Table 1 can be realised using a three-qubit system initialised in the GHZ state $|\psi_{\mathsf{GHZ}}\rangle = 1/\sqrt{2}\,(|000\rangle + |111\rangle)$. The measurements x_0, x_1 (resp. y_0, y_1 and z_0, z_1) at each site correspond to measuring the corresponding qubit in the $X = \{|+\rangle, |-\rangle\}$ or $Y = \{|+i\rangle, |-i\rangle\}$ bases, with the outcomes labelled 0 and 1.[2]

But we stress that such details about the quantum realisation are not essential for what follows: the key takeaway is that this seemingly paradoxical empirical model is physically realisable. Throughout, we work at the level of observable behaviour (the statistics of measurement outcomes and the algebraic constraints they satisfy) rather than the Hilbert-space formalism, treating the quantum system as a black box. This operational, theory-independent approach lays bare the ingredients driving computational advantage, independently of any particular physical model.

Computing with AvN paradoxes: the Anders–Browne construction. Anders and Browne [14] showed how the GHZ empirical model can be used as a computational resource to enhance the power of a limited classical computer. Specifically, they showed that a parity computer, restricted to performing affine operations over $\mathbb{Z}_2$, can compute non-affine Boolean functions when given access to this empirical model. Such access takes the form of *measurement-based computation*: the

[2]The outcomes 0 and 1 correspond to the $+1$ and -1 eigenvalues of the respective Pauli operators via the group homomorphism $r \mapsto (-1)^r$.

classical controller queries the model by choosing measurement settings and obtaining corresponding outcomes.

We present the construction for the (non-affine) OR function that computes the disjunction of two bits, $\vee : \mathbb{Z}_2^2 \to \mathbb{Z}_2 :: (a, b) \mapsto a \vee b = a + b + a \cdot b$. Given input bits $a, b \in \mathbb{Z}_2$, the parity control computer first applies the linear preprocessing function $Q : \mathbb{Z}_2^2 \to \mathbb{Z}_2^3 :: (a, b) \mapsto (a, b, a + b)$. This encodes the input into a three-bit vector that determines the measurement settings for the three sites: the computer queries the resource by measuring the context $\{x_a, y_b, z_{a+b}\}$. The three measurement outcomes are then post-processed by a linear function $Z : \mathbb{Z}_2^3 \to \mathbb{Z}_2$ that sums them, producing a single output bit.

The fact that this protocol *deterministically* computes the OR function follows from the linear equations in eq. (1), which can be condensed as

$$x_a + y_b + z_{a+b} = a \vee b \qquad (a, b \in \mathbb{Z}_2). \tag{2}$$

Note that the empirical model can be regarded as a probabilistic function $e : \mathbb{Z}_2^3 \to \mathcal{D}(\mathbb{Z}_2^3)$, where $\mathcal{D}$ is the probability distribution monad. The action of this protocol is then the Kleisli composition $Z \bullet e \bullet Q$, which coincides with $\eta \circ \vee$, where η is the unit of the monad $\mathcal{D}$ embedding deterministic values as point distributions.

One can use this construction as a building block to compute arbitrary Boolean functions, as OR together with affine operations is functionally complete.[3] Each OR gate in a logical circuit is then implemented by this construction. Achieving this in general requires *adaptivity* or *feed-forward* controlled by the parity computer: later measurement settings must depend (affinely) on prior measurement outcomes, allowing one to compose OR gates sequentially.

The GHZ algebraic paradox underlies the computation of the OR function. This raises a natural question: for an arbitrary adaptive $\mathbb{Z}_2$-linear measurement-based computation, is there always such an AvN argument witnessing the contextuality being exploited as a computational resource? At first sight, this appears challenging: if the protocol has nontrivial adaptivity, then the description of its output (the left-hand side of eq. (2)) will not be a simple linear expression in measurements performed on the resource system. For a minimal example, consider the composition of two Anders–Browne constructions by feeding the output of the first as input to the second to implement a three-wise OR. The output is described by

$$x'_{x_a+y_b+z_{a+b}} + y'_c + z'_{x_a+y_b+z_{a+b}+c}$$

where the primed variables refer to measurements on the second copy of the GHZ model. With greater adaptive depth, such expressions become increasingly complex.

In this paper, we resolve this question affirmatively. The key insight is to consider AvN arguments over adaptive measurement protocols, which we treat as first-class measurements in their own right. This allows us to construct an AvN argument explicitly for any adaptive $\mathbb{Z}_2$-linear MBQC, building the inconsistent systems of equations inductively from the structure of the computation tree.

Roadmap. The remainder of the paper is organised as follows. Section 2 recalls the background on contextuality, in particular on $\mathbb{Z}_2$-linear AvN arguments. Section 3 introduces $\mathbb{Z}_2$-linear measurement protocols on a Bell scenario, makes precise what it means for an MBQC to compute a Boolean function deterministically, and states our main result (Theorem 3.6) together with a proof sketch. Section 4 works through three examples of increasing complexity, recovering the PR-box and GHZ paradoxes and then treating a genuinely adaptive depth-three protocol in which each step of the general argument has a concrete counterpart.

[3] In fact, [14] shows that access to polynomially-many GHZ states promotes the parity computer from (a subclass of) the complexity class $\oplus L$ to P.

The proof itself is set up in Section 5: it calls for a richer class of measurement protocols than the basic linear ones – protocols that can halt on undesired outcomes or measure several sites simultaneously – which are assembled into the scenario $\mathsf{MP}(B_{\mathcal{I}}, \mathbb{Z}_2)$; the proof strategy is then presented, reducing the main theorem to a single identity (Theorem 5.11) whose proof proceeds by induction (roughly on protocol depth) requiring three key lemmas. Section 6 proves the three lemmas underpinning this inductive argument, establishing in each case that protocol trees can be rearranged, condensed, or decomposed while preserving the sums at the heart of the argument.

Section 7 draws further consequences: most notably, it resolves Raussendorf's open question by extending cohomological witnesses of contextuality to the adaptive MBQC setting, in both the sheaf-cohomological (Theorem 7.1) and the group-cohomological frameworks, with the Čech witness propagating all the way back to the original scenario. Section 8 concludes with directions for further work.

2. Background on contextuality

We give a rapid précis of the basic notions from [12] and [10].

2.1. **Scenarios, empirical models, and contextuality.** A contextuality *scenario* is a structure $S := (X, \Sigma, O)$, where X is a finite set of measurement labels or *variables*; Σ is a family of subsets of X giving the sets of *compatible measurements* or *contexts*; and O is a set of measurement *outcomes*. The family Σ is required to be downward closed (subsets of compatible measurements are compatible) and to contain all singletons (any single measurement can be performed), i.e. Σ is an abstract simplicial complex on X.[4]

An important class of scenarios are the *Bell scenarios*, which arise in non-locality arguments. These are scenarios for which the set of measurements is partitioned as $X = \bigsqcup_{i \in \mathcal{I}} X_i$ and $C \in \Sigma$ iff $|C \cap X_i| \leq 1$ for all $i \in \mathcal{I}$. Thus $\mathcal{I}$ is a set of sites or agents (Alice, Bob, etc.), X_i are the measurements available at site i, and the contexts arise by choosing at most one measurement from each site.

The *event sheaf* over a scenario associates to each context $C \in \Sigma$ the set $\mathscr{E}(C) = O^C$ of joint outcomes for C, with restriction maps $\rho_C^{C'} : \mathscr{E}(C') \to \mathscr{E}(C)$ when $C \subseteq C'$ given by the obvious projection. This is a sheaf $\mathscr{E} : P^{\mathsf{op}} \to \mathbf{Set}$ defined on the poset category $P = (\mathscr{P}(X), \subseteq)$. We write $h|_C$ for $\rho_C^{C'}(h)$ when $h \in \mathscr{E}(C')$.

An *empirical model* on a contextuality scenario is a family $e = \{e_C\}_{C \in \Sigma}$, where $e_C \in \mathcal{D}\mathscr{E}(C)$ is a distribution over the joint outcomes for context C. We write $e : S$ to express that e is an empirical model on the scenario S. Here $\mathcal{D}$ is the distributions functor that assigns to each set X the set of discrete (finitely supported) distributions on X. This yields a presheaf $\mathcal{D}\mathscr{E} : P^{\mathsf{op}} \to \mathbf{Set}$.

To spell this out more explicitly, let R be a commutative semiring of "weights". An R-distribution on X is a function $d : X \to R$ with finite support $\mathsf{supp}(d) := \{x \in X \mid d(x) \neq 0\}$ which satisfies the normalisation condition $\sum_{x \in X} d(x) = 1$. The distribution functor $\mathcal{D}_R$, parameterised on R, sends X to the set of all R-distributions on X. The restriction maps now correspond to marginalisation. Note that, if we take R to be the non-negative reals, we get the standard discrete probability distributions, while if we take R to be the booleans, we get "possibility distributions", i.e. non-empty subsets, where non-emptiness corresponds to the normalisation condition. This allows a unified treatment of probabilistic and "probability-free" non-locality and contextuality arguments.

[4] More generally, one may make O into a dependent type, $O = \{O_x\}_{x \in X}$, allowing the set of outcomes to depend on the measurement. This is the setting of [9, 17, 37], though it will not be needed here.

Empirical models are required to satisfy the compatibility property: for all $C, C' \in \Sigma$, $e_C|_{C\cap C'} = e_{C'}|_{C\cap C'}$. This says that the distributions over contexts have consistent marginals on their overlaps. This is known as the "no-disturbance" condition, and specialises on Bell scenarios to the standard "no-signalling" condition.

Whereas the event sheaf $\mathscr{E}$ satisfies the sheaf property, so that compatible families over $\mathscr{E}$ can always be glued together into a unique global section (an assignment of outcomes to all the measurements in the scenario), the presheaf $\mathcal{D}_R\mathscr{E}$ is *not* a sheaf. Counter-examples to gluing correspond exactly to contextuality.[5] More explicitly, an empirical model e is *non-contextual* iff there is a "global distribution" $d \in \mathcal{D}_R\mathscr{E}(X)$ over global outcome assignments such that $d|_C = e_C$ for all $C \in \Sigma$, i.e. iff a gluing exists for the compatible family. Such a global distribution can be seen as a canonical form of "hidden-variable model" or "ontological model"; see [12].

In the case that R is the non-negative reals, this yields the notion of *probabilistic contextuality*, covering the usual examples from non-locality, typified by Bell's theorem [19]. The case where R is the booleans was dubbed *logical contextuality* in [12]. This covers examples of "probability-free" non-locality proofs such as the Hardy paradox [33], as observed in [12]. Note that logical contextuality is strictly stronger than probabilistic contextuality. Indeed, a semiring homomorphism $R \to S$ induces a natural transformation $\mathcal{D}_R \Rightarrow \mathcal{D}_S$, which in turn lifts to a mapping from empirical models of R-distributions to empirical models of S-distributions. Evidently, there is such a homomorphism from the non-negative reals to the booleans, which sends positive elements to 1.[6]

There is a yet stronger, and in a suitable sense maximally strong, form of contextuality, known as strong contextuality. An empirical model e is *strongly contextual* iff there is no deterministic global assignment $g \in \mathscr{E}(X)$ such that $g|_C \in \mathsf{supp}(e_C)$ for all $C \in \Sigma$. To compare this with logical contextuality, note that the latter holds iff for *some* local section $h \in \mathsf{supp}(e_C)$, there is no global assignment $g \in \mathscr{E}(X)$ extending h such that $g|_C \in \mathsf{supp}(e_C)$ for all $C \in \Sigma$. By contrast, strong contextuality holds iff this is the case for *all* local sections. Thus we obtain a strict hierarchy of notions of contextuality:

$$\text{probabilistic} \;<\; \text{logical} \;<\; \text{strong}.$$

2.2. **AvN contextuality.** Now we assume that there is algebraic structure on the outcomes, specifically $O = \mathbb{Z}_2$, where we regard $\mathbb{Z}_2$ as an abelian group.[7] This algebraic structure can be used to define the $\mathbb{Z}_2$-*linear theory* of an empirical model e. Given a context C, a $\mathbb{Z}_2$-linear equation over C has the form (D, a), where $D = \{x_1, \ldots, x_d\} \subseteq C$, and $a \in \mathbb{Z}_2$. A section $h \in \mathscr{E}(C)$ satisfies such an equation if $\sum_{x\in D} h(x) = a$. We write this as $h \vDash \sum_{x\in D} x = a$, and for a model e, we write $e \vDash \sum_{x\in D} x = a$ iff $h \vDash \sum_{x\in D} x = a$ for all h in the support of e_C. The $\mathbb{Z}_2$-linear theory of an empirical model e is the union over all contexts C of the set of $\mathbb{Z}_2$-linear equations with variables in C that are satisfied by all the sections in the support of e_C.

This leads in turn to the notion of *All-versus-Nothing contextuality* [10]. (The term "All versus Nothing" was first used in [46].) An empirical model with outcomes $\mathbb{Z}_2$ is $\mathsf{AvN}_{\mathbb{Z}_2}$-contextual if its linear theory is inconsistent, i.e. there is no global assignment of values in $\mathbb{Z}_2$ to all the variables which satisfies all the equations in the linear theory. Writing $\vDash^*$ for the closure of $\vDash$ under classical reasoning, such a model e is $\mathsf{AvN}_{\mathbb{Z}_2}$ iff $e \vDash^* 0 = 1$. This is a more structured version of the theme of

[5]The fact that $\mathscr{E}$ is a sheaf says that no-signalling deterministic models are non-contextual.

[6]More generally, there is such a homomorphism from any zero-sum free semiring. Note, however, that there is no such homomorphism from the reals, or indeed any ring.

[7]More generally, O need only be a module over a ring for the linear theory to be defined; any abelian group is a module over $\mathbb{Z}$; cf. [10].

"local consistency, global inconsistency" [3]. AvN contextuality is strictly stronger than strong contextuality [10].

3. $\mathbb{Z}_2$-linear MBQC and the main result

In this section, we introduce the model of computation studied in this paper: measurement-based computation with $\mathbb{Z}_2$-linear classical control. We define $\mathbb{Z}_2$-linear measurement protocols over a Bell scenario $B_{\mathcal{I}}$ (Section 3.1) and how such a protocol, acting on an empirical model, computes a Boolean function (Section 3.2). We then state our main result – that deterministically computing a non-affine function requires AvN contextuality – and sketch an informal proof outline (Section 3.3).

3.1. Linear protocols.

For a finite set of sites $\mathcal{I}$, let $B_{\mathcal{I}}$ be the Bell scenario whose measurements are given by $X = \bigsqcup_{i \in \mathcal{I}} \mathbb{Z}_2$ with outcomes in $\mathbb{Z}_2$. When useful, we denote the two measurements available at site i by $x_{i,0}$ and $x_{i,1}$, although we often represent arbitrary measurements simply by x, y, z, or by $x_1, x_2, \ldots$ when referring to several at once. We write $s : X \to \mathcal{I}$ for the map sending a measurement to its site.

Given $U \subseteq V$, we write $\pi_U : \mathbb{Z}_2^V \to \mathbb{Z}_2^U$ for the projection and $\iota_V : \mathbb{Z}_2^U \to \mathbb{Z}_2^V$ for the inclusion, leaving the domain implicit. When U is a singleton $\{i\}$, we write π_i instead of $\pi_{\{i\}}$.

Roughly speaking, $\mathbb{Z}_2$-linear MBQC protocols operate on the scenario $B_{\mathcal{I}}$ for some set of sites $\mathcal{I}$, going through (some of) the sites in some fixed order, with later measurement choices depending linearly on past outcomes. We represent such a protocol as a tree

$$\begin{array}{ccc} & x & \\ 0 \swarrow & & \searrow 1 \\ q & & v.q \end{array}$$

where x is the initial measurement, q is the continuation protocol, and v is a vector specifying how to update the measurement settings in the continuation upon obtaining outcome 1 for the measurement x. The operational interpretation of such a tree is then a protocol that starts by measuring x, and then proceeds to either q or $v.q$ depending on the observed outcome.

This suggests an inductive definition, where one represents such a tree by the triple (x, v, q). The following definition, our version of [52, Definition 1], makes this precise.

Definition 3.1. For a set of sites $U \subseteq \mathcal{I}$, we define the set MP_U of *linear measurement protocols* that measure the sites U. We overload notation by writing s for the function sending a protocol p to the set $s(p) \subseteq \mathcal{I}$ of sites it measures. These notions are defined inductively as follows:

- We set $\mathsf{MP}_\emptyset := \{()\}$, i.e. there is a unique measurement protocol measuring no sites, namely the empty protocol $()$.
- If $q \in \mathsf{MP}_U$ is a linear measurement protocol, v is a vector in $\mathbb{Z}_2^U$ and x is a measurement at a site not measured by q, i.e. $s(x) \notin U$, then $p := (x, v, q)$ is a linear measurement protocol and $s(p) = \{s(x)\} \cup U$, so that $p \in \mathsf{MP}_{\{s(x)\} \cup U}$.

The above definition encodes the data needed to describe a linear measurement protocol. However, we also need to specify how suitably-sized vectors act on measurement protocols, so that the continuation $v.q$ is well defined.

Definition 3.2. The regular action of $\mathbb{Z}_2$ on itself induces a group action of $\mathbb{Z}_2$ on the set of measurements $\bigsqcup_{i \in \mathcal{I}} \mathbb{Z}_2$ of $B_{\mathcal{I}}$: explicitly, this action is given by $v.x_{i,a} = x_{i,a+v}$ for $v \in \mathbb{Z}_2$. Note that this action preserves the site the measurement lives on.

We define an action of $\mathbb{Z}_2^U$ on MP_U by induction on the structure of the protocol. The action of $\mathbb{Z}_2^\emptyset \cong \{0\}$ is uniquely defined on $\mathsf{MP}_\emptyset$. The action of $w \in \mathbb{Z}_2^{s(p)}$ on $p = (x, v, q)$ is defined as $w.p := \big(\pi_{s(x)}(w).x, v, \pi_{s(q)}(w).q\big)$.

We call these linear measurement protocols since the adaptivity in them is $\mathbb{Z}_2$-linear. While we later generalise this definition somewhat (allowing for some amount of non-linearity), we do not define measurement protocols in the full generality used in [9].

The set $s(p)$ is linearly ordered by the protocol p, with the order defined inductively for (x, v, q) by setting $s(x) < s(q)$, i.e. $s(x)$ is below every element of $s(q)$, followed by the order on $s(q)$ induced by q. We will use this *intrinsic order* on $s(p)$ throughout.

We often represent linear measurement protocols as trees as in the beginning of this section. We use such visual notations more expressively in the sequel. For example, we write

$$\begin{array}{ccc} & p & \\ & | & \\ & x & \\ 0 \swarrow & & \searrow 1 \\ q & & v.q \end{array}$$

to mean a protocol that performs the protocol p, then proceeds to x and then to q or $v.q$ depending on the outcome of x. However, this notation suppresses the fact that the outcome of p may influence the measurement settings of x and q, isolating just the action of x on the sequel while reminding the reader that there indeed was a past prior to x.

Definition 3.3. Given a linear measurement protocol p, we define its *adaptivity matrix* T_p (or simply T), an $s(p) \times s(p)$ matrix over $\mathbb{Z}_2$, by induction on the structure of p. If p is the empty protocol, then T_p is the unique 0×0 matrix. If $p = (x, v, q)$, we define

$$T_p := \begin{array}{c} \\ s(x) \\ s(q) \end{array} \begin{array}{c} \begin{array}{cc} s(x) & s(q) \end{array} \\ \begin{bmatrix} 0 & 0 \\ v & T_q \end{bmatrix} \end{array}$$

Note that when $|s(p)| = 1$, the definition above gives $T_p = \big[0\big]$. The operational meaning of the matrix is that the ith column specifies how the measurement outcome at site i affects other measurement settings. By construction, the matrix T_p is strictly lower triangular when one orders $s(p)$ in the intrinsic order. This constraint ensures that a measurement outcome can only affect the settings of (strictly) later measurements.

Remark 3.4. The matrix T_p does not fix the protocol p uniquely. First, the same matrix might be lower-triangular with respect to many different orderings. Any such ordering is the intrinsic order of some linear protocol with the same matrix. However, even if we fix the linear order, we are still missing information. This is because the indexing set only keeps track of the measurement site but not of the measurement setting. This missing information can be provided by a vector $w_p \in \mathbb{Z}_2^{s(p)}$ specifying the default measurement settings (i.e. those used when all prior outcomes are 0). Then one can represent p uniquely by $s(p)$ equipped with the intrinsic order of p, T_p, and w_p. The action $(v, p) \mapsto v.p$ translates in this representation to $v.(T_p, w_p) = (T_p, v + w_p)$. In [52], the default measurement setting is (in a sense) fixed to be 0, which corresponds to working with linear rather than affine functions of prior outcomes, so this issue does not arise.

3.2. **Computing with linear protocols.** We discuss how such protocols, given access to an empirical model $e : B_{\mathcal{I}}$, can be used to compute a Boolean function $f : \mathbb{Z}_2^l \to \mathbb{Z}_2^o$. To do so, we first fix a base protocol p and a linear input encoding $Q : \mathbb{Z}_2^l \to \mathbb{Z}_2^{s(p)}$: the idea is that given an input $w \in \mathbb{Z}_2^l$, we run the protocol $Q(w).p$ on our empirical model, obtaining a vector $\mathbb{Z}_2^{s(p)}$ of measurement outcomes (an outcome for each measurement performed while running p). This outcome is then post-processed by another linear map $Z : \mathbb{Z}_2^{s(p)} \to \mathbb{Z}_2^o$ to produce the overall output of the computation. This whole computation is then captured by fixing the empirical model $e : B_{\mathcal{I}}$, the base protocol $p \in \mathsf{MP}$, and the pre- and post-processing maps Q and Z. We say that the MBQC (e, p, Q, Z) computes f deterministically if, regardless of the (in general probabilistic) outcomes obtained from e, every input w in $\mathbb{Z}_2^l$ is always mapped to the correct output $f(w) \in \mathbb{Z}_2^o$.

This can be understood more formally as follows: once the protocol p has been fixed, the empirical model induces a stochastic map $e_p : \mathbb{Z}_2^{s(p)} \to \mathbb{Z}_2^{s(p)}$ by sending a vector $v \in \mathbb{Z}_2^{s(p)}$ to the joint outcome in $\mathbb{Z}_2^{s(p)}$ obtained when running $v.p$. Then, (e, p, Q, Z) computing f deterministically boils down to commutativity of

$$\begin{array}{ccc} \mathbb{Z}_2^l & \xrightarrow{\quad f \quad} & \mathbb{Z}_2^o \\ Q\downarrow & & \uparrow Z \\ \mathbb{Z}_2^{s(p)} & \xrightarrow[e_p]{} & \mathbb{Z}_2^{s(p)} \end{array}$$

Now, roughly speaking, [52, Theorem 2] states that if an MBQC (e, p, Q, Z) computes a non-affine function f, then e is strongly contextual. [8] We strengthen this to show that e is AvN-contextual when extended to a suitable scenario of measurement protocols. We first make some further observations and reductions.

As a function $f : \mathbb{Z}_2^l \to \mathbb{Z}_2^o$ is affine iff each of its coordinate functions is, it suffices to study functions $f : \mathbb{Z}_2^l \to \mathbb{Z}_2$ returning a single output bit. We now show that we can further reduce to two input bits. We say that a function $f : \mathbb{Z}_2^l \to \mathbb{Z}_2$ has *even (odd) weight* if the preimage of 1 under f has even (odd) cardinality, i.e. $\sum_{v \in \mathbb{Z}_2^l} f(v) = 0$ $(= 1)$. While every affine function $\mathbb{Z}_2^l \to \mathbb{Z}_2$ has even weight when $l \geq 2$, the converse does not hold in general. However, when $l = 2$, all functions with even weight are in fact affine. Moreover, since affineness is a condition stated with two variables, it can be tested on each two-dimensional subspace, so that a function is affine iff its restrictions to two-dimensional subspaces are affine. We thus obtain the following.

Lemma 3.5. *A function $f : \mathbb{Z}_2^l \to \mathbb{Z}_2$ is affine if and only if for every two-dimensional subspace $V \subseteq \mathbb{Z}_2^l$ the restriction $f|_V$ has even weight.*

Consider an MBQC (e, p, Q, Z) that deterministically computes a non-affine function $f : \mathbb{Z}_2^l \to \mathbb{Z}_2$. As f is not affine, by Lemma 3.5 some restriction to a two-dimensional subspace has odd weight. We can then precompose both f and Q with the inclusion $\mathbb{Z}_2^2 \hookrightarrow \mathbb{Z}_2^l$ of this subspace, resulting in a function $\mathbb{Z}_2^2 \to \mathbb{Z}_2$ with odd weight computed deterministically by the resulting measurement protocol. Thus we have reduced to the case when $l = 2$.

3.3. **Main result and proof sketch.** To get an AvN argument in the sense of [10], we need a measurement scenario where all measurements have outcomes valued in a single ring (or module). The scenario of all linear measurement protocols on $B_{\mathcal{I}}$ will not do, since different measurement protocols can return bit-strings of

[8] More precisely, Raussendorf shows that e is strongly contextual even when restricted to those joint measurements that could occur when running p.

different lengths. This can be fixed by incorporating the post-processing Z into the measurements of the scenario. However, in order to obtain an AvN argument, we also need to generalise our notion of a measurement protocol. A crucial modification is to include certain subprotocols of linear measurement protocols as measurements in the scenario: this allows us to decompose protocols into sums of smaller protocols that may in turn satisfy further linear equations. The resulting scenario is defined later in Definition 5.7 and denoted $\mathsf{MP}(B_{\mathcal{I}}, \mathbb{Z}_2)$: any model $e : B_{\mathcal{I}}$ extends to a model $\mathsf{MP}(e, \mathbb{Z}_2) : \mathsf{MP}(B_{\mathcal{I}}, \mathbb{Z}_2)$.

Let us state our main result:

Theorem 3.6. *If an MBQC (e, p, Q, Z) deterministically computes a non-affine function $f : \mathbb{Z}_2^l \to \mathbb{Z}_2^o$ then the model $\mathsf{MP}(e, \mathbb{Z}_2)$ on $\mathsf{MP}(B_{\mathcal{I}}, \mathbb{Z}_2)$ induced by e has an inconsistent $\mathbb{Z}_2$-linear theory.*

The basic premise of the proof is simple. We first reduce to a function $\mathbb{Z}_2^l \to \mathbb{Z}_2$ by picking a non-affine coordinate function of f, and then restrict to a function $\hat{f} : \mathbb{Z}_2^2 \to \mathbb{Z}_2$ with odd weight by Lemma 3.5.

Then, an MBQC (e, p, Q, Z) deterministically computes a function $\hat{f} : \mathbb{Z}_2^2 \to \mathbb{Z}_2$ iff

$$\mathsf{MP}(e, \mathbb{Z}_2) \vDash Q(w).p; Z = \hat{f}(w) \quad \text{for all } w \in \mathbb{Z}_2^2, \tag{3}$$

where we (temporarily) write $Q(w).p; Z$ for the measurement in $\mathsf{MP}(B_{\mathcal{I}}, \mathbb{Z}_2)$ which post-processes $Q(w).p$ by Z. When $\hat{f}$ has odd weight, adding these four equations yields 1 on the right-hand side, so that

$$\mathsf{MP}(e, \mathbb{Z}_2) \vDash^* \sum_{w \in \mathbb{Z}_2^2} Q(w).p; Z = 1.$$

Hence, to prove Theorem 3.6, it suffices to show that

$$\mathsf{MP}(e, \mathbb{Z}_2) \vDash^* \sum_{w \in \mathbb{Z}_2^2} Q(w).p; Z = 0 \tag{4}$$

and that's where most of the effort lies. In fact, this equation holds in the $\mathbb{Z}_2$-linear theory of $\mathsf{MP}(e, \mathbb{Z}_2)$ for *any* $e : B_{\mathcal{I}}$ (see Proposition 5.9). As a result, we will slightly abuse notation and often omit $\mathsf{MP}(e, \mathbb{Z}_2) \vDash$ and $\mathsf{MP}(e, \mathbb{Z}_2) \vDash^*$ from the equations we state.

We close this section by sketching the overall structure of the inductive argument for eq. (4). However, the reader should be aware that the sketch below is morally but not formally correct: roughly speaking, making it fully rigorous requires proving a more general result with a sufficiently strong induction hypothesis making everything work.

Proof sketch. The proof proceeds by induction on $n := |s(p)|$, the number of sites in p.

If $n = 1$, then $Q : \mathbb{Z}_2^2 \to \mathbb{Z}_2^1$ is not injective, since its domain has strictly larger dimension than its codomain. As a result, each summand in eq. (4) appears an even number of times and hence cancels out, so eq. (4) holds trivially.

For $n > 1$, the result follows from the following three lemmas:

(1) If $\operatorname{rank} T_p = n$, we show in Lemma 6.5 that there is another measurement protocol q such that

$$\sum_{w \in \mathbb{Z}_2^2} Q(w).p; Z \;=\; \sum_{w \in \mathbb{Z}_2^2} Q(w).q; Z$$

and $\operatorname{rank} T_q < n$. In other words, without loss of generality, we can assume that T_p is not of maximal rank.

(2) If $\operatorname{rank} T_p < n$, we show in Lemma 6.7 that we can condense p to have strictly fewer levels without affecting the value of the sum $\sum_{w \in \mathbb{Z}_2^2} Q(w).p; Z$. The cost of this is that we need to allow *higher-arity nodes*, which, rather than measuring a single measurement, measure the total parity $\sum x_i$ of a set of measurements in one step.
(3) Given a measurement protocol, we show in Lemma 6.10 that we can systematically replace all higher-arity branching nodes by sums of protocols with unary nodes. This will let us write the overall sum as a sum of smaller sums $\sum_w Q_i(w).p_i; Z_i$, where each p_i has strictly fewer than n sites, therefore reducing to the induction hypothesis. However, the cost to pay is that we need to allow measurement protocols that have *multiplicative nodes* which, instead of branching into two continuations, continue only if a certain fixed outcome is observed, otherwise returning 0. While such nodes allow one to implement non-affine functions, it turns out that each level with such nodes increases the input dimension by 1 as well, and the result remains true provided the input dimension is sufficiently large compared to the number of such levels. □

4. Illustrative examples

In this section, we give examples of AvN arguments that arise in this manner. We are not as precise here as in the formal treatment of the following sections, since the goal is to illustrate rather than to prove carefully.

Let $|\mathcal{I}| = 2$, and denote by x_0, x_1 the measurements at the first site and by y_0, y_1 those at the second site. Let the base protocol p just measure x_0 and then y_0 regardless of the outcome, with the post-processing Z returning the sum of the two outcome bits, and set the pre-processing Q to be identity on $\mathbb{Z}_2^2$, i.e. the first input bit determines the measurement choice for x and the second input bit the measurement choice for y. If $e : B_{\mathcal{I}}$ is an empirical model such that (e, p, Q, Z) deterministically computes the AND function, equation (3) specialises to the following system of equations:

$$\begin{aligned} x_0 + y_0 &= 0 & \qquad x_0 + y_1 &= 0 \\ x_1 + y_0 &= 0 & \qquad x_1 + y_1 &= 1 \end{aligned}$$

The only empirical model satisfying these equations is the Popescu–Rohrlich (PR) box [50], and thus Theorem 3.6 recovers the usual AvN argument of the PR box (obtained by summing the four equations) as a special case.

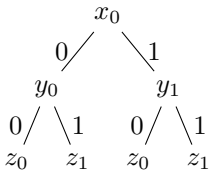
Now, let $|\mathcal{I}| = 3$, denoting the measurements at the first, second, and third sites by x_a, y_a, and z_a for $a \in \mathbb{Z}_2$, respectively. Let the base protocol p measure x_0, then y_0, and finally z_0 (once again, adaptivity is trivial), with the post-processing Z returning the sum of the three outcome bits, and let Q be the linear preprocessing function $Q : \mathbb{Z}_2^2 \to \mathbb{Z}_2^3 :: (a, b) \mapsto (a, b, a + b)$ from Section 1. If $e : B_{\mathcal{I}}$ is an empirical model such that (e, p, Q, Z) deterministically computes the OR function, equation (3) specialises to

$$\begin{aligned} x_0 + y_0 + z_0 &= 0 & \qquad x_0 + y_1 + z_1 &= 1 \\ x_1 + y_0 + z_1 &= 1 & \qquad x_1 + y_1 + z_0 &= 1 \end{aligned}$$

i.e. to (1) from Section 1. Thus Theorem 3.6 recovers the usual AvN argument of the GHZ state as a special case.

We now turn to an example in which adaptivity plays a genuine role, again with $|\mathcal{I}| = 3$. The argument establishing (4) is correspondingly more involved: we sketch it here, indicating how each step of the argument corresponds to a general result formalised in the following sections. We adopt the same notation as in the GHZ example, and let the base protocol p start from x_0, then proceed to y, and finally to

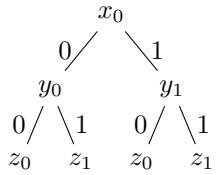
z, with the measurement settings for y and z given by the immediately preceding outcome. The tree representing p and the matrix T_p (which has rank 2) are given by



and $T_p = \begin{array}{c} \\ x \\ y \\ z \end{array} \begin{array}{c} \begin{array}{ccc} x & y & z \end{array} \\ \begin{bmatrix} 0 & 0 & 0 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{bmatrix} \end{array}$.

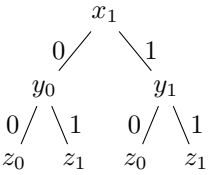
We set the outcome post-processing to return only the final bit, i.e. the outcome of the z_a measurement, and define the input encoding $Q : \mathbb{Z}_2^2 \to \mathbb{Z}_2^3$ by $(a, b) \mapsto (a, b, b)$. Now, if we assume that we are computing the AND function using these choices for p, Q, Z, we obtain the following equations:

(5–i)



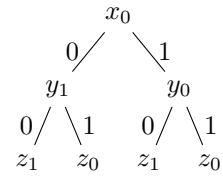
$= 0$

(5–ii)



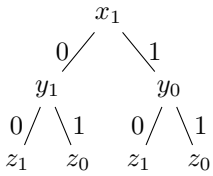
$= 0$

(5–iii)



$= 0$

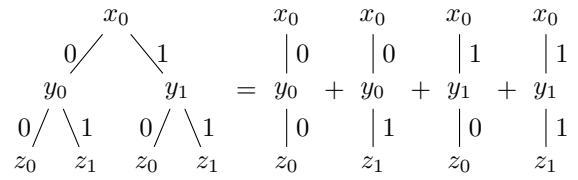
(5–iv)



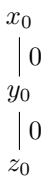
$= 1$

These equations are slightly informal and in particular, they omit the additional detail that the protocols return only a single bit (the outcome of the final z_a measurement), so we must remember this ourselves. Again, the right-hand sides sum to 1; we thus wish to show that the left-hand sides sum to 0, establishing that eq. (4) holds. But to do so, we need to evaluate such sums of trees.

The first key identity is that a tree equals the sum of its branches; for the first tree in (5), this reads:



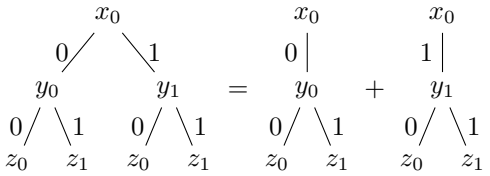
To explain the meaning of such an equation, let us first look at the interpretation of a single branch, say the leftmost one:



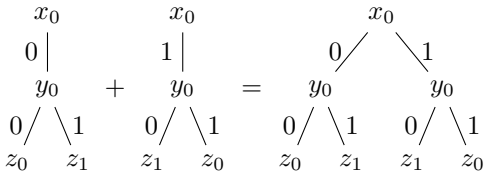
It depicts a (no longer linear) measurement protocol, which starts from the top and attempts to obtain the indicated outcomes in order: if it succeeds, it returns the outcome of the final measurement z_0; if at any point a measurement returns a non-indicated outcome (e.g. measuring x_0 yields outcome 1), the protocol halts and returns 0. Both the tree itself and its branches are measurements in their own right in $\mathsf{MP}(B_{\mathcal{I}}, \mathbb{Z}_2)$. Asserting an equation like this is to claim that the support of

our empirical model (extended to $\mathsf{MP}(B_{\mathcal{I}}, \mathbb{Z}_2)$) satisfies it, i.e. that upon measuring both sides of the equation, the obtained joint outcome satisfies the equation.

While all equations we discuss could be expressed by breaking every protocol into its branches, it is conceptually clearer to group branches into larger subprotocols. Hence our base protocol also satisfies the equation



asserting that the whole protocol equals the sum of its left and right halves. A similar equation is satisfied by each of the four trees in (5). As the aim is to show that the sum of these four trees is zero, when evaluating this sum we can regroup these halves as we wish. In particular, if we add the left half of the base protocol (5–i) and the right half of (5–iii), we obtain another linear measurement protocol:

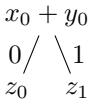


Such re-groupings are formalised in Lemmas 6.4 and 6.5.

This recombined measurement protocol has a different adaptivity structure from the original ones: after measuring x_0, it measures y_0 regardless of the outcome, while the final measurement setting for z is determined by the parity of the first two outcomes. The corresponding adaptivity matrix is

$$\begin{array}{c} \\ x \\ y \\ z \end{array} \begin{array}{c} \begin{array}{ccc} x & y & z \end{array} \\ \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 1 & 1 & 0 \end{bmatrix} \end{array},$$

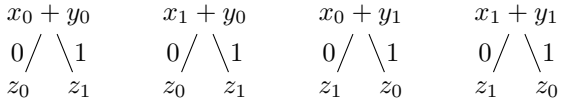
which now has rank 1. We could therefore condense this information and represent the measurement protocol as



which first measures both x_0 and y_0, then, based on their parity $x_0 + y_0$, selects the setting to use for z, and finally returns the outcome of this last measurement. Such measurement protocols that may measure a sum of basic measurements in a single step are defined formally in Definition 5.1, and this operation of condensing a longer protocol into fewer levels is formalised in Lemma 6.6.

Performing this regrouping for all four trees in (5), we find that showing equation (4) holds for them is equivalent to showing that it holds for the following four trees:

(6)



In summary, we have regrouped the four original measurement protocols (without affecting their total sum) to obtain four protocols with simpler adaptivity, at the

cost of allowing protocols that measure a sum of basic measurements in a single step.

Now, to show that these sum to zero, we first assert the equation

$$\overset{x_a + y_b}{\underset{z_c}{\big|}} r \quad = \quad \overset{x_a}{\underset{z_c}{\big|}} r \; + \; \overset{y_b}{\underset{z_c}{\big|}} 1$$

(true in a measurement context of the scenario of measurement protocols on $B_{\mathcal{I}}$; see proof of Lemma 6.8) for each branch. Applying this to the first tree in (6) yields

$$\overset{x_0 + y_0}{\underset{z_0}{0/}} \; \overset{}{\underset{z_1}{\backslash 1}} \quad = \quad \overset{x_0 + y_0}{\underset{z_0}{\big|}} 0 \; + \; \overset{x_0 + y_0}{\underset{z_1}{\big|}} 1 \quad = \quad \overset{x_0}{\underset{z_0}{\big|}} 0 + \overset{y_0}{\underset{z_0}{\big|}} 1 + \overset{x_0}{\underset{z_1}{\big|}} 1 + \overset{y_0}{\underset{z_1}{\big|}} 1 \quad = \quad \overset{x_0}{\underset{z_0}{0/}} \; \overset{}{\underset{z_1}{\backslash 1}} + \overset{y_0}{\underset{z_0}{\big|}} 1 + \overset{y_0}{\underset{z_1}{\big|}} 1$$

This kind of manipulation is demonstrated more carefully and in greater generality in Lemma 6.8. Proceeding similarly for all four trees in (6) results in

$$\overset{x_0 + y_0}{\underset{z_0}{0/}} \; \overset{}{\underset{z_1}{\backslash 1}} \quad = \quad \overset{x_0}{\underset{z_0}{0/}} \; \overset{}{\underset{z_1}{\backslash 1}} \; + \; \overset{y_0}{\underset{z_0}{\big|}} 1 \; + \; \overset{y_0}{\underset{z_1}{\big|}} 1 \qquad \overset{x_1 + y_0}{\underset{z_0}{0/}} \; \overset{}{\underset{z_1}{\backslash 1}} \quad = \quad \overset{x_1}{\underset{z_0}{0/}} \; \overset{}{\underset{z_1}{\backslash 1}} \; + \; \overset{y_0}{\underset{z_0}{\big|}} 1 \; + \; \overset{y_0}{\underset{z_1}{\big|}} 1$$

$$\overset{x_0 + y_1}{\underset{z_1}{0/}} \; \overset{}{\underset{z_0}{\backslash 1}} \quad = \quad \overset{x_0}{\underset{z_1}{0/}} \; \overset{}{\underset{z_0}{\backslash 1}} \; + \; \overset{y_1}{\underset{z_1}{\big|}} 1 \; + \; \overset{y_1}{\underset{z_0}{\big|}} 1 \qquad \overset{x_1 + y_1}{\underset{z_1}{0/}} \; \overset{}{\underset{z_0}{\backslash 1}} \quad = \quad \overset{x_1}{\underset{z_1}{0/}} \; \overset{}{\underset{z_0}{\backslash 1}} \; + \; \overset{y_1}{\underset{z_1}{\big|}} 1 \; + \; \overset{y_1}{\underset{z_0}{\big|}} 1$$

Therefore, to evaluate the sum of the four trees in (6), we may instead add the right-hand sides of the four equations above. These split neatly into protocols that start from x and others that start from y.

The latter sum to zero as each branch occurs twice. As for the summands starting with x, they form a sum similar to the one we started with, but involving fewer variables. In the context of the general proof, this is then handled by the induction hypothesis. Unwinding the resulting construction, we proceed by repeating the earlier trick of splitting trees into halves and rearranging. Specifically, when adding the two trees starting with x_0 we can compute as follows:

$$\overset{x_0}{\underset{z_0}{0/}} \; \overset{}{\underset{z_1}{\backslash 1}} \; + \; \overset{x_0}{\underset{z_1}{0/}} \; \overset{}{\underset{z_0}{\backslash 1}} \quad = \quad \overset{x_0}{\underset{z_0}{\big|}} 0 \; + \; \overset{x_0}{\underset{z_1}{\big|}} 1 \; + \; \overset{x_0}{\underset{z_1}{\big|}} 0 \; + \; \overset{x_0}{\underset{z_0}{\big|}} 1$$

$$= \quad \overset{x_0}{\underset{z_0}{\big|}} 0 \; + \; \overset{x_0}{\underset{z_0}{\big|}} 1 \; + \; \overset{x_0}{\underset{z_1}{\big|}} 0 \; + \; \overset{x_0}{\underset{z_1}{\big|}} 1 \quad = \quad \overset{x_0}{\underset{z_0}{0/}} \; \overset{}{\underset{z_0}{\backslash 1}} \; + \; \overset{x_0}{\underset{z_1}{0/}} \; \overset{}{\underset{z_1}{\backslash 1}} \quad = z_0 + z_1,$$

where we used the equation $\overset{x_0}{\underset{z_c}{0/}} \; \overset{}{\underset{z_c}{\backslash 1}} = z_c$. This boils down to the fact that if a protocol returns the outcome of z_c in every branch, then it might as well just measure z_c directly. Similarly, the two trees starting with x_1 sum to $z_0 + z_1$, cancelling the above. Taken together, these computations show that the four trees in (5) sum to zero, concluding the AvN argument. This kind of step is systematised in Lemma 6.10.

The general proof proceeds along similar lines, i.e. by simplifying the expressions and regrouping them until each grouping sums to zero because every expression in it occurs an even number of times. The main ways of simplifying expressions amount to:

- splitting an expression into two halves, and recombining the halves with those from another tree (Lemma 6.5)
- simplifying the adaptivity at the cost of measuring sums of several measurements at once (Lemma 6.7)
- splitting a level involving a sum of variables into a sum of expressions (Lemma 6.8) and using this to write the whole sum as a sum of smaller things (Lemma 6.10).

We make these steps rigorous in the next sections.

5. Formalising the main result

In this section, we formalise our main result. We begin by extending the class of measurement protocols from Section 3, then assemble these into a contextuality scenario, and finally state and prove the key equation (4) on which Theorem 3.6 rests, modulo some lemmas deferred to the next section.

5.1. **Extended measurement protocols.**

We change our measurement protocols in three ways:

- We incorporate the single output bit returned by the protocol (i.e. the function $Z : \mathbb{Z}_2^{s(p)} \to \mathbb{Z}_2$) into the protocol itself.
- We allow for two kinds of nodes: *branching nodes* (those considered so far), which split (linearly) into two continuations, and *multiplicative nodes*, which continue only if a particular outcome is observed, otherwise returning 0.
- We allow nodes to have arbitrary arity, simultaneously performing a finite set $D = \{x_1, \ldots, x_d\}$ of measurements at distinct sites, rather than a single measurement at a single site. The operational meaning is that D stands for performing all the measurements in D and returning only their total parity $x_1 + \cdots + x_d \in \mathbb{Z}_2$ (rather than their individual outcomes).

The incorporation of the output bit and higher-arity nodes are technical conveniences that simplify the inductive argument; multiplicative nodes, on the other hand, genuinely extend the expressive power of protocols and play an essential role in the argument. Although these are less general than the measurement protocols of [9], we call them simply *measurement protocols*, as the more general notion does not arise in this work.

From now on, given sets U and V, we write $U + V$ for their disjoint union, and moreover we denote a singleton set by 1. In particular, given a set of sites U, $U + 1$ stands for U with an additional point $*$, which we use to index the output bit of a protocol.

Definition 5.1. For sets of sites $U, V \subseteq \mathcal{I}$ (not necessarily disjoint) we define the set $\mathsf{MP}_{U,V}$ of measurement protocols with multiplicative nodes at sites in U and branching nodes at sites in V. For a protocol $p \in \mathsf{MP}_{U,V}$, we write $s(p) := U \cup V$ for the set of sites it measures, and also define:

- the set $S_p \subseteq \mathscr{P}(s(p))$ of jointly measured sets of sites;
- the set $M_p \subseteq X$ of possibly performed measurements.

These notions are defined inductively. Call a set of measurements $D = \{x_1, \ldots, x_d\}$ on distinct sites $s(D) := \{s(x_1), \ldots, s(x_d)\}$ *admissible for* $q \in \mathsf{MP}_{U,V}$ if

(1) $s(D) \notin S_q$, i.e. a measurement on the exact same set of sites is not performed in q;
(2) $\{x \in M_q \mid s(x) \in s(D)\} \subseteq D$, i.e. if a site in $s(D)$ is possibly measured in q, it is always measured with the same setting as in D.

The inductive clauses are then as follows:

- We set $\mathsf{MP}_{\emptyset,\emptyset} := \mathbb{Z}_2 = \{0, 1\}$. For $p \in \mathsf{MP}_{\emptyset,\emptyset}$, we set $S_p, M_p := \emptyset$.

- If $q \in \mathsf{MP}_{U,V}$, $v \in \mathbb{Z}_2^{s(q)+1}$, and $D = \{x_1, \ldots, x_d\}$ is admissible for q and $\pi_{s(D)\cap s(q)} v = 0$, then $p := (D, v, q)$ is in $\mathsf{MP}_{U,V\cup s(D)}$. We call this a d-ary *branching node*. We set $S_p := S_q \cup \{s(D)\}$ and $M_p := M_q \cup v.M_q \cup D$, where $v.M_q := \{\pi_{s(x)}(v).x \mid x \in M_q\}$
- If $q \in \mathsf{MP}_{U,V}$, $r \in \mathbb{Z}_2$, and $D = \{x_1, \ldots, x_d\}$ is admissible for q, then $p := \langle D, r, q\rangle$ is a measurement protocol in $\mathsf{MP}_{U\cup s(D),V}$. We call this a d-ary *multiplicative node*. We set $S_p := S_q \cup \{s(D)\}$ and $M_p := M_q \cup D$.

We call a node (whether multiplicative or branching) *unary* if $d = 1$. A *branch* is a measurement protocol with only multiplicative nodes ending in 1.

We now extend the action from Definition 3.2.

Definition 5.2. We define an action of $\mathbb{Z}_2^{U\cup V+1}$ on $\mathsf{MP}_{U,V}$ by induction on the structure of measurement protocols.

- For the base case, the action of $w \in \mathbb{Z}_2^1$ on $p \in \mathsf{MP}_{\emptyset,\emptyset} = \mathbb{Z}_2$ is given by $w.p := w+p$, i.e. the regular action of $\mathbb{Z}_2$ on itself.
- For a branching node, the action of $w \in \mathbb{Z}_2^{s(p)+1}$ on $p = (D, v, q)$ is defined as

$$w.p := \left(\{\pi_{s(x)}(w).x \mid x \in D\}, v, \pi_{s(q)+1}(w).q\right).$$

- For a multiplicative node, the action of $w \in \mathbb{Z}_2^{s(p)+1}$ on $p = \langle D, r, q\rangle$ is defined as

$$w.p := \left\langle \{\pi_{s(x)}(w).x \mid x \in D\}, r, \pi_{s(q)+1}(w).q \right\rangle.$$

Measurement protocols (D, v, q) (starting with a branching node) and $\langle D, r, q\rangle$ (starting with a multiplicative node), with $D = \{x_1, \ldots, x_d\}$, are respectively represented as

$$\begin{array}{ccc} x_1 + \cdots + x_d & & x_1 + \cdots + x_d \\ 0\big/ \quad \big\backslash 1 & \text{and} & \big| \, r \\ q \quad v.q & & q \end{array}$$

In the first case, the operational interpretation is that the protocol starts by measuring $x_1, \ldots, x_d$, then proceeds to either q or $v.q$ depending on the total parity $x_1 + \cdots + x_d$. Note that the side-condition that $\pi_{s(D)\cap s(q)} v = 0$ ensures that the outcome of D cannot affect the measurement settings of sites in $s(D)$ (which have already been measured). Under the assumption that D is admissible for q, this is equivalent to D also being admissible for $v.q$.

In the second case, the interpretation is that p measures $x_1, \ldots, x_d$: if the total parity equals r, it proceeds to q; otherwise it returns 0 immediately. For the cases $r = 0$ and $r = 1$, we could (but won't) depict a protocol $\langle D, r, q\rangle$ starting with a multiplicative node as

$$\begin{array}{ccccc} x_1 + \cdots + x_d & & x_1 + \cdots + x_d & & \\ 0\big/ \quad \big\backslash 1 & \text{and} & 0\big/ \quad \big\backslash 1 & & \text{respectively.} \\ q \quad 0 & & 0 \quad q & & \end{array}$$

Definition 5.3. For a protocol p, we define the *number of multiplicative levels* $m(p)$ and the *number of branching levels* $n(p)$ by induction on the structure of p:

$$m(p) := \begin{cases} 0 & \text{if } p \in \mathsf{MP}_{\emptyset,\emptyset}, \\ m(q) & \text{if } p = (D, v, q), \\ m(q) + 1 & \text{if } p = \langle D, r, q\rangle; \end{cases} \qquad n(p) := \begin{cases} 0 & \text{if } p \in \mathsf{MP}_{\emptyset,\emptyset}, \\ n(q) + 1 & \text{if } p = (D, v, q), \\ n(q) & \text{if } p = \langle D, r, q\rangle. \end{cases}$$

We now slightly generalise the definition of the matrix associated to a linear measurement protocol, to accommodate unary multiplicative nodes as well.

Definition 5.4. Given a measurement protocol $p \in \mathsf{MP}_{U,V}$ with all nodes unary,[9] we define the *adaptivity matrix* T_p (or simply T) of p, a $(U + V + 1) \times V$ matrix over $\mathbb{Z}_2$, by induction on the structure of p:

- If $p \in \mathsf{MP}_{\emptyset,\emptyset}$, then T_p is the unique 1×0 matrix.
- If $p = (x, v, q)$ starts with a branching node where $q \in \mathsf{MP}_{U,V}$, we define

$$T_p := \begin{array}{r} \\ s(x) \\ U + V + 1 \end{array} \begin{array}{c} \begin{array}{cc} s(x) & V \end{array} \\ \begin{bmatrix} 0 & 0 \\ v & T_q \end{bmatrix} \end{array}.$$

- If $p = \langle x, r, q \rangle$ starts with a multiplicative node where $q \in \mathsf{MP}_{U,V}$, we define

$$T_p := \begin{array}{r} \\ s(x) \\ U + V + 1 \end{array} \begin{array}{c} \begin{array}{c} V \end{array} \\ \begin{bmatrix} 0 \\ T_q \end{bmatrix} \end{array}.$$

The largest possible rank for the adaptivity matrix T_p is $n(p)$, the number of branching levels; we say that T_p (or p) is *of maximal rank* if its rank is exactly $n(p)$.

Note that for protocols p with all nodes unary, the sets $s(p)$ remain intrinsically linearly ordered: for $p = (x, v, q)$ or $p = \langle x, r, q \rangle$, the site $s(x)$ precedes those in $s(q)$. This intrinsic order on $s(p)$ extends to $s(p) + 1$ by placing the additional output point last; we use this extended order throughout.

By construction, T_p is strictly lower triangular in the intrinsic order of p. As T_p is not a square matrix, this needs further clarification. The rows of T_p are indexed by $U + V + 1$ and the columns are indexed by $V \subseteq U + V + 1$. By strict lower triangularity we mean that for each $i \in V$, all entries on the ith column up to i in the intrinsic order are zero; explicitly, $T_{j,i} = 0$ for all $i \in V$ and $j \leq i$ in this order.

The operational meaning of T_p is the same as in Section 3.1 for linear protocols: the column indexed by branching site $i \in V$ specifies how the outcome at site i affects the measurement settings at other sites. Strict lower triangularity ensures that a measurement outcome can only influence the settings of strictly later measurements.

Remark 5.5. As in Remark 3.4, specifying a $(U+V+1) \times V$-matrix T that is strictly lower triangular in some ordering of $U + V + 1$ (with the output point last) does not uniquely determine a protocol $p \in \mathsf{MP}_{U,V}$ with unary nodes: the same matrix may be strictly lower triangular in several orderings, and additional data is required even once an ordering is fixed. Specifically, to reconstruct p, we must also provide a vector $w_p \in \mathbb{Z}_2^{s(p)+1}$ of default measurement settings (and default output bit) and a vector $r_p \in \mathbb{Z}_2^U$ of required outcomes for the multiplicative nodes. Together with the ordering and the adaptivity matrix T_p, these data determine a unique measurement protocol $p \in \mathsf{MP}_{U,V}$. As before, the action $(v, p) \mapsto v.p$ translates in this representation to $v.(T_p, w_p, r_p) = (T_p, v + w_p, r_p)$.

5.2. **The scenario of measurement protocols.** We now want to define the scenario whose measurements are the measurement protocols. For this, we require an auxiliary notion to define compatibility for measurement protocols. Intuitively, $W(p, g)$ is the set of measurements that are actually performed when running a protocol p with outcomes determined by a global assignment g, following the path through the tree that g selects at each node. In parallel, we define $o(p, g) \in \mathbb{Z}_2$ to record the output bit produced by that run.

[9]Since all nodes are unary, U and V are necessarily disjoint, so $s(p) = U \cup V \cong U + V$. The restriction to unary nodes is not necessary to define the matrix, but it does make things technically simpler and is sufficiently general for our purposes.

Definition 5.6. Let p be a measurement protocol on the scenario $B_{\mathcal{I}}$ and $g : \coprod_{i\in\mathcal{I}} \mathbb{Z}_2 \to \mathbb{Z}_2$ a global assignment of outcomes to the measurements of $B_{\mathcal{I}}$. We define the set $W(p,g)$ of *measurements performed by p given g* and the *output of p given g*, $o(p,g) \in \mathbb{Z}_2$, by simultaneous induction on the structure of p:

- If $p \in \mathsf{MP}_{\emptyset,\emptyset}$, set $W(p,g) = \emptyset$ and $o(p,g) = p$ (viewing $p \in \mathbb{Z}_2$).
- If $p = (D, v, q)$ starts with a branching node,

$$W(p,g) = \begin{cases} D \cup W(q,g) & \text{if } \sum_{x\in D} g(x) = 0, \\ D \cup W(v.q,g) & \text{otherwise;} \end{cases}$$

$$o(p,g) = \begin{cases} o(q,g) & \text{if } \sum_{x\in D} g(x) = 0, \\ o(v.q,g) & \text{otherwise.} \end{cases}$$

- If $p = \langle D, r, q\rangle$ starts with a multiplicative node,

$$W(p,g) = \begin{cases} D \cup W(q,g) & \text{if } \sum_{x\in D} g(x) = r, \\ D & \text{otherwise.} \end{cases}$$

$$o(p,g) = \begin{cases} o(q,g) & \text{if } \sum_{x\in D} g(x) = r, \\ 0 & \text{otherwise.} \end{cases}$$

Note that $o(p,g)$ depends on g only through $g|_{W(p,g)}$, i.e. the output is determined by the outcomes of the measurements actually performed in the run. By the admissibility requirements in Definition 5.1, the set $W(p,g)$ is always a context of $B_{\mathcal{I}}$.

Definition 5.7. Given the Bell scenario $B_{\mathcal{I}}$ on a set of sites $\mathcal{I}$, we define the *scenario* $\mathsf{MP}(B_{\mathcal{I}}, \mathbb{Z}_2)$ *of measurement protocols on* $B_{\mathcal{I}}$:

- The measurements are given by measurement protocols on $B_{\mathcal{I}}$ (in the sense of Definition 5.1).
- A set P of measurement protocols forms a *context* if, for every global assignment g of outcomes to the measurements of $B_{\mathcal{I}}$, the set $W(P,g) := \bigcup_{p\in P} W(p,g)$ is a context of $B_{\mathcal{I}}$.
- The outcome set is fixed to be $\mathbb{Z}_2$.

Every empirical model $e : B_{\mathcal{I}}$ extends to an empirical model $\mathsf{MP}(e, \mathbb{Z}_2) : \mathsf{MP}(B_{\mathcal{I}}, \mathbb{Z}_2)$, defined by assigning to each context P the distribution over joint outputs obtained by jointly running each protocol $p \in P$ on e and recording the output bit. We briefly set out its definition below; see [9, Definition 16] for a precise definition with a more detailed explanation in an analogous setting – the main difference being that there the full sequence of outcomes observed throughout a run is recorded, rather than a single output bit; the model we define here is thus a coarse-graining thereof.

Recall from Section 2 the definition of the event sheaf: $\mathscr{E}(U)$ denotes the set of outcome assignments to the measurements in a context U. For the scenario $B_{\mathcal{I}}$, $\mathscr{E}(U)$ is the set of functions $U \to \mathbb{Z}_2$, where U is a subset of the measurement set $X = \coprod_{i\in\mathcal{I}} \mathbb{Z}_2$ of $B_{\mathcal{I}}$.

Given a context P of $\mathsf{MP}(B_{\mathcal{I}}, \mathbb{Z}_2)$, write

$$\mathcal{F}_P := \{g|_{W(P,g)} \in \mathscr{E}(W(P,g)) \mid g \in \mathscr{E}(X)\}$$

for the set of partial assignments of outcomes to measurements that may be observed in a (joint) run of the protocols in P. Since, as remarked right after its definition, $o(p,g)$ depends on g only through $g|_{W(p,g)}$, each map $o(p,\cdot) : \mathscr{E}(X) \to \mathbb{Z}_2$ factors through the restriction $g \mapsto g|_{W(P,g)}$, yielding a well-defined map $\tilde{o}(p,\cdot) : \mathcal{F}_P \to \mathbb{Z}_2$.

Definition 5.8. Given an empirical model $e : B_{\mathcal{I}}$, the induced empirical model $\mathsf{MP}(e, \mathbb{Z}_2) : \mathsf{MP}(B_{\mathcal{I}}, \mathbb{Z}_2)$ is defined as follows: for each context P of $\mathsf{MP}(B_{\mathcal{I}}, \mathbb{Z}_2)$, the probability of observing an outcome assignment $r : P \to \mathbb{Z}_2$ is

$$\mathsf{MP}(e, \mathbb{Z}_2)_P(r) = \sum_{\substack{(h:U \to \mathbb{Z}_2) \in \mathcal{F}_P \\ \forall p \in P,\, \tilde{o}(p,h) = r(p)}} e_U(h) \ .$$

In much of what follows we prove equations that hold in the $\mathbb{Z}_2$-linear theory of $\mathsf{MP}(e, \mathbb{Z}_2)$ for *every* empirical model $e : B_{\mathcal{I}}$, i.e. equations in the intersection of the $\mathbb{Z}_2$-linear theories of $\mathsf{MP}(e, \mathbb{Z}_2)$ over all $e : B_{\mathcal{I}}$. This common theory may be thought of as the $\mathbb{Z}_2$-linear theory of the scenario $\mathsf{MP}(B_{\mathcal{I}}, \mathbb{Z}_2)$ itself. Of course, the theory common to all empirical models on any given measurement scenario is trivial, consisting only of tautologies. Here it is not, because the models of the form $\mathsf{MP}(e, \mathbb{Z}_2)$ for $e : B_{\mathcal{I}}$ constitute a proper subclass of all the empirical models on $\mathsf{MP}(B_{\mathcal{I}}, \mathbb{Z}_2)$ seen as a bare measurement scenario: it contains only those arising by *running* measurement protocols on some underlying $B_{\mathcal{I}}$-model. Thus the MP construction implicitly constrains the class of empirical models deemed valid in the scenario. The following proposition characterises this common theory, showing that validity of an equation reduces to a condition on global outcome assignments.

Proposition 5.9. *Let P be a context of $\mathsf{MP}(B_{\mathcal{I}}, \mathbb{Z}_2)$ and $a \in \mathbb{Z}_2$. The following are equivalent:*

(i) $\mathsf{MP}(e, \mathbb{Z}_2) \vDash \sum_{p \in P} p = a$ *for every empirical model* $e : B_{\mathcal{I}}$*;*
(ii) for every global assignment $g : \coprod_{i \in \mathcal{I}} \mathbb{Z}_2 \to \mathbb{Z}_2$, $\sum_{p \in P} o(p, g) = a$.

Proof. ($\Rightarrow$) Suppose $\mathsf{MP}(e, \mathbb{Z}_2) \vDash \sum_{p \in P} p = a$ for every $e : B_{\mathcal{I}}$. In particular, any global assignment g determines a (deterministic, non-contextual) empirical model $\delta_g : B_{\mathcal{I}}$ defined by $(\delta_g)_C := \delta_{g|_C}$, which produces outcomes according to g with certainty. Such a deterministic model lifts to a deterministic model on $\mathsf{MP}(B_{\mathcal{I}}, \mathbb{Z}_2)$ given by $\mathsf{MP}(\delta_g, \mathbb{Z}_2) = \delta_{o(\cdot, g)}$, with $o(\cdot, g)$ being a global assignment for the scenario $\mathsf{MP}(B_{\mathcal{I}}, \mathbb{Z}_2)$. The equation holding for δ_g thus gives $\sum_{p \in P} o(p, g) = a$.

($\Leftarrow$) Suppose $\sum_{p \in P} o(p, g) = a$ for every g, and let $e : B_{\mathcal{I}}$ be an arbitrary empirical model. An outcome assignment $r : P \to \mathbb{Z}_2$ in the support of $\mathsf{MP}(e, \mathbb{Z}_2)_P$ arises from some $h \in \mathcal{F}_P$ with $\tilde{o}(p, h) = r(p)$ for all $p \in P$. In turn, such h is the restriction of some global assignment g with $W(P, g) = \mathrm{dom}(h)$ (one may freely assign outcomes to measurements outside $\mathrm{dom}(h)$ to obtain such an extension), so that $o(p, g) = \tilde{o}(p, h) = r(p)$ for all $p \in P$. By assumption, $\sum_{p \in P} o(p, g) = a$, hence $\sum_{p \in P} r(p) = a$. □

Both the construction of the induced model $\mathsf{MP}(e, \mathbb{Z}_2)$ (Definition 5.8) and the characterisation of the theory common to all such models on a given scenario (Proposition 5.9) extend verbatim to arbitrary measurement protocols in the broader sense considered in [9], where the possible continuations of a branching node may be arbitrary, rather than linearly related. Specifically for the measurement protocols considered in this paper – where adaptivity is *linear*, in that the outcome of a branching node acts on the continuation by adding a vector to later measurement settings – the following proposition gives a more concrete operational description using the adaptivity matrix. We restrict to protocols with all nodes unary, since the adaptivity matrix has only been defined in that setting; this is a technical convenience that keeps the exposition simple, and could be lifted at the cost of some extra complexity. In the statement below, the vector b records the measurement outcomes that are (or would be) observed at each branching site, agreeing with the value assigned by $g|_{W(p,g)}$ above.

Proposition 5.10. *Let $p \in \mathsf{MP}_{U,V}$ be a protocol with all nodes unary, with adaptivity matrix $T = T_p$, default measurement settings and output $w = w_p \in \mathbb{Z}_2^{s(p)+1}$ and required outcomes $r = r_p \in \mathbb{Z}_2^U$ for the multiplicative sites. Given a global assignment $g : \coprod_{i\in\mathcal{I}} \mathbb{Z}_2 \to \mathbb{Z}_2$, define the vector $b = b(p,g) \in \mathbb{Z}_2^V$ of measurement outcomes at the branching sites by recursion along the intrinsic order: for each branching site $j \in V$,*

$$b_j := g\big(x_{j,\,(w+Tb)_j}\big) = g\big(x_{j,\,w_j+\sum_{i\in V,\,i<j} T_{j,i}b_i}\big),$$

which is well defined since T is strictly lower triangular.

The run of p against g measures $x_{j,(w+Tb)_j}$ at each visited site j. If $g(x_{k,(w+Tb)_k}) = r_k$ for every multiplicative site $k \in U$, then the run completes without aborting:

$$W(p,g) = \{\, x_{j,(w+Tb)_j} \mid j \in s(p) \,\}, \qquad o(p,g) = (w+Tb)_*,$$

with $*$ the output point. Otherwise, letting k^* be the least $k \in U$ such that $g\big(x_{k,(w+Tb)_k}\big) \neq r_k$, the run aborts at k^*:

$$W(p,g) = \{\, x_{j,(w+Tb)_j} \mid j \in s(p),\ j \le k^* \,\}, \qquad o(p,g) = 0.$$

Proof. The proof is straightforward, by induction on the structure of p, using the definitions of T_p (Definition 5.4) and of $W(p,g)$ and $o(p,g)$ (Definition 5.6). □

5.3. **The main inductive argument.** As noted in Section 3.2, to prove Theorem 3.6, it suffices to establish equation (4). We now state the latter more precisely.

Theorem 5.11. *Let $p \in \mathsf{MP}_{U,V}$ be a measurement protocol on $B_\mathcal{I}$ with all nodes unary, and $e : B_\mathcal{I}$ be an empirical model. If $Q : \mathbb{Z}_2^l \to \mathbb{Z}_2^{s(p)+1}$ is a linear map where $l \ge m(p) + 2$, then*

$$\mathsf{MP}(e, \mathbb{Z}_2) \vDash^* \sum_{w\in\mathbb{Z}_2^l} Q(w).p = 0\ . \tag{7}$$

Ultimately we care about the case when $m(p) = 0$: the added generality afforded by multiplicative nodes is not of intrinsic interest here, but is needed to ensure a sufficiently strong induction hypothesis.

We now give the overall structure of the proof, formalising the proof sketch of Section 3.2; the key lemmas are stated and proved in Section 6.

Proof of Theorem 5.11. We prove the claim by induction on the number of branching levels $n(p)$.

If $n(p) = 0$, then the codomain of Q has dimension $m(p)+n(p)+1 = m(p)+1 < l$, so Q is not injective. As a result, each summand in eq. (7) appears an even number of times and hence cancels out, so that the equation holds.

For $n(p) > 0$, if Q fails to be injective, eq. (7) holds by the argument above. If Q is injective, the result follows from the following three lemmas:

(1) If $\operatorname{rank} T_p = n(p)$, Lemma 6.5 shows that there is another measurement protocol $q \in \mathsf{MP}_{U,V}$ with all nodes unary such that

$$\sum_{w\in\mathbb{Z}_2^l} Q(w).p \;=\; \sum_{w\in\mathbb{Z}_2^l} Q(w).q$$

and $\operatorname{rank} T_q < n(p) = n(q)$. In other words, without loss of generality, we can assume that T_p is not of maximal rank.

(2) If $\operatorname{rank} T_p < n(p)$, Lemma 6.7 shows that we can condense p to have strictly fewer branching levels without affecting the value of the sum $\sum_w Q(w).p$, at the cost of increasing the arity of some nodes.

(3) Given a measurement protocol, Lemma 6.10 shows that we can systematically replace all higher-arity branching levels by sums of protocols with unary nodes without increasing the number of branching levels. Combined with Step 2, this lets us write the overall sum as a sum of smaller sums $\sum_v Q_i(v).p_i$, where each p_i has strictly fewer than $n(p)$ branching levels, therefore reducing to the induction hypothesis. □

6. Tools for the proof

In this section, we establish the key lemmas on which the proof of Theorem 5.11 rests. After a preliminary result on decomposing protocols into branches (Section 6.1), each of the three main subsections addresses one step of the proof: splicing to reduce the rank of the adaptivity matrix (Section 6.2), condensing protocols with non-maximal rank into fewer branching levels (Section 6.3), and removing higher-arity branching nodes by expressing them as sums of simpler protocols (Section 6.4). Each of these three follows a common pattern: a key construction is introduced and a general lemma about it established, which is then applied in the specific context of Theorem 5.11.

6.1. Decomposing into branches.

We begin by establishing that every measurement protocol decomposes as the sum of its branches, a fact used throughout the sequel.

Recall from Definition 5.1 that a branch is a measurement protocol with only multiplicative nodes, ending in 1. It records a single deterministic path through a protocol tree, given by a sequence of parity measurements each with a required outcome, that returns 1 when every one of them is observed, and 0 otherwise.

Definition 6.1. For a measurement protocol p, we define its set of branches, $\mathrm{Br}(p)$, by induction on the structure of p:

- For $p \in \mathsf{MP}_{\emptyset,\emptyset}$, we set $\mathrm{Br}(0) = \emptyset$ and $\mathrm{Br}(1) = \{1\}$.
- If $p = \langle D, r, q\rangle$ starts with a multiplicative node, then
$$\mathrm{Br}(p) := \{\langle D, r, t\rangle \mid t \in \mathrm{Br}(q)\}\,.$$
- If $p = (D, v, q)$ starts with a branching node, then
$$\mathrm{Br}(p) := \mathrm{Br}(\langle D, 0, q\rangle) \cup \mathrm{Br}(\langle D, 1, v.q\rangle).$$

Lemma 6.2. *The $\mathbb{Z}_2$-linear theory of $\mathsf{MP}(e, \mathbb{Z}_2)$ respects pre-composition by a multiplicative node: if $\sum_i p_i = \sum_j q_j$ and each*

$$\langle D, r, p_i\rangle = \begin{array}{c} x_1 + \cdots + x_d \\ \big| \, r \\ p_i \end{array} \quad \text{and} \quad \langle D, r, q_j\rangle = \begin{array}{c} x_1 + \cdots + x_d \\ \big| \, r \\ q_j \end{array}$$

is a well-defined protocol, then $\sum_i \langle D, r, p_i\rangle = \sum_j \langle D, r, q_j\rangle$, i.e.

$$\sum_i \begin{array}{c} x_1 + \cdots + x_d \\ \big| \, r \\ p_i \end{array} = \sum_j \begin{array}{c} x_1 + \cdots + x_d \\ \big| \, r \\ q_j \end{array}\,.$$

Proof. This follows from Proposition 5.9 by a straightforward case split, for each global assignment g, on whether $\sum_{x\in D} g(x) = r$. □

Lemma 6.3. *Every measurement protocol p equals the sum of its branches:*

$$p = \sum_{s \in \mathrm{Br}(p)} s\ . \tag{8}$$

Proof. We prove this by induction on the structure of p.

- If $p \in \mathsf{MP}_{\emptyset,\emptyset}$, the claim is clear.
- If $p = \langle D, r, q\rangle$ starts with a multiplicative node, where $D = \{x_1, \dots, x_d\}$, then

$$\overset{x_1+\cdots+x_d}{\underset{q}{\big|\, r}} = \sum_{t\in \mathrm{Br}(q)} \overset{x_1+\cdots+x_d}{\underset{t}{\big|\, r}} = \sum_{s\in \mathrm{Br}(p)} s$$

  where the first equality follows from the induction hypothesis by Lemma 6.2 and the second by the multiplicative-node case of the definition of $\mathrm{Br}(p)$ (Definition 6.1).
- If $p = (D, v, q)$ starts with a branching node, where $D = \{x_1, \dots, x_d\}$, then

$$\overset{x_1+\cdots+x_d}{\underset{q}{0\big/}\ \underset{v.q}{\big\backslash 1}} = \overset{x_1+\cdots+x_d}{\underset{q}{\big|\, 0}} + \overset{x_1+\cdots+x_d}{\underset{v.q}{\big|\, 1}}$$

$$= \sum_{t\in \mathrm{Br}(q)} \overset{x_1+\cdots+x_d}{\underset{t}{\big|\, 0}} + \sum_{t\in \mathrm{Br}(v.q)} \overset{x_1+\cdots+x_d}{\underset{t}{\big|\, 1}} = \sum_{s\in \mathrm{Br}(p)} s$$

  where the first equality holds for all models on $B_{\mathcal{I}}$, because both sides measure $x_1+\cdots+x_d$ and then depending on the outcome return either q or $v.q$, the second follows from the induction hypothesis by Lemma 6.2, and the final one by the branching-node case of the definition of $\mathrm{Br}(p)$ (Definition 6.1).

□

6.2. **Splicing.** We now address Step 1 of the proof of Theorem 5.11: given a protocol whose adaptivity matrix has maximal rank, we produce an equivalent one of strictly lower rank. We achieve this via a *splicing* construction, which operates on a pair of related protocols, rearranging their branches to modify the adaptivity structure while preserving their sum.

The main construction of this subsection operates on a measurement protocol p with unary nodes and its translate $v.p$ by a vector v with sufficiently many leading zeroes. Let x be the measurement at any branching node of p that precedes the first nonzero entry of v (in the intrinsic order of p). Since v has leading zeroes up to and including $s(x)$, p and $v.p$ agree up to x, diverging only in the continuations that follow it. Thus, schematically, these two protocols look like

$$p = \begin{array}{c} p_1 \\ | \\ x \\ 0\big/\ \big\backslash 1 \\ p_2 \quad v'.p_2 \end{array} \quad \text{and} \quad v.p = \begin{array}{c} p_1 \\ | \\ x \\ 0\big/\ \big\backslash 1 \\ \pi_{s(p_2)+1}(v).p_2 \quad (\pi_{s(p_2)+1}(v)+v').p_2 \end{array}$$

where p_1 denotes an initial subprotocol on which v acts trivially.

We split both protocols along x and recombine their halves, obtaining the two protocols below. Denoting the first of these by q, the second is then $v.q$:

$$q := \begin{array}{c} p_1 \\ | \\ x \\ 0\big/\ \big\backslash 1 \\ p_2 \quad (\pi_{s(p_2)+1}(v)+v').p_2 \end{array} \quad \text{and} \quad v.q = \begin{array}{c} p_1 \\ | \\ x \\ 0\big/\ \big\backslash 1 \\ \pi_{s(p_2)+1}(v).p_2 \quad v'.p_2 \end{array}$$

We call these two protocols the result of *splicing p and $v.p$ along x.*[10]

The multiset of branches across the two resulting protocols remains the same, so their total sum is preserved. Moreover, the spliced pair of protocols q and $v.q$ still differ by the same vector v as the original pair p and $v.p$. However, while in p and $v.p$ the outcome of x affects later settings via v', in q and $v.q$, it does so via $v' + \pi_{s(p_2)+1}(v)$. Choosing a suitable splicing thus lets us rewrite $\sum_w Q(w).p$ as $\sum_w Q(w).q$ with a precise description of how the adaptivity matrix T_q differs from T_p. We collect this into the following statement.

Lemma 6.4. *Let p be a measurement protocol with all nodes unary, x a branching node of p, and $v \in \mathbb{Z}_2^{s(p)+1}$ such that $\pi_i(v) = 0$ for all sites $i \leq s(x)$ in the intrinsic order of p. Then splicing p and $v.p$ along x results in protocols q and $v.q$ satisfying*

(1) $p + v.p = q + v.q$.
(2) The adaptivity matrix T_q differs from T_p only in column $s(x)$, where it equals the corresponding column of T_p plus v. Equivalently, for $u \in \mathbb{Z}_2^V$, $T_q(u) = T_p(u) + \pi_{s(x)}(u)\, v$.

We apply Lemma 6.4 in the context of Theorem 5.11. The following lemma implements Step 1. A suitable splicing is found via a dimension count: the images of T_p and Q together exceed the ambient dimension, so they must intersect nontrivially, yielding the required vector v.

Lemma 6.5. *Let $p \in \mathsf{MP}_{U,V}$ be a measurement protocol with all nodes unary and T_p of maximal rank, and let $Q : \mathbb{Z}_2^l \to \mathbb{Z}_2^{s(p)+1}$ be an injective linear map where $l \geq m(p)+2$. Then there exists a measurement protocol $q \in \mathsf{MP}_{U,V}$ with $n(q) = n(p)$ such that $\mathrm{rank}(T_q) < n(q)$ and*

$$\sum_{w \in \mathbb{Z}_2^l} Q(w).p \;=\; \sum_{w \in \mathbb{Z}_2^l} Q(w).q.$$

Proof. The space $\mathbb{Z}_2^{s(p)+1}$ has dimension $m(p) + n(p) + 1$. By the maximal rank assumption, $\mathrm{Im}\, T_p$ is a subspace of dimension $n(p)$ while since Q is injective, $\mathrm{Im}\, Q$ is a subspace of dimension $l \geq m(p) + 2$. Since

$$\dim(\mathrm{Im}\, T_p) + \dim(\mathrm{Im}\, Q) = n(p) + l \geq m(p) + n(p) + 2 > \dim(\mathbb{Z}_2^{s(p)+1}),$$

these subspaces must intersect nontrivially, so we can find $u^* \in \mathbb{Z}_2^V$ and $w^* \in \mathbb{Z}_2^l$ such that $T_p u^* = Q w^* \neq 0$. Let x be the measurement corresponding to the first (in the intrinsic order) nonzero component of u^*, which must exist since $T_p(u^*) \neq 0$ and so $u^* \neq 0$. By strict lower triangularity of T_p, the first nonzero component of $T_p(u^*) = Q(w^*)$ lies strictly after $s(x)$. Hence, p and $Q(w^*).p$ can be spliced along x, resulting in protocols q and $Q(w^*).q$ as in Lemma 6.4.

To show that

$$\sum_{w \in \mathbb{Z}_2^l} Q(w).p \;=\; \sum_{w \in \mathbb{Z}_2^l} Q(w).q,$$

[10]While we could give a detailed inductive definition and accompanying proofs for the results in this subsection, this would be notationally heavy. We believe that the slightly less formal exposition here is more illuminating.

express $\mathbb{Z}_2^l$ as $\mathbb{Z}_2 w^* \oplus W$. Then

$$\begin{aligned}\sum_{w\in\mathbb{Z}_2^l} Q(w).p &= \sum_{w\in W} Q(w).p + Q(w+w^*).p \\ &= \sum_{w\in W} Q(w).(p + Q(w^*).p) \\ &= \sum_{w\in W} Q(w).(q + Q(w^*).q) \\ &= \sum_{w\in W} Q(w).q + Q(w+w^*).q \\ &= \sum_{w\in\mathbb{Z}_2^l} Q(w).q\end{aligned}$$

where we used Lemma 6.4–1 when moving to the third line.

To see that $\operatorname{rank}(T_q) < n(q)$, it suffices to exhibit a nonzero vector in the kernel of T_q. As T_q differs from T_p only at the column of $s(x)$, the leading 1 of u^*, Lemma 6.4–2 gives us $T_q(u^*) = T_p(u^*) + \pi_{s(x)}(u^*)\, Q(w^*) = T_p(u^*) + Q(w^*) = 0$ as desired. □

6.3. **Condensing.** We now address Step 2 of the proof of Theorem 5.11: given a protocol not of maximal rank, we condense it into an equivalent one with strictly fewer branching levels, at the cost of introducing higher-arity nodes. The key point is that a non-maximal-rank adaptivity matrix effectively encodes fewer independent branching decisions than its depth suggests; we exploit this redundancy to condense it into a shallower equivalent.

Lemma 6.6. *Let $p \in \mathsf{MP}_{U,V}$ be a measurement protocol with all nodes unary. Then there exists a measurement protocol $q \in \mathsf{MP}_{U,V'}$ for some $V' \subseteq V$ that has $n(q) = \operatorname{rank} T_p$ branching levels and satisfies $\mathsf{MP}(e, \mathbb{Z}_2) \vDash q = p$ for any $e : B_{\mathcal{I}}$.*

Proof. The construction relies on choosing particular rows from the adaptivity matrix, resulting in a convenient basis for the row space. Each row in the basis corresponds to a branching level, determining the sum of measurements to be performed; multiplicative levels are interleaved wherever they appear, with the choice of the basis ensuring that the resulting protocol is well defined. In a bit more detail, we build the protocol by going through each row of the matrix from bottom to top. For each row, what we do depends on (i) whether the row in question was included in the basis and (ii) whether the row in question is for a multiplicative or a branching site. When the row in question was included in the basis, we add a corresponding branching node. When the row in question was for a multiplicative node, we add the same multiplicative node. When both conditions hold, we do both constructions, with the branching node before the multiplicative one. Let us now fill in the details.

Let T be the $(U + V + 1) \times V$ adaptivity matrix of p, $w \in \mathbb{Z}_2^{s(p)+1}$ the vector of default measurement settings of p, and $r \in \mathbb{Z}_2^U$ the vector of required outcomes for the multiplicative nodes. For $j \in U + V + 1$, write t_j for the jth row of T. Call a row t_j a pivot if t_j is not in the span of the set $\{t_i \mid i < j\}$ of strictly earlier rows. The pivot rows form a basis of the row space with the property that every row t_j can be uniquely written as a sum $t_j = \sum_i c_{j,i} t_i$ of pivot rows weakly before the jth site so that $c_{j,i} = 0$ whenever $i > j$. For a row t_j, let D_j be the set of measurements at the sites indexing the non-zero entries of t_j, with default measurement settings determined by w, i.e. $D_j = \{x_{i,w_i} \mid T_{j,i} \neq 0\}$.

Let $1, \ldots, N$ enumerate $s(p) + 1$ in the intrinsic order, with $N = *$ the output point. We construct protocols $q_N, \ldots q_1$ inductively, with $q := q_1$. We first set

$$q_N := \begin{cases} (D_N, 1, w_N) & \text{if } t_N \text{ is a pivot row,} \\ w_N & \text{otherwise.} \end{cases}$$

For $i < N$, we first define $v_i \in \mathbb{Z}_2^{s(q_{i+1})+1}$ by $v_i := (c_{j,i})_{j \in s(q_{i+1})+1}$ and then set

$$q_i := \begin{cases} \begin{array}{ccc} & D_i & \\ 0 \swarrow & & \searrow 1 \\ q_{i+1} & & v_i.q_{i+1} \end{array} & \text{if } t_i \text{ is a pivot row and } i \text{ is a branching site.} \\ \begin{array}{ccc} & D_i & \\ 0 \swarrow & & \searrow 1 \\ x_{i,w_i} & & 1.x_{i,w_i} \\ r_i \,\big| & & \big|\, r_i \\ q_{i+1} & & v_i.q_{i+1} \end{array} & \text{if } t_i \text{ is a pivot row and } i \text{ is a multiplicative site.} \\ q_{i+1} & \text{if } t_i \text{ is not a pivot row and } i \text{ is a branching site.} \\ \begin{array}{c} x_{i,w_i} \\ \big| \, r_i \\ q_{i+1} \end{array} & \text{if } t_i \text{ is not a pivot row and } i \text{ is a multiplicative site.} \end{cases}$$

We first argue that each q_i is a well-formed measurement protocol. By construction, the branching nodes correspond to basis vectors and hence operate on pairwise distinct sets of sites. The multiplicative nodes are already multiplicative nodes in p, and hence pairwise distinct. Moreover, U and V are disjoint (as all nodes in p were unary), so that all branching nodes of each q_i are also distinct from the multiplicative nodes. Furthermore, $c_{j,i} = 0$ whenever $i > j$ in the intrinsic order, and lower triangularity of T implies that if $x \in D_i$ for some i, then $s(x) < i$. Therefore, the coordinates of later action vectors v_j for $j \geq i$ vanish on $s(D_i)$, so that every occurrence in q_i of a site in D_i is with the same measurement setting as in D_i. This ensures admissibility, so that each q_i is indeed a measurement protocol. By construction $q \in \mathsf{MP}_{U,V'}$ for some $V' \subseteq V$, and q has precisely $n(q) = \operatorname{rank} T$ branching levels as these correspond to a basis of the row space of T.

To see that $\mathsf{MP}(e, \mathbb{Z}_2) \vDash q = p$ for any $e : B_{\mathcal{I}}$, by Proposition 5.9 it suffices to show that for all global assignments on $B_{\mathcal{I}}$, both q and p result in the same outcome. For that purpose, fix a global assignment $g : \coprod_{i \in \mathcal{I}} \mathbb{Z}_2 \to \mathbb{Z}_2$. Let $b \in \mathbb{Z}_2^V$ be defined as in Proposition 5.10, which applies to p. While the proposition as stated does not apply to q (which has non-unary nodes and hence we haven't defined its matrix), q behaves similarly to p in that, barring them aborting early, the vector giving the measurement settings and the final outcome bit of both p and q is given by $w + Tb$. Moreover, as p and q have the same multiplicative nodes with the same expected outcomes, p aborts iff q aborts, in which case they both return 0. Thus $W(q, g) \subseteq W(p, g)$ and in particular p and q are compatible. Finally, as the final output bit of both p and q is given by $(w + Tb)_*$ when they don't abort, we have concluded that p and q return the same outcome for every global assignment, so that $\mathsf{MP}(e, \mathbb{Z}_2) \vDash q = p$. □

We apply Lemma 6.6 in the context of Theorem 5.11. The following lemma implements Step 2, condensing each protocol in the sum uniformly along the action of Q.

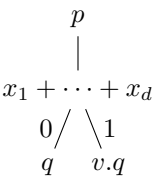
Lemma 6.7. *Let p be a measurement protocol with all nodes unary and $Q : \mathbb{Z}_2^l \to \mathbb{Z}_2^{s(p)+1}$ be a linear map where $l \geq m(p) + 2$. Denote by q the protocol produced by Lemma 6.6. Then*

$$\sum_{w \in \mathbb{Z}_2^l} Q(w).p \;=\; \sum_{w \in \mathbb{Z}_2^l} (\pi_{s(q)+1} \circ Q)(w).q \;.$$

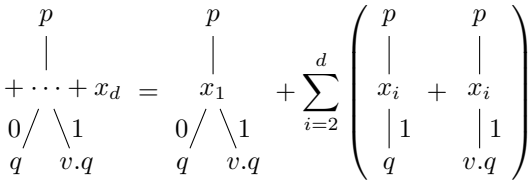
Proof. As the construction in Lemma 6.6 reads only the matrix and the order, with the default measurement settings carried along, it commutes with the action of the input in the sense that $Q(w).p = (\pi_{s(q)+1} \circ Q)(w).q$, giving the result. □

6.4. **Removing higher-arity nodes.** We now address Step 3 of the proof of Theorem 5.11: eliminating higher-arity branching nodes by expressing $\sum_w Q(w).p$ as a sum of smaller sums over protocols with all nodes unary. Each such reduction introduces new multiplicative nodes while correspondingly increasing the input dimension, so that the bound $l \geq m(p) + 2$ is maintained throughout.

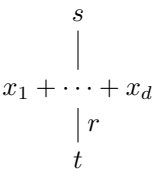
Lemma 6.8. *For any measurement protocol of the form*



we have

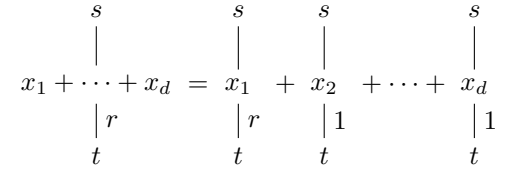
(9)


Proof. We first break the protocol into the sum of its branches using Lemma 6.3. Any branch of this protocol is of the form



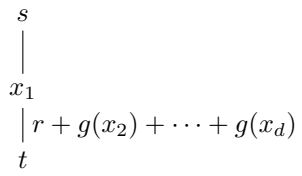
for some branches $s \in \mathrm{Br}(p)$ and $t \in \mathrm{Br}(q)$ and a bit $r \in \mathbb{Z}_2$.

We first argue that the following equation holds for each branch:

(10)


Observe that the above is an equation in a context of $\mathsf{MP}(B_{\mathcal{I}}, \mathbb{Z}_2)$ (in fact, all the measurements appearing in it, namely the x_i as well as those occurring in s and t, necessarily form a context of $B_{\mathcal{I}}$), so we can apply Proposition 5.9. For a global assignment $g : \coprod_{i \in \mathcal{I}} \mathbb{Z}_2 \to \mathbb{Z}_2$, both sides of eq. (10) have, under g, the same output

as the protocol

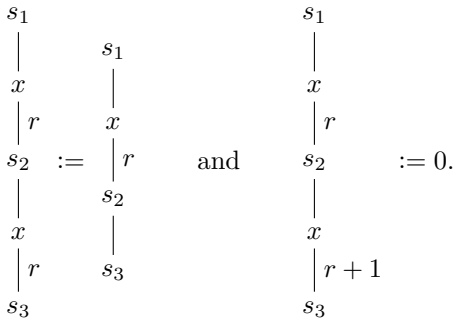


Thus eq. (10) holds for all global assignments, and hence is satisfied by the $\mathbb{Z}_2$-linear theory of $\mathsf{MP}(e, \mathbb{Z}_2)$ by Proposition 5.9.

The claim then follows by expressing the original protocol as a sum of branches using Lemma 6.3, applying eq. (10) to each branch, and then recollecting the terms. □

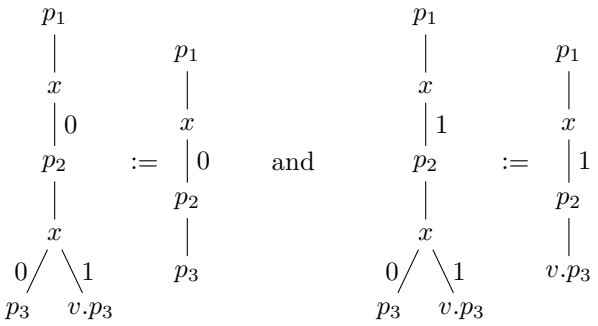
We pause to address a subtlety in the previous proof. Namely, by assumption, the left-hand sides of these equations denote well-defined protocols. However, the same is not obvious for terms appearing on the right-hand side. The issue has to do with repeated measurements: while the contexts appearing in a measurement protocol are distinct, they may share variables (allowing such overlaps is in fact crucial for the overall proof). When applying Lemma 6.8 to expand two different higher-arity branching levels of the protocol, it could happen that a variable x_i occurs as a unary node twice in the same protocol in the end result, which Definition 5.1 does not permit. Rather than introducing yet another generalisation of measurement protocols ("measurement protocols with potentially repeating nodes"), we treat each such expression unambiguously as shorthand for an honest protocol in the sense of Definition 5.1, as we now explain.

In a nutshell, the convention for branches is: given repeated occurrences of the same variable, we keep only the first one, provided they agree on the required outcome value, and otherwise the protocol equals zero. Slightly more formally, we assert the following equations as definitions: for a measurement x, branches s_1, s_2, s_3, and an outcome $r \in \mathbb{Z}_2$,

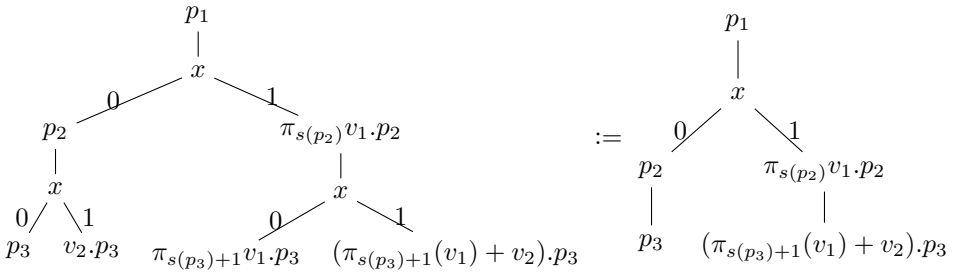


Informally, the conventions for protocols more complicated than branches follow from this: just decompose them into branches and apply this convention branch-wise. Formally, this is not fully rigorous without defining these protocols with repeated measurements and proving a version of Lemma 6.3 decomposing them into branches. Hence we extend our conventions to cover cases in which the repeated measurement x occurs once or twice as a branching node. When the first occurrence is multiplicative

and the second is branching, we assert the following equations as definitions:



Analogous equations are asserted when the first occurrence is branching and the second is multiplicative. When both occurrences are branching, only the first branching node remains as is, while the vectors that act on the continuation upon obtaining outcome 1 are added together (after mapping them to the correct space):



In the above, we think of the left-hand side as arising from an honest measurement protocol via Lemma 6.8, where v_1 (resp. v_2) is the vector affecting the continuation of the first (resp. second) occurrence of x. The admissibility requirement in Definition 5.1 ensures that the component of v_1 at $s(x)$ is zero.

What if, after repeated application of the equations above, we end up with more than two occurrences of a given variable in a single protocol? In some sense, the issue does not arise: if we apply Lemma 6.8 to an honest protocol, we end up with at most two occurrences of a given variable. Moreover, since the definitions above always prioritise the first occurrence, the order in which repeated pairs are handled does not matter (and in particular, neither does the order in which higher-arity nodes are eliminated via Lemma 6.8): each variable ultimately resolves to its earliest occurrence, much like the operation of projecting a pair to its first coordinate defines the associative operation of selecting the head of any non-empty list.

We apply Lemma 6.8 in the context of Theorem 5.11. The following two lemmas implement Step 3 of removing higher-arity branching nodes: the first eliminates a single higher-arity branching node, and Lemma 6.10 removes all such nodes by iteration.

Lemma 6.9. *Let p be a measurement protocol with all multiplicative nodes unary, and $Q : \mathbb{Z}_2^l \to \mathbb{Z}_2^{s(p)+1}$ be a linear map where $l \geq m(p) + 2$. If p has a d-ary branching node with $d > 1$, then there exist measurement protocols $p_1, \ldots, p_d$, each with $n(p_i) \leq n(p)$, strictly fewer higher-arity branching levels than p, and all multiplicative nodes unary, and linear maps $Q_i : \mathbb{Z}_2^{l_i} \to \mathbb{Z}_2^{s(p_i)+1}$ with $l_i \geq m(p_i) + 2$ for $i = 1, \ldots, d$ such that*

$$\sum_{w \in \mathbb{Z}_2^l} Q(w).p \;=\; \sum_{i=1}^{d} \sum_{w \in \mathbb{Z}_2^{l_i}} Q_i(w).p_i. \tag{11}$$

Proof. The key idea is to apply Lemma 6.8 across the action of Q. Fixing a d-ary level of p that measures $x_1 + \cdots + x_d$, we can write p as

$$\begin{array}{c} q \\ | \\ x_1 + \cdots + x_d \\ 0/\ \ \backslash 1 \\ q' \quad v.q' \end{array}$$

and use Lemma 6.8 to deduce

$$\begin{array}{c} q \\ | \\ x_1 + \cdots + x_d \\ 0/\ \ \backslash 1 \\ q' \quad v.q' \end{array} = \begin{array}{c} q \\ | \\ x_1 \\ 0/\ \ \backslash 1 \\ q' \quad v.q' \end{array} + \sum_{i=2}^{d} \left(\begin{array}{c} q \\ | \\ x_i \\ |1 \\ q' \end{array} + \begin{array}{c} q \\ | \\ x_i \\ |1 \\ v.q' \end{array} \right).$$

We then define p_1 and p_i for $i = 2, \ldots, d$ by setting

$$p_1 := \begin{array}{c} q \\ | \\ x_1 \\ 0/\ \ \backslash 1 \\ q' \quad v.q' \end{array} \qquad \text{and} \qquad p_i := \begin{array}{c} q \\ | \\ x_i \\ |1 \\ q' \end{array}.$$

We set $l_1 := l$ and $l_i := l+1$ for $i = 2, \ldots, d$. By construction, each p_i satisfies $n(p_i) \leq n(p)$, has strictly fewer higher-arity branching levels than p and all multiplicative nodes unary, and $l_i \geq m(p_i) + 2$.

It remains to define the linear maps Q_i and verify eq. (11). We set $Q_1 := \pi_{s(p_1)+1} \circ Q$ and for $i = 2, \ldots, d$ we set

$$Q_i := \left[\pi_{s(p_i)+1} \circ Q \quad \iota_{s(p_i)+1}(v) \right].$$

With these definitions, we now have that

$$p = p_1 + \sum_{i=2}^{d} (p_i + \iota_{s(p_i)+1}(v).p_i) = p_1 + \sum_{i=2}^{d} (p_i + Q_i(e_{l+1}).p_i)$$

where $e_{l+1} = (0, \ldots, 0, 1) \in \mathbb{Z}_2^{l+1}$. This decomposition does not depend on the initial choices of measurement settings, and hence goes through verbatim for $Q(w).p$. This then gives

$$Q(w).p = Q_1(w).p_1 + \sum_{i=2}^{d} \left(Q_i(\iota_{l+1}(w)).p_i + Q_i(\iota_{l+1}(w) + e_{l+1}).p_i \right).$$

As $\mathbb{Z}_2^{l+1} = \mathrm{Im}(\iota_{l+1}) \oplus \mathbb{Z}_2 e_{l+1}$, summing over all $w \in \mathbb{Z}_2^l$ gives (11), concluding the proof. □

Repeating the previous Lemma recursively until all branching levels are unary yields the following.

Lemma 6.10. *Let p be a measurement protocol with all multiplicative nodes unary, and $Q : \mathbb{Z}_2^l \to \mathbb{Z}_2^{s(p)+1}$ be a linear map where $l \geq m(p) + 2$. Then there are measurement protocols p_i with all nodes unary and $n(p_i) \leq n(p)$, and linear maps $Q_i : \mathbb{Z}_2^{l_i} \to \mathbb{Z}_2^{s(p_i)+1}$ with $l_i \geq m(p_i) + 2$ for $i = 1, \ldots, d$ such that*

$$\sum_{w \in \mathbb{Z}_2^l} Q(w).p \ = \sum_{i=1}^{d} \sum_{w \in \mathbb{Z}_2^{l_i}} Q_i(w).p_i.$$

7. (Cohomo)logical consequences

In this section we draw further consequences from our main result. In particular, we use it to answer the question raised by Robert Raussendorf [53, 54] on how to lift the cohomological criterion for contextuality to the adaptive case. We do so in two distinct cohomological frameworks for contextuality: the sheaf-theoretic (Čech) one of [10, 13] and the group-cohomological one of [48, 53, 56], in which Raussendorf originally posed his question. We treat each in turn, and then also examine the flattening construction through the lens of Raussendorf's de-iffification [54] of conditional AvN arguments. While the full details would take us outside the scope of this article, we sketch the main ideas here, which we intend to develop further in a follow-up paper.

7.1. **Čech-cohomological witnesses.** In the sheaf-theoretic framework of [10, 13, 24, 25], strong contextuality is witnessed by a non-vanishing Čech cohomology class. The following result states that an AvN argument over measurement protocols implies the existence of such a witness not just at the level of measurement protocols, but already over the original scenario.

To state the theorem, we need to introduce some notation not used elsewhere in this paper: if S is a scenario, we write $\mathsf{MP}(S)$ for the set of general measurement protocols over S as defined in [9], and $\mathsf{MP}_{\mathbb{Z}_2}(S)$ for the scenario of $\mathbb{Z}_2$-coarse-grained measurement protocols over S. Finally, for an empirical model e the notation $\mathsf{CSC}_{\mathbb{Z}_2}(e)$ means that e is cohomologically strongly contextual with coefficients in $\mathbb{Z}_2$, as defined in [10, 13] – that is, the strong contextuality of e is witnessed by a non-vanishing Čech cohomology class.

Theorem 7.1. *Let $S = (X, \Sigma, \mathbb{Z}_2)$ be a scenario and $e : S$ an empirical model. Then $\mathsf{AvN}_{\mathbb{Z}_2}(\mathsf{MP}_{\mathbb{Z}_2}(e)) \Rightarrow \mathsf{CSC}_{\mathbb{Z}_2}(e)$. That is, if the induced empirical model $\mathsf{MP}_{\mathbb{Z}_2}(e)$ on the scenario of $\mathbb{Z}_2$-coarse-grained measurement protocols exhibits AvN contextuality, then the original model e is cohomologically strongly contextual.*

(Proof sketch). We prove the chain of implications

$$\mathsf{AvN}_{\mathbb{Z}_2}(\mathsf{MP}_{\mathbb{Z}_2}(e)) \Rightarrow \mathsf{CSC}_{\mathbb{Z}_2}(\mathsf{MP}_{\mathbb{Z}_2}(e)) \Rightarrow \mathsf{CSC}_{\mathbb{Z}_2}(\mathsf{MP}(e)) \Rightarrow \mathsf{CSC}_{\mathbb{Z}_2}(e)$$

The first implication is [10, Theorem 21]. The second follows from the existence of a canonical map of scenarios $\mathsf{MP}(S) \to \mathsf{MP}_{\mathbb{Z}_2}(S)$ (in the sense of [9]), together with the fact that cohomology cooperates suitably with maps of scenarios.

Finally, the vanishing of the cohomological obstruction to extending a local section of e corresponds to a certain cochain being a coboundary in the relative cohomology. Concretely, this amounts to a compatible family of $\mathbb{Z}_2$-linear combinations of local sections of e extending the given local section, where compatibility is an analogue of no-signalling [3, 13]. Therefore, just as an empirical model $e : S$ extends canonically to an empirical model $\mathsf{MP}(e) : \mathsf{MP}(S)$, such a family extends to corresponding data for $\mathsf{MP}(e)$, witnessing the vanishing of the cohomological obstruction over $\mathsf{MP}(e)$. Taking contrapositives gives the final implication. □

The part of this argument requiring significant elaboration is the verification of the functoriality and naturality properties that ensure cohomology behaves well with respect to maps of scenarios. While essentially straightforward, the details are somewhat lengthy.

By the theorem above, AvN contextuality of $\mathsf{MP}(e)$ implies the non-vanishing of the Čech-cohomological invariant on e. Combining this with our main result (Theorem 3.6) yields an answer to Raussendorf's question within the sheaf-cohomological framework: cohomological witnesses for contextuality extend to adaptive MBQC

protocols. Moreover, in this framework, the witness lives already over the original scenario.

7.2. **Group-cohomological witnesses.** Raussendorf originally formulated his question in a different cohomology framework [48, 53, 56] from the one considered above, based on group cohomology. Those group-cohomological witnesses were formulated for sets of Pauli/Weyl operators closed under commuting products. Aasnæss [1, 2] made a direct comparison between the two frameworks, first generalising the group-cohomological approach to a more abstract setting and then showing that a non-trivial group-cohomological obstruction implies a non-trivial Čech obstruction. Note that this implication runs from group-cohomological to Čech-cohomological obstructions, so Theorem 7.1, which establishes non-triviality of the Čech cohomology class, does not directly translate to group-cohomological witnesses; we can nonetheless derive such witnesses for $\mathsf{MP}(e)$ from our main result, as we now describe.

The abstract formulation in [2] is in terms of G-bundle scenarios which are scenarios whose measurement set carry algebraic structure: each context is a commutative monoid on which the group G of outcomes acts compatibly. We briefly describe how the measurement scenarios we consider can be interpreted as such, and draw the corresponding consequences from our main result. The key idea is that AvN arguments are formulated for scenarios with algebraic structure on the *outcomes*; the move is to lift that structure to the *measurements* themselves. Recall the setting of Section 2.2: a measurement scenario $S = (X, \Sigma, O)$ where each measurement has outcomes valued in the abelian group $O = \mathbb{Z}_2$. [11] The abelian group structure of $\mathbb{Z}_2$ can be lifted from outcomes to measurements, giving rise to a partial algebraic structure. Operationally, given two compatible $\mathbb{Z}_2$-valued measurements x and y, one may form a new $\mathbb{Z}_2$-valued measurement $x + y$ by jointly measuring x and y and adding their outcomes. We have in fact been using this operation implicitly throughout, since it underlies the higher-arity nodes that arise in the condensing construction of Section 6.3. One may also adjoin, for each $a \in \mathbb{Z}_2$, a constant measurement c_a that always returns outcome a. The measurement $x + y$ is itself compatible with both x and y, while each c_a is compatible with every measurement, so any context may be closed under these operations (binary addition and the two constants).

Under such closure and identification of operationally indistinguishable measurement procedures, each context gives rise to a *pointed Boolean group* $G(C)$: an abelian group in which every element has order 2, with neutral element c_0, with an additional distinguished element c_1. Equivalently, a Boolean group can be thought of as a $\mathbb{Z}_2$-vector space and the distinguished element singles out a vector, corresponding to moving to an affine space. [12] This same structure admits an equivalent description via a $\mathbb{Z}_2$-action $\theta : \mathbb{Z}_2 \times G(C) \to G(C)$, required to be a monoid homomorphism (as per [2, Definitions 3.1.4 and 3.3.1]), or equivalently, to satisfy $\theta(a,\, m + m') \;=\; m + \theta(a,\, m')$ for all $a \in \mathbb{Z}_2$ and $m, m' \in G(C)$ (as per [4]). Note that this condition forces the action to be translation by constants (which form an image of $\mathbb{Z}_2$ inside $G(C)$): setting $m' = c_0$ gives $\theta(a, m) = m + \theta(a, c_0)$ for all m, so θ is entirely determined by constants $c_a := \theta(a, c_0)$. Operationally, the

[11] More generally, as per footnote 7, one may take the outcome set O to be an arbitrary abelian group, or even any R-module over a ring R, yielding analogous constructions; see Remark 7.2.

[12] From a logical perspective, this structure captures the linear fragment of a Boolean algebra: the group addition is exclusive-or, the constants c_0 and c_1 are the constants false and true, with negation being recovered as $m \mapsto m + c_1$; the multiplicative connectives (conjunction and disjunction) are absent.

non-trivial element $1 \in \mathbb{Z}_2$ acts as $m \mapsto m + c_1$, flipping the outcome (that is, the negation of the logical interpretation in footnote 12).

These pointed Boolean groups $G(C)$ on each context patch together into a *partial pointed Boolean group* $G(S)$, a partial algebra in the same spirit as partial Boolean algebras [5, 41]. Namely, a partial pointed Boolean group is a set G equipped with a reflexive symmetric (*commeasurability*) relation $\odot$, a partial binary (*addition*) operation $+ : \odot \to G$, and constants $c_0, c_1 \in G$, such that every set of pairwise commeasurable elements extends to a set of pairwise commeasurable elements forming a (total) pointed Boolean group under restriction of the operations. [13] The $\mathbb{Z}_2$-action extends compatibly to this partial algebraic structure, making $G(S)$ a $\mathbb{Z}_2$-bundle scenario in the sense of [2, Chapter 3].

This structure $G(S)$ is the free partial pointed Boolean group generated by the scenario S. The linear theory of an empirical model $e : S$ can now be incorporated into this structure. Recall that a linear equation over a context C has the form $\sum_{x \in D} x = a$ where $D \subseteq C$ and $a \in \mathbb{Z}_2$ (see Section 2.2). Since all $\iota(x)$ for $x \in D$ are compatible in $G(S)$ (because D is a context of S), this equation makes sense as an equation between elements of $G(S)$: $\sum_{x \in D} \iota(x) = c_a$. Quotienting $G(S)$ by the linear theory of e thus yields a further partial pointed Boolean group $G(S, e) := G(S)/\operatorname{Th}(e)$.

A global assignment $g : X \to \mathbb{Z}_2$ extends uniquely to a homomorphism $G(S) \to \mathbb{Z}_2$ of partial pointed Boolean groups by freeness, and it satisfies the linear theory of e precisely when this homomorphism factors through $G(S, e)$. Therefore, the model e is $\mathsf{AvN}_{\mathbb{Z}_2}$-contextual, i.e. its $\mathbb{Z}_2$-linear theory is inconsistent, if and only if no such homomorphism $\tilde{g} : G(S, e) \to \mathbb{Z}_2$ exists. This is equivalent to saying that the sequence

$$\mathbb{Z}_2 \xhookrightarrow{c} G(S, e) \xrightarrow{[\cdot]_\theta} G(S, e)/\mathbb{Z}_2 \tag{12}$$

where c is the inclusion $a \mapsto c_a$ and $[\cdot]_\theta$ is the quotient by the $\mathbb{Z}_2$-translation action, has no left splitting, that is, no homomorphism $\tilde{g} : G(S, e) \to \mathbb{Z}_2$ of partial pointed Boolean groups satisfying $\tilde{g} \circ c = \mathrm{id}_{\mathbb{Z}_2}$. Since any homomorphism of partial pointed Boolean groups to $\mathbb{Z}_2$ sends $c_a \mapsto a$ automatically, this is precisely a satisfying assignment for the linear theory of e. Non-existence of such a splitting is exactly a cohomological obstruction in the sense of [2]. [14]

Applying this to $\mathsf{MP}_{\mathbb{Z}_2}(e)$: our main result (Theorem 3.6) produces an AvN argument over $\mathsf{MP}_{\mathbb{Z}_2}(e)$, which by the above witnesses a non-trivial group-cohomological obstruction for $G(\mathsf{MP}_{\mathbb{Z}_2}(S), \mathsf{MP}_{\mathbb{Z}_2}(e))$. In contrast to the Čech-cohomological case (Theorem 7.1), this obstruction does *not* propagate back to a group-cohomological obstruction for $G(S, e)$: the reason is the same as discussed there, namely that our construction yields an AvN argument over measurement protocols, not necessarily over the original scenario.

---

[13] Concretely, this means that the operations (nullary c_0, c_1 and binary $+$) respect commeasurability – $c_0 \odot m$ and $c_1 \odot m$ for every m, and $m_1 + m_2 \odot m'$ whenever m_1, m_2, m' are pairwise commeasurable – and that the axioms of pointed Boolean groups hold among any pairwise commeasurable elements.

The partial pointed Boolean group $G(S)$ associated to a measurement scenario $S = (X, \Sigma, O)$ is the free partial pointed Boolean group on the reflexive graph of measurement compatibility: generated by $\iota(x)$ for $x \in X$, subject to $\iota(x) \odot \iota(y)$ whenever $\{x, y\} \in \Sigma$. The explicit inductive construction of the free partial Boolean algebra on a reflexive graph given in [5, 6] adapts directly to this setting, giving the free partial pointed Boolean group.

[14] This property is phrased as a splitting of the whole sequence (12), rather than as a retraction of c alone: a version of the splitting lemma [2, Lemma 3.1.1] establishes a bijection between such left splittings and *trivialisations* of the bundle $[\cdot]_\theta : G(S, e) \twoheadrightarrow G(S, e)/\mathbb{Z}_2$. It is the (non-)existence of a trivialisation of the bundle that the group-cohomological obstruction of [2] measures.

Remark 7.2. The analysis above generalises beyond the $\mathbb{Z}_2$ setting considered here (see footnotes 7 and 11). For O an arbitrary abelian group, the construction above goes through almost verbatim: one adjoins constants c_a for each element $a \in O$, and closes each context under pairwise addition and constants, exactly as before. This yields a partial abelian group $G(S)$ containing, via the homomorphism $a \mapsto c_a$, an image of O commeasurable with every element of $G(S)$. The O-action is again given by translation $m \mapsto m + c_a$. This makes $G(S)$ an O-bundle scenario in the sense of [2, Chapter 3]; the analysis above carries through in this setting.

More generally, if O is an R-module over a commutative ring R, one lifts not only module addition and the constants c_a for elements of O, but additionally closes each context under scalar multiplication by ring elements, equipping $G(S)$ with a total, unary operation $r\cdot : G(S) \to G(S)$ for each $r \in R$. $G(S)$ then becomes a partial R-module containing an image of O as a distinguished R-linear submodule, commeasurable with every element. A similar analysis can be carried through in this setting.

7.3. **De-iffifying adaptive AvN arguments.** Our flattening construction also admits a logical perspective that sharpens the picture above. Adaptive protocols allow for systems of *conditional* equations providing AvN arguments, where the conditioning on variables corresponds to the adaptivity. A concrete example of such a system of conditional equations appears in [7]. Such arguments have been dubbed "iffy proofs" by Raussendorf [54]. Thus we can interpret our flattening construction as a process of "de-iffification".

The situation is, however, somewhat more subtle: arbitrary conditional systems can capture arbitrary supports, hence in that setting inconsistency coincides with strong contextuality. However, Theorem 7.1 shows that all-versus-nothing contextuality over measurement protocols is at most as strong as Čech-cohomological contextuality. This leads to a further consequence: $\mathsf{AvN}_{\mathbb{Z}_2}(e)$ is *not* implied by $\mathsf{AvN}_{\mathbb{Z}_2}(\mathsf{MP}_{\mathbb{Z}_2}(e))$, and therefore not by $\mathsf{CSC}_{\mathbb{Z}_2}(e)$ either. In particular, the adaptive AvN arguments produced by our construction are genuinely new: they go beyond what non-adaptive AvN arguments on the underlying scenario can detect. This is witnessed by the iffy proof in [7, Section 5], since one can verify that the ordinary $\mathbb{Z}_2$-linear theory of the underlying model in that example is consistent.

By Section 7.2, the same AvN argument also witnesses a non-trivial group-cohomological obstruction for $G(\mathsf{MP}_{\mathbb{Z}_2}(S), \mathsf{MP}_{\mathbb{Z}_2}(e))$. Unlike the Čech case, this cannot in general propagate back to a group-cohomological obstruction for $G(S, e)$ itself: such an obstruction would amount to $\mathsf{AvN}_{\mathbb{Z}_2}(e)$, which the iffy proof example just discussed already shows is too strong.

8. Further directions

We close by outlining some directions for further work suggested by our results:

- Does the converse to Theorem 7.1 hold? That is, does (Čech-)cohomological strong contextuality of e imply that $\mathsf{MP}_{\mathbb{Z}_2}(e)$ exhibits algebraic contextuality? Equivalently, does every cohomological proof of strong contextuality give rise to an AvN argument over measurement protocols?
- Our construction produces AvN arguments for the scenario $\mathsf{MP}(B_{\mathcal{I}}, \mathbb{Z}_2)$ of measurement protocols. What do these translate to when expressed in terms of the original measurements of $B_{\mathcal{I}}$? More precisely, which conditional AvN arguments ('iffy proofs', in the sense of [7, 54]) on $B_{\mathcal{I}}$ correspond to the AvN arguments our construction produces?
- In [11], a quantitative refinement of Raussendorf's result was established, with an inequality relating the success probability of an MBQC, the nonlinearity of the computed Boolean function, and the contextual fraction (a measure of

contextuality) of the underlying empirical model. Can an analogous quantitative refinement of our main result be developed for adaptive protocols?

- Here we focussed on dichotomic measurements with outcomes in $\mathbb{Z}_2$, typically quantum-realised using qubits. A natural extension considers measurements with outcomes in $\mathbb{Z}_d$ for $d > 2$, corresponding to qudit-based MBQC. Contextuality and AvN arguments have been studied for that setting in [28–30]. Do analogous results hold for adaptive $\mathbb{Z}_d$-linear protocols?
- In [8], there is a complete classification of all quantum AvN arguments for stabiliser states. Does an analogous classification exist for adaptive AvN arguments of the kind produced by our construction?
- We have focussed on algebraic AvN paradoxes. How do these relate to the logical paradoxes (contradictions verified in partial Boolean algebras) studied by Kochen and Specker in their seminal work on contextuality [5, 41]? In particular, how can one capture adaptive measurements on partial Boolean algebras? Can an analogue of our flattening construction be developed in that setting?
- The scenario $\mathsf{MP}(B_{\mathcal{I}}, \mathbb{Z}_2)$ is constructed from $B_{\mathcal{I}}$. Can this be phrased for arbitrary algebraic contextuality scenarios, rather than just those of Bell type? Is this construction functorial with respect to morphisms of scenarios, which give rise to simulations between empirical models [18, 37], and does our main result interact naturally with such morphisms?
- Building on the previous point: in [9] a more general notion of measurement protocol was structured as a comonad on a category of empirical models. Can the version of MP studied here be given a similar comonadic structure, possibly graded by the depth or other parameters of the protocol?
- Throughout this paper we have worked within the MBQC paradigm. Can our results be related to the circuit model of quantum computation, perhaps via the notion of sequential contextuality [44]?
- Our construction yields explicit AvN witnesses for adaptive protocols. Can these be used for self-testing [45, 60] of quantum resources, certifying not only the contextuality but the specific quantum state and measurements? There is reason to expect so: AvN arguments, and the closely related linear constraint system games [26] (which are in a sense a state-independent analogue of AvN arguments), have already enabled self-testing in the non-adaptive setting [27, 36], including scalable self-testing of graph states [16] and CSS codewords [34]. Whether adaptive AvN witnesses admit analogous results, where the feed-forward structure itself must be certified, is an open problem.
- There is a connection between contextuality and quantum error correction: [38] show that, given any subsystem stabiliser code with two or more gauge qubits, the closure of its measurements as a partial abelian group yields a strong contextuality argument. Such contextuality, in the spirit of Kirby and Love [39], is algebraic in nature, albeit state-independent. They further show that code-switching schemes for universal fault tolerance – such as the doubled colour codes of Bravyi and Cross [21] – inherit such contextuality. However, their argument is *static*, concerning the fixed algebraic structure of the code-switching gadget as a whole, rather than the protocol's adaptive unfolding. Could the techniques developed in this work deepen this connection by witnessing contextuality directly in the code-switching protocol itself?
- Contextuality has been used for quantum hardware benchmarking: in [31], AvN-type constraints are exploited for error detection beyond standard randomised benchmarking. Our result provides explicit AvN witnesses for any adaptive $\mathbb{Z}_2$-linear MBQC computing a non-affine function. Can such witnesses be used

to certify or benchmark the correct functioning of adaptive quantum hardware, where one must account for the feed-forward structure?

Acknowledgements

This work was supported by the Horizon Europe project FoQaCiA, *Foundations of Quantum Computational Advantage*, GA no. 101070558 (UCL and INL); the Engineering and Physical Sciences Research Council (EPSRC, U.K.) fellowship *Resources in Computation*, EP/V040944/1 (UCL); and the Fundação para a Ciência e Tecnologia (FCT, Portugal) project KaleidosQope, *Contextual partial views: a logical foundation for quantum computational advantage*, 2023.13603.PEX (INL).

References

[1] Sivert Aasnæss. Cohomology and the algebraic structure of contextuality in measurement based quantum computation. In Bob Coecke and Matthew Leifer, editors, *16th International Conference on Quantum Physics and Logic (QPL 2019)*, volume 318 of *Electronic Proceedings in Theoretical Computer Science*, pages 242–253. Open Publishing Association, May 2020. doi:10.4204/EPTCS.318.15.

[2] Sivert Aasnæss. *Comparing two cohomological obstructions for contextuality, and a generalised construction of quantum advantage with shallow circuits*. DPhil thesis, University of Oxford, 2021. URL: https://ora.ox.ac.uk/objects/uuid:a31ce72c-3d5e-4ca2-afa3-a38afbb93833/, arXiv:2212.09382 [quant-ph].

[3] Samson Abramsky. Contextuality: At the borders of paradox. In Elaine Landry, editor, *Categories for the Working Philosopher*, chapter 11, pages 262–285. Oxford University Press, November 2017. doi:10.1093/oso/9780198748991.003.0011.

[4] Samson Abramsky. Contextuality, algebraic structure, and state-independence. In Luca Reggio, Sam van Gool, and Wesley Fussner, editors, *Mai Gehrke on Logic, Algebra and Duality*, Outstanding Contributions to Logic. Springer, 2026. To appear.

[5] Samson Abramsky and Rui Soares Barbosa. The logic of contextuality. In Christel Baier and Jean Goubault-Larrecq, editors, *29th EACSL Annual Conference on Computer Science Logic (CSL 2021)*, volume 183 of *Leibniz International Proceedings in Informatics (LIPIcs)*, pages 5:1–5:18. Schloss Dagstuhl – Leibniz-Zentrum für Informatik, January 2021. doi:10.4230/LIPIcs.CSL.2021.5.

[6] Samson Abramsky and Rui Soares Barbosa. Contextuality in logical form: Duality for transitive partial CABAs. Talk presented at 10th Conference on Topology, Algebra, and Categories in Logic (TACL 2022), June 2022. URL: https://www.mat.uc.pt/~tacl2022/pdfs/TACL_2022_paper_19.pdf.

[7] Samson Abramsky, Rui Soares Barbosa, Giovanni Carù, Nadish de Silva, Kohei Kishida, and Shane Mansfield. Minimum quantum resources for strong non-locality. In Mark M Wilde, editor, *12th Conference on the Theory of Quantum Computation, Communication and Cryptography (TQC 2017)*, volume 73 of *Leibniz International Proceedings in Informatics (LIPIcs)*, pages 9:1–9:20. Schloss Dagstuhl – Leibniz-Zentrum für Informatik, March 2018. doi:10.4230/LIPICS.TQC.2017.9.

[8] Samson Abramsky, Rui Soares Barbosa, Giovanni Carù, and Simon Perdrix. A complete characterization of all-versus-nothing arguments for stabilizer states. *Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences*, 375(2106):20160385, November 2017. doi:10.1098/rsta.2016.0385.

[9] Samson Abramsky, Rui Soares Barbosa, Martti Karvonen, and Shane Mansfield. A comonadic view of simulation and quantum resources. In *34th Annual ACM/IEEE Symposium on Logic in Computer Science (LICS 2019)*, pages 23:1–23:12. IEEE, August 2019. doi:10.1109/LICS.2019.8785677.

[10] Samson Abramsky, Rui Soares Barbosa, Kohei Kishida, Raymond Lal, and Shane Mansfield. Contextuality, cohomology and paradox. In Stephan Kreutzer, editor, *24th EACSL Annual Conference on Computer Science Logic (CSL 2015)*, volume 41 of *Leibniz International Proceedings in Informatics (LIPIcs)*, pages 211–228. Schloss Dagstuhl – Leibniz-Zentrum für Informatik, September 2015. doi:10.4230/LIPIcs.CSL.2015.211.

[11] Samson Abramsky, Rui Soares Barbosa, and Shane Mansfield. Contextual fraction as a measure of contextuality. *Physical Review Letters*, 119(5):050504, August 2017. doi:10.1103/PhysRevLett.119.050504.

[12] Samson Abramsky and Adam Brandenburger. The sheaf-theoretic structure of non-locality and contextuality. *New Journal of Physics*, 13(11):113036, November 2011. doi:10.1088/1367-2630/13/11/113036.

[13] Samson Abramsky, Shane Mansfield, and Rui Soares Barbosa. The cohomology of non-locality and contextuality. In Bart Jacobs, Peter Selinger, and Bas Spitters, editors, *8th International Workshop on Quantum Physics and Logic (QPL 2011)*, volume 95 of *Electronic Proceedings in Theoretical Computer Science*, pages 1–14. Open Publishing Association, October 2012. doi:10.4204/EPTCS.95.1.

[14] Janet Anders and Dan E. Browne. Computational power of correlations. *Physical Review Letters*, 102(5):050502, February 2009. doi:10.1103/PhysRevLett.102.050502.

[15] Alex Arkhipov. Extending and characterizing quantum magic games, September 2012. arXiv:1209.3819 [quant-ph].

[16] Flavio Baccari, Remigiusz Augusiak, Ivan Šupić, Jordi Tura, and Antonio Acín. Scalable Bell inequalities for qubit graph states and robust self-testing. *Physical Review Letters*, 124(2):020402, January 2020. doi:10.1103/PhysRevLett.124.020402.

[17] Rui Soares Barbosa. *Contextuality in quantum mechanics and beyond*. DPhil thesis, University of Oxford, 2015.

[18] Rui Soares Barbosa, Martti Karvonen, and Shane Mansfield. Closing Bell: Boxing black box simulations in the resource theory of contextuality. In Alessandra Palmigiano and Mehrnoosh Sadrzadeh, editors, *Samson Abramsky on Logic and Structure in Computer Science and Beyond*, volume 25 of *Outstanding Contributions to Logic*, chapter 13, pages 475–529. Springer, August 2023. doi:10.1007/978-3-031-24117-8_13.

[19] John S Bell. On the Einstein Podolsky Rosen paradox. *Physics Physique Fizika*, 1(3):195–200, November 1964. doi:10.1103/PhysicsPhysiqueFizika.1.195.

[20] John S Bell. On the problem of hidden variables in quantum mechanics. *Reviews of Modern Physics*, 38(3):447–452, July 1966. doi:10.1103/RevModPhys.38.447.

[21] Sergey Bravyi and Andrew Cross. Double color codes, September 2015. arXiv:1509.03239 [quant-ph].

[22] Sergey Bravyi, David Gosset, and Robert König. Quantum advantage with shallow circuits. *Science*, 362(6412):308–311, October 2018. doi:10.1126/science.aar3106.

[23] Hans J. Briegel, Dan E. Browne, Wolfgang Dür, Robert Raussendorf, and Maarten van den Nest. Measurement-based quantum computation. *Nature Physics*, 5(1):19–26, January 2009. doi:10.1038/nphys1157.

[24] Giovanni Carù. On the cohomology of contextuality. In Ross Duncan and Chris Heunen, editors, *13th International Conference on Quantum Physics and Logic (QPL 2016)*, volume 236 of *Electronic Proceedings in Theoretical Computer Science*, pages 21–39. Open Publishing Association, January 2017. doi:10.4204/EPTCS.236.2.

[25] Giovanni Carù. Towards a complete cohomology invariant for non-locality and contextuality, July 2018. arXiv:1807.04203 [quant-ph].

[26] Richard Cleve and Rajat Mittal. Characterization of binary constraint system games. In Javier Esparza, Pierre Fraigniaud, Thore Husfeldt, and Elias Koutsoupias, editors, *41st International Colloquium on Automata, Languages, and Programming (ICALP 2014)*, volume 8572 of *Lecture Notes in Computer Science*, pages 320–331. Springer, June 2014. doi:10.1007/978-3-662-43948-7_27.

[27] Andrea Coladangelo and Jalex Stark. Robust self-testing for linear constraint system games, September 2017. arXiv:1709.09267 [quant-ph].

[28] Markus Frembs, Cihan Okay, and Ho Yiu Chung. No quantum solutions to linear constraint systems from monomial measurement-based quantum computation in odd prime dimension, January 2026. arXiv:2601.11367 [quant-ph].

[29] Markus Frembs, Sam Roberts, and Stephen D Bartlett. Contextuality as a resource for measurement-based quantum computation beyond qubits. *New Journal of Physics*, 20(10):103011, October 2018. doi:10.1088/1367-2630/aae3ad.

[30] Stefano Gogioso and William Zeng. Generalised Mermin-type non-locality arguments. *Logical Methods in Computer Science*, 15(2):3:1–3:51, April 2019. doi:10.23638/lmcs-15(2:3)2019.

[31] Google Quantum AI (Shashwat Kumar, Eliott Rosenberg, Alejandro Grajales Dau, Rodrigo Cortiñas, et al.). Quantum-classical separation in bounded-resource tasks arising from measurement contextuality, December 2025. arXiv:2512.02284 [quant-ph].

[32] Daniel M Greenberger, Michael A Horne, and Anton Zeilinger. Going beyond Bell's theorem. In Menas Kafatos, editor, *Bell's Theorem, Quantum Theory and Conceptions of the Universe*, volume 37 of *Fundamental Theories of Physics*, pages 69–72. Springer, October 1989. doi:10.1007/978-94-017-0849-4_10.

[33] Lucien Hardy. Nonlocality for two particles without inequalities for almost all entangled states. *Physical Review Letters*, 71(11):1665–1668, September 1993. doi:10.1103/PhysRevLett.71.1665.

[34] Oliver Hart, David T Stephen, Evan Wickenden, and Rahul Nandkishore. Many-body contextuality and self-testing quantum matter via nonlocal games, December 2025. arXiv:2512.16886 [quant-ph].

[35] Mark Howard, Joel Wallman, Victor Veitch, and Joseph Emerson. Contextuality supplies the 'magic' for quantum computation. *Nature*, 510(7505):351–355, June 2014. doi:10.1038/nature13460.

[36] Jędrzej Kaniewski. Analytic and nearly optimal self-testing bounds for the Clauser-Horne-Shimony-Holt and Mermin inequalities. *Physical Review Letters*, 117(7):070402, August 2016. doi:10.1103/PhysRevLett.117.070402.

[37] Martti Karvonen. Categories of empirical models. In Peter Selinger and Giulio Chiribella, editors, *15th International Conference on Quantum Physics and Logic (QPL 2018)*, volume 287 of *Electronic Proceedings in Theoretical Computer Science*, pages 239–252. Open Publishing Association, January 2019. doi:10.4204/EPTCS.287.14.

[38] Derek Khu, Andrew Tanggara, Chao Jin, and Kishor Bharti. Contextuality of quantum error-correcting codes. *PRX Quantum*, 7(1):010319, January 2026. doi:10.1103/7zb5-vs4x.

[39] William M. Kirby and Peter J. Love. Contextuality test of the nonclassicality of variational quantum eigensolvers. *Physical Review Letters*, 123(20):200501, November 2019. doi:10.1103/PhysRevLett.123.200501.

[40] William M Kirby and Peter J Love. Classical simulation of noncontextual Pauli Hamiltonians. *Physical Review A*, 102(3):032418, September 2020. doi:10.1103/PhysRevA.102.032418.

[41] Simon Kochen and Ernst P Specker. The problem of hidden variables in quantum mechanics. *Journal of Mathematics and Mechanics*, 17(1):59–87, July 1967. doi:10.1512/iumj.1968.17.17004.

[42] Benjamin P Lanyon, Petar Jurcevic, Michael Zwerger, Cornelius Hempel, Esteban A Martinez, Wolfgang Dür, Hans J Briegel, Rainer Blatt, and Christian F Roos. Measurement-based quantum computation with trapped ions. *Physical Review Letters*, 111(21):210501, November 2013. doi:10.1103/PhysRevLett.111.210501.

[43] Harashta Tatimma Larasati and Byung-Soo Choi. Circuit-based vs. measurement-based quantum computing: a comparative analysis, layered metrics, and decision flow for approach selection. *EPJ Quantum Technology*, 13(1):39, February 2026. doi:10.1140/epjqt/s40507-026-00483-1.

[44] Shane Mansfield and Elham Kashefi. Quantum advantage from sequential-transformation contextuality. *Physical Review Letters*, 121(23):230401, December 2018. doi:10.1103/PhysRevLett.121.230401.

[45] Dominic Mayers and Andrew Yao. Self testing quantum apparatus. *Quantum Information and Computation*, 4(4):273–286, July 2004. doi:10.26421/QIC4.4-3.

[46] N David Mermin. Extreme quantum entanglement in a superposition of macroscopically distinct states. *Physical Review Letters*, 65(15):1838–1840, October 1990. doi:10.1103/PhysRevLett.65.1838.

[47] N David Mermin. Simple unified form for the major no-hidden-variables theorems. *Physical Review Letters*, 65(27):3373–3376, December 1990. doi:10.1103/PhysRevLett.65.3373.

[48] Cihan Okay, Sam Roberts, Stephen D. Bartlett, and Robert Raussendorf. Topological proofs of contextuality in quantum mechanics. *Quantum Information and Computation*, 17(13 & 14):1135–1166, November 2017. doi:10.26421/QIC17.13-14-5.

[49] Asher Peres. Incompatible results of quantum measurements. *Physics Letters A*, 151(3):107–108, December 1990. doi:10.1016/0375-9601(90)90172-K.

[50] Sandu Popescu and Daniel Rohrlich. Quantum nonlocality as an axiom. *Foundations of Physics*, 24(3):379–385, March 1994. doi:10.1007/BF02058098.

[51] Robert Prevedel, Philip Walther, Felix Tiefenbacher, Pascal Böhi, Rainer Kaltenbaek, Thomas Jennewein, and Anton Zeilinger. High-speed linear optics quantum computing using active feed-forward. *Nature*, 445(7123):65–69, January 2007. doi:10.1038/nature05346.

[52] Robert Raussendorf. Contextuality in measurement-based quantum computation. *Physical Review A*, 88(2):022322, August 2013. doi:10.1103/PhysRevA.88.022322.

[53] Robert Raussendorf. Cohomological framework for contextual quantum computations. *Quantum Information and Computation*, 19(13 & 14):1141–1170, November 2019. doi:10.26421/qic19.13-14-4.

[54] Robert Raussendorf. Putting paradoxes to work: Contextuality in measurement-based quantum computation. In Alessandra Palmigiano and Mehrnoosh Sadrzadeh, editors, *Samson Abramsky on Logic and Structure in Computer Science and Beyond*, volume 25 of

Outstanding Contributions to Logic, chapter 16, pages 595–622. Springer, August 2023. doi:10.1007/978-3-031-24117-8_16.

[55] Robert Raussendorf and Hans J Briegel. A one-way quantum computer. *Physical Review Letters*, 86(22):5188–5191, May 2001. doi:10.1103/PhysRevLett.86.5188.

[56] Robert Raussendorf, Cihan Okay, Michael Zurel, and Polina Feldmann. The role of cohomology in quantum computation with magic states. *Quantum*, 7:979, April 2023. doi:10.22331/q-2023-04-13-979.

[57] Amy J Searle. *Correlation and combinatorics: causal contextuality and spin systems*. DPhil thesis, University of Oxford, 2024. URL: https://ora.ox.ac.uk/objects/uuid:7f3e70f3-5136-4a64-89e8-858034004181.

[58] Amy J Searle, Rui Soares Barbosa, and Samson Abramsky. Mapping temporal correlations to contextuality correlations. Talk presented at 21st International Conference on Quantum Physics and Logic (QPL 2024), July 2024. URL: https://qpl2024.dc.uba.ar/abstracts/100.pdf.

[59] William Slofstra and Lu-Ming Zhang. Operator solutions of linear systems and small cancellation, December 2024. arXiv:2412.10305 [quant-ph].

[60] Ivan Šupić and Joseph Bowles. Self-testing of quantum systems: a review. *Quantum*, 4:337, September 2020. doi:10.22331/q-2020-09-30-337.

[61] Philip Walther, Kevin J Resch, Terry Rudolph, Emmanuel Schenck, Harald Weinfurter, Vlatko Vedral, Markus Aspelmeyer, and Anton Zeilinger. Experimental one-way quantum computing. *Nature*, 434(7030):169–176, March 2005. doi:10.1038/nature03347.

[62] Wanbing Zhao, H W Shawn Liew, Wen Wei Ho, Chunxiao Liu, and Vir B Bulchandani. Scalable tests of quantum contextuality from stabilizer-testing nonlocal games, December 2025. arXiv:2512.16654 [quant-ph].

Samson Abramsky
Department of Computer Science, University College London
London, United Kingdom
Email address: s.abramsky@ucl.ac.uk

Rui Soares Barbosa
International Iberian Nanotechnology Laboratory
Braga, Portugal
Email address: rui.soaresbarbosa@inl.int

Carmen Constantin
Department of Computer Science, University College London
London, United Kingdom
Email address: c.constantin@ucl.ac.uk

Martti Karvonen
Department of Computer Science, University College London
London, United Kingdom
From August 2026: Department of Computer Science, University of Bath
Bath, United Kingdom
Email address: martti.karvonen@ucl.ac.uk