Causal inequalities witness non-stabilizerness

Leonardo Vaglini[1], Nasra Daher Ahmed[1], and Ravi Kunjwal[1]

[1]Aix-Marseille University, CNRS, LIS, Marseille, France

This is an extended abstract for a non-proceedings talk. The paper is not publicly available, but a draft is attached to this abstract

Introduction

The Gottesman-Knill theorem [1,5,9] sets a boundary in quantum computation theory that must be overcome in order to achieve a quantum speedup over classical computers. It establishes that any quantum computation performed using stabilizer operations (SO), i.e. operations built out of Clifford operations, stabilizer states preparation and Pauli measurements performed adaptively, can be efficiently classically simulated. Therefore, observing a quantum advantage requires the use of non-Clifford gates, such as the Toffoli gate, or the preparation of non-stabilizer states, also known as magic states, which are the resourceful states in a resource theory of magic where SO represent the free operations.

Within the qubit stabilizer formalism, multiqubit states are eigenvectors of tensor product of Pauli operators. However, even though such kind of states can be easily prepared through SO, non-stabilizer operations might be necessary in order to perfectly discriminate them, thus exhibiting non-stabilizerness without magic [8]. Notably, the set of states used in [8] to reveal this phenomenon is related to the so called SHIFT basis, a complete orthonormal set of unentangled vectors for a three-qubit system which is known to exhibit quantum non-locality without entanglement (QNLWE) [3]. This means that the vectors composing the basis are not perfectly discriminable through LOCC protocols, despite being product vectors. From a resource theoretic perspective, non-stabilizerness without magic is the analogous of QNLWE in the resource theory of entanglement. In the latter case, there exists a set of free states (product states) that cannot be perfectly discriminated via free (LOCC) operations. Correspondigly, in a resource theory of magic where free operations are stabilizer operations, there exists stabilizer product states that cannot be discriminate via free operations.

Remarkably, such states can be perfectly discriminated by a protocol in which the classical communication takes place without a definite causal order [7], thus revealing a connection between causality and non-locality. All these results highlight a pattern which seems to link together topics that are, in principle, unrelated: causality, non-locality and non-stabilizerness. The aim of the present work is to strengthen the relationship between causality and non-stabilizerness, as described below.

Methods and Results

In this section, we establish a fundamental link between classical indefinite causal order and the resource theory of magic by showing that multiqubit product stabilizer bases generated by non-causal process functions cannot be perfectly discriminated using stabilizer operations alone, despite containing no magic states. Since every non-causal process function violates a causal inequality, causal inequalities thus provide an operational witness of non-stabilizerness. We further extend this connection to the multiqudit case by introducing a local d-colourability condition, which ensures that a broad class of (possibly ambiguous) product bases can still be represented through quasi-process functions.

The process function framework

A well-established framework to rigourously study the consequences of dropping the assumption of a definite causal order in a classical communication scenario is the process functions framework [2]. Consider n parties, each one is in his/her own laboratory, who are allowed to perform one round of classical communication exchanging classical systems with the environment. Each party receives an input $i_k \in \mathcal{I}_k$ from the environment only once and similarly they feed one output $o_k \in \mathcal{O}_k$ in it. Moreover, each of them has access to an input x_k, the local setting, and produces an outcome a_k. The basic assumption is that classical probability theory holds locally, meaning that the local operations they can apply to their input variables are stochastic processes $p(o_k, a_k|i_k x_k)$. The environment essentially encodes how the parties communicate with each other and, in the deterministic case, it is mathematically modeled by a function $\omega : \times_k \mathcal{O}_k \to \times_k \mathcal{I}_k$. Combining this function with the local intervention of the parties yields the observed correlations $P(\boldsymbol{a}|\boldsymbol{x}) = \sum_{\boldsymbol{o},\boldsymbol{i}} \delta_{\omega(\boldsymbol{o}),\boldsymbol{i}} \prod_k p(o_k, a_k|i_k x_k)$. A process function is a function ω that gives a legitimate probability distribution $P(\boldsymbol{a}|\boldsymbol{x})$, for any arbitrary choice of local interventions; otherwise the function is called a quasi-process function. This constraint also implies that $\omega_k(\boldsymbol{a}) \equiv \omega_k(\boldsymbol{a}_{\setminus k})$, where $\boldsymbol{a}_{\setminus k} = (a_1, \ldots, a_{k-1}, a_{k+1}, \ldots, a_n)$ (non-self signaling condition). The process function is the central object in this framework: it completely determines whether the communication between the parties takes place with a definite causal order. According to the characterization provided in [4] a process function is non-causal if there exists at least a subset of parties such that the input received by each party depends non trivially from the outputs of the other ones. In other words, for such a subset, there is no party in the global past of all the other ones. Notice that this condition still allows for a partial causal order among some of the parties.

Causal nequalities as witnesses of non-stabilizerness (Multiqubit Case)

In the multiqubit case, there is a tight connection between product bases and boolean process functions $\omega : \{0,1\}^n \to \{0,1\}^n$. Indeed, it has been shown that every product basis can be mapped to a unique process function via the following formula [4]

$$\mathcal{S} = \left\{ \bigotimes_k U_k^{\omega_k(\boldsymbol{a}_{\setminus k})} |\boldsymbol{a}\rangle \, | \boldsymbol{a} \in \{0,1\}^n \right\} \tag{1}$$

where $U_k^{\omega(\boldsymbol{a}_{\setminus k})}$ are local controlled unitaries. Depending on the values $\boldsymbol{a}_{\setminus k}$ of the other qubits, either U_k^0 or U_k^1 is applied, and when $U_k^{x_k} \in \{\mathbb{1}, H, SH\}$, the set $\mathcal{S}$ corresponds to a product basis made of stabilizer states (notice that SH rotates the computational basis in the Y eigenbasis). Now, in the resource theory of magic where free operations are SO, a discrimination protocol is made of the following building blocks: i) preparing ancillary quibits in some stabilizer state; ii) applying Clifford unitaries; iii) performing Pauli measurements. The operations can be applied adaptively, depending on the outcomes of previous measurements. We then say that a set of product stabilizer states $\mathcal{S}$ exhibits non-stabilizerness without magic if it cannot be discriminated through stabilizer operations.

A simple example of a basis that can be perfectly discriminated via stabilizer operations is the two-qubit basis $\{|00\rangle, |01\rangle, |1+\rangle, |1-\rangle\}$. A trivial LOCC protocol consists of Alice measuring her qubits and sending the result to Bob: if she obtains 0, then Bob measures in the Z basis, otherwise he measures in the X basis. Notice that, even though the measurement protocol is a legitimate stabilizer protocol, the unitary implementing the measurement in a circuital representation is a controlled Hadamard, which is a non-Clifford unitary. As already mentioned, the SHIFT basis cannot be discriminated using stabilizer operations, and it can be expressed via the formula (1) where ω is the Lugano process [2]. Therefore, the result shown for the SHIFT basis might suggest that the non-causality of the process function in (1) plays a key role in determining the discriminability of the associated stabilizer basis. It turns out that this is indeed the case, as proved by the following theorem

Theorem 1. *Consider a product basis* $\mathcal{S} = \left\{ \bigotimes_k U_k^{\omega_k(\boldsymbol{a}_{\setminus k})} |\boldsymbol{a}\rangle \, | \boldsymbol{a} \in \{0,1\}^n \right\}$. *If the process function* ω *is non-causal,* $\mathcal{S}$ *exhibits non-stabilizerness without magic.*

It is known that every non-causal process function violates the "guess the process function" causal inequality [4]. As a corollary of the previous theorem we then conclude that causal inequalities provide an operational witness for non-stabilizerness.

Extension to the multiqudit case: local colourability

Let $\mathcal{S} = \{\left|\psi^i\right\rangle = \otimes_{k=1}^n \left|\psi_k^i\right\rangle\}_{i=1}^{|\mathcal{S}|}$ be an orthonormal product basis. For each party k, we can group the local vectors into a set $\mathcal{S}^{(k)} = \{\left|\psi_k^i\right\rangle\}_{i=1}^{s_k}$ of $s_k \leq |\mathcal{S}|$ distinct elements. This set can be further divided into $|\mathcal{X}_k|$ subsets $\mathcal{S}_{x_k}^{(k)}$ such that $\mathcal{S}^{(k)} = \cup_{x_k=0}^{|\mathcal{X}_k|-1} \mathcal{S}_{x_k}^{(k)}$ and all vectors in a given subset are mutually orthogonal. Since $\mathcal{S}$ is a basis, we have that each subset $\mathcal{S}_{x_k}^{(k)}$ contains exactly d_k (local dimension) elements and is itself a local basis.

When the product basis in unambiguous, meaning that for every party k $\langle\psi_k^i|\psi_k^j\rangle = 0$ iff they belong to the same setting, there always exists a process function and a set of local unitaries such that $\mathcal{S}$ can be expressed as

$$\mathcal{S} = \{\otimes_k U_k^{\omega_k(\boldsymbol{a})} |\boldsymbol{a}\rangle \, | \boldsymbol{a} \in \{0,1\}^n\} \tag{2}$$

This property is a key ingredient used in the construction given in [4]. Here we observe that unambiguity can be relaxed in a way that still allows to write $\mathcal{S}$ as above. For each local vector $|\psi_k^i\rangle$ we choose a setting $S_{x_k}^{(k)}$. When $|\psi_k^i\rangle$ appears in multiple settings we can choose to assign the same setting to each occurence, even if this is not necessary. In any case, in this way, each vector of the product basis $\left|\psi^i\right\rangle \in \mathcal{S}$ is associated with a unique string of settings $(x_1, \ldots, x_n)$. Now for every $\mathcal{S}_{x_k}^{(k)} = \{|\psi_k^{i_{x_k}}\rangle\}_{x_k \in \mathcal{X}_k}$ there exists a unitary operator $U_k^{x_k}$ that maps the computational basis $\mathcal{B}_k$ to $\mathcal{S}_{x_k}^{(k)}$. We consider the following assumption

Assumption 1 (Local colourability)**.** *For every party k there exists a labeling function $f_k : \mathcal{S}^{(k)} \to \{0, \ldots d_k - 1\}$ such that whenever $|\phi\rangle$, $|\psi\rangle$ are orthogonal $f_k(|\phi\rangle) \neq f_k(|\psi\rangle)$.*

This assumption is equivalent to requiring that the orthogonality graph of the local set $S^{(k)}$ is d_k-colourable for every k. On the one hand, this condition is implied by the aforementioned unambiguity condition. On the other hand this represents a weaker constraint since it applies also to ambiguous bases, such as the Domino basis [3]. However, there might be some cases in which this assumption is too strong, since one can build local sets that are not d-colourable. A counterexample for $d = 3$ can be easily built using the thirteen states proposed in [11]. Effectively, the constraint of local d-colourability amounts to a noncontextual assignment of "colours" to the local sets. In the case of $d = 2$, such a colouring always exists: this fact underlies the Kochen-Specker hidden-variable model for a qubit [6, 10] and it explains why all unentangled multiqubit bases are "unambiguous" [4], defining process functions.

An immediate consequence of the local colourability of the local sets is that the unitaries $U_k^{x_k}$ can be explicitly written as

$$U_k^{x_k} = \sum_{i_{x_k}} |\psi_k^{i_{x_k}}\rangle\langle f(\psi_k^{i_{x_k}})| \tag{3}$$

namely, whatever setting x_k a given vector $|\psi_k^i\rangle \in \mathcal{S}^{(k)}$ has been assigned, we have $|\psi_k^i\rangle = U_k^{x_k}|f(\psi_k^i)\rangle$. This readily implies the existence of a function that leads to (2), as stated in the following proposition

Proposition 1. *For every locally colourable product basis $\mathcal{S}$ there always exists a quasi-process function*

$$\boldsymbol{\omega} : \times_{k=1}^n \{0, \ldots, d_k - 1\} \to \boldsymbol{\mathcal{X}} \tag{4}$$

such that

$$\mathcal{S} = \{\bigotimes_{k=1}^n U_k^{\omega_k(\boldsymbol{a})} |\boldsymbol{a}\rangle \, | \boldsymbol{a} \in \times_{k=1}^n \{0, \ldots, d_k - 1\}\} \tag{5}$$

This proposition provides a recipe to build a quasi-process function that can be used to derive an explicit representation of the basis $\mathcal{S}$.

References

[1] Scott Aaronson and Daniel Gottesman. Improved simulation of stabilizer circuits. *Phys. Rev. A*, 70:052328, Nov 2004.

[2] Ämin Baumeler and Stefan Wolf. The space of logically consistent classical processes without causal order. *New Journal of Physics*, 18(1):013036, jan 2016.

[3] Charles H. Bennett, David P. DiVincenzo, Christopher A. Fuchs, Tal Mor, Eric Rains, Peter W. Shor, John A. Smolin, and William K. Wootters. Quantum nonlocality without entanglement. *Phys. Rev. A*, 59:1070–1091, Feb 1999.

[4] Hippolyte Dourdent, Kyrylo Simonov, Andreas Leitherer, Emanuel-Cristian Boghiu, Ravi Kunjwal, Saronath Halder, Remigiusz Augusiak, and Antonio Acín. Paradox-free classical non-causality and unambiguous non-locality without entanglement are equivalent, 2025.

[5] Daniel Gottesman. The heisenberg representation of quantum computers, 1998.

[6] Simon Kochen and E. P. Specker. The problem of hidden variables in quantum mechanics. *Journal of Mathematics and Mechanics*, 17(1):59–87, 1967.

[7] Ravi Kunjwal and Ämin Baumeler. Trading causal order for locality. *Phys. Rev. Lett.*, 131:120201, Sep 2023.

[8] Hyukjoon Kwon. Nonstabilizerness without magic: Classically simulatable quantum states that are indistinguishable by classically simulatable quantum circuits, 2025.

[9] Maarten Nest. Classical simulation of quantum computation, the gottesmann-knill theorem, and slightly beyond. *Quantum Information & Computation*, 10:258–271, 03 2010.

[10] Victoria J Wright and Ravi Kunjwal. Contextuality in composite systems: the role of entanglement in the Kochen-Specker theorem. *Quantum*, 7:900, January 2023.

[11] Sixia Yu and C. H. Oh. State-independent proof of kochen-specker theorem with 13 rays. *Phys. Rev. Lett.*, 108:030402, Jan 2012.

Causal inequalities witness non-stabilizerness

Leonardo Vaglini, Nasra Daher Ahmed, and Ravi Kunjwal
Aix-Marseille University, CNRS, LIS, Marseille, France

(Dated: February 2026)

We study non-stabilizerness without magic, where multiqubit stabilizer product states cannot be perfectly discriminated using only stabilizer operations. Drawing a parallel with quantum non-locality without entanglement, we show that this limitation arises from classical processes with indefinite causal order. Using the process functions framework, we demonstrate that product stabilizer bases generated by non-causal process functions are operationally indistinguishable by stabilizer protocols. Since non-causal process functions violate causal inequalities, these inequalities provide an operational witness of non-stabilizerness, establishing a direct link between causal structures and resource-theoretic quantum advantage.

I. INTRODUCTION

The Gottesman-Knill theorem [1, 11, 12, 20] sets a boundary in quantum computation theory that must be overcome in order to achieve a quantum speedup over classical computers. It establishes that any quantum computation performed using stabilizer operations (SO), i.e. operations built out of Clifford operations, stabilizer states preparation and Pauli measurements performed adaptively, can be efficiently classically simulated. Therefore, observing a quantum computational advantage requires the use of non-Clifford gates, such as the Toffoli gate, or the preparation of non-stabilizer states, also known as magic states, which are the resourceful states in a resource theory of magic [9, 23] where SO represent the free operations.

Within the qubit stabilizer formalism, multiqubit states are eigenvectors of tensor product of Pauli operators. However, even though such kind of states can be easily prepared through SO, non-stabilizer operations might be necessary in order to perfectly discriminate them, thus exhibiting non-stabilizerness without magic [18]. This result somewhat mimics the separation between the set of separable operations and LOCC transformations, which is revealed by the existence of QNLWE product bases. Such bases are made of vectors that are not perfectly discriminable through Local Operations and Classical Communication (LOCC) protocols, despite being product vectors. Then, from a resource theoretic perspective, non-stabilizerness without magic is the analogous of Quantum Non-Locality Without Entanglement (QNLWE) in the resource theory of entanglement. In the latter case, there exists a set of free states (product states) that cannot be perfectly discriminated via free (LOCC) operations. Correspondigly, in an operational resource theory of magic where free operations are stabilizer operations, there exists stabilizer product states that cannot be discriminated via free operations. Notably, the set of states used in [18] to reveal this phenomenon is related to the so called SHIFT basis, a complete orthonormal set of unentangled vectors for a three-qubit system which is known to exhibit QNLWE [6].

Recently, the QNLWE property of product bases has also been studied in connection with the process function framework [5, 10, 21], where the assumption of a definite causal order is abandoned. The process function framework corresponds to deterministic classical counterpart of the process martrix formalism developed within the quantum setting [21]. One models multipartite classical operations without assuming a fixed causal order, while ensuring that the overall dynamics remain logically consistent, i.e., producing legitimate probability distributions for the correlations among the parties for every choice of their local operations. It has been shown that non-causal quantum processes provide operational advantages in a variety of information-theoretic tasks [2, 7, 8, 14, 19]. Along these lines, it has been proved that the SHIFT basis can be perfectly discriminated by a protocol in which the classical communication takes place without a definite causal order [17]. This result has been strengthened in [10], where the authors prove that every non-causal process function can be encoded on a QNLWE basis, and viceversa every QNLWE basis is mapped to a non-causal process function, thus revealing a deep connection between causality and non-locality. All these results highlight a pattern which seems to link together topics that are, in principle, unrelated: causality, non-locality and non-stabilizerness. The aim of the present work is to strengthen the relationship between causality and non-stabilizerness.

The remainder of the paper is organized as follows. In Section II, we review the process-function framework and recall the characterization of causal and non-causal process functions, together with their reduction properties. We then discuss the correspondence between complete product bases and (quasi-)process functions, introducing the notions of unambiguity and local colourability, and proving that locally colourable product bases can always be encoded in terms of suitable local unitaries and a quasi-process function. In Section II we further clarify the relation between this construction and quantum non-locality without entanglement. In Section III, after reviewing the stabilizer formalism and adaptive stabilizer discrimination protocols, we prove our main result: every multiqubit stabilizer product basis associated with a non-

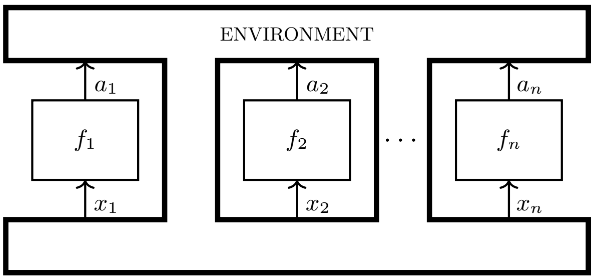

FIG. 1: A pictorial representation of the scenario described in the main text. Each party is allowed to locally apply classical channels, mathematically represented by functions $f_k : \mathcal{X}_k \to \mathcal{A}_k$ in the deterministic case. The way in which the parties communicate with each other is ruled by the environment, mathematically described by a process function $\boldsymbol{\omega} : \boldsymbol{\mathcal{A}} \to \boldsymbol{\mathcal{X}}$

causal process function exhibits non-stabilizerness without magic. In section IV We draw our conclusions.

II. PRELIMINARIES

In this section, we give a brief introduction of the process function framework and collect the results relevant for the present work.

A. Process function framework

The process functions framework [5] provides a well-established approach to rigorously study the consequences of dropping the assumption of a definite causal order in a classical communication scenario. The correlational scenario is defined as follows: consider n parties, each one is in his/her own laboratory, who are allowed to perform one round of classical communication exchanging classical systems with the environment. Each party receives an input $x_k \in \mathcal{X}_k$ from the environment only once and similarly feeds one output $a_k \in \mathcal{A}_k$ into it. Moreover, each of them has access to an input $i_k \in \mathcal{I}_k$, the local setting, and produces an outcome $o_k \in \mathcal{O}_k$. The basic assumption is that classical probability theory holds locally, which means that the local operations they can apply to their input variables are stochastic processes $p(o_k, a_k|i_k x_k)$.

The environment essentially encodes how the parties communicate with each other and, in the deterministic case, it is mathematically modeled by a function $\omega : \times_k \mathcal{O}_k \to \times_k \mathcal{I}_k$. Combining this function with the local intervention of the parties yields the observed correlations

$$P(\boldsymbol{o}|\boldsymbol{i}) = \sum_{\boldsymbol{a},\boldsymbol{x}} \delta_{\omega(\boldsymbol{a}),\boldsymbol{x}} \prod_k p(o_k, a_k|i_k x_k). \tag{1}$$

In figure 1 we reproduce a pictorial representation of the correlational scenario. A process function is a function ω that gives a legitimate probability distribution $P(\boldsymbol{a}|\boldsymbol{x})$, for any arbitrary choice of local interventions; otherwise the function is called a quasi-process function. This constraint also implies that $\omega_k(\boldsymbol{a}) \equiv \omega_k(\boldsymbol{a}_{\setminus k})$, where $\boldsymbol{a}_{\setminus k} = (a_1, \ldots, a_{k-1}, a_{k+1}, \ldots, a_n)$, known as the non-self signaling condition. In the remainder, we will denote the cartesian product of all the outcomes set $\mathcal{A}_k$ by $\mathcal{A} = \bigtimes_k \mathcal{A}_k$, and similarly for the settings $\mathcal{X} = \bigtimes_k \mathcal{X}_k$.

The process function is the central object in this framework: it completely determines whether the communication between the parties takes place with a definite causal order. According to the characterization provided in [10] a process function is non-causal if there exists at least a subset of parties such that the input received by each party depends non trivially on the outputs of the other ones. In other words, for such a subset, there is no party in the global past of all the other ones. Notice that this condition still allows for partial causal order among some of the parties. To introduce the formal characterization of causal process functions, we introduce a standard procedure to obtain a $n-1$-partite process function from a n-partite one. Assume to have a process function $\omega : \bigtimes_k \mathcal{A}_k \to \bigtimes_k \mathcal{X}_k$. Now, fix a local deterministic intervention $f : \mathcal{X}_k \to \mathcal{A}_k$ for a given party k and define $\omega^{f^k} : \mathcal{A}_{\setminus k} \to \mathcal{X}_{\setminus k}$ as follows

$$\boldsymbol{\omega}^{f_k}(\boldsymbol{a}_{\setminus k}) := \boldsymbol{\omega}_{\setminus k}(\boldsymbol{a}_{\setminus k}, f_k(\omega_k(a_{\setminus k}))). \tag{2}$$

A crucial point here is that a process function is non-self signaling, which guarantees the well-posedness of the above definition. Indeed, the new function does not depend on the outcomes of the other parties, and it is a process function called reduced process function. A special case of this is when the local intervention is a constant function, i.e. $f(x_k) = a_k \forall x_k$. In this case $\boldsymbol{\omega}^{a_k}$ is called output-reduced process function. This latter notion has been used in a recent characterization of causal process functions. We quote it here for the reader's convenience.

Theorem 1 (Theorem 4 [10]). *Let $\boldsymbol{\omega} : \mathcal{A} \to \mathcal{X}$ be a n-partite process function. Then, $\boldsymbol{\omega}$ is causal if and only if the following conditions are satisfied:*

1. *there exists a party k_1 such that no other party can signal to it, i.e., the input received by the environment is constant*

$$x_{k_1} = \omega_{k_1}(a_{\setminus k_1}) = \overline{x}_{k_1}. \tag{3}$$

2. *For every such party and for every $a_{k_1} \in \mathcal{A}_{k_1}$, the output-reduced process functions $\boldsymbol{\omega}^{a_{k_1}]}$ are themselves causal $n-1$-partite process functions.*

By contrapositive, non-causal process functions are those that violate one of the two conditions above. If condition 1 is violated, there is no party who receives a constant from the environment. The input of every party depends on the output of at least another party. Formally speaking, for every k there exist $\boldsymbol{a}^0_{\backslash k}$, $\boldsymbol{a}^1_{\backslash k}$ such that

$$\omega_k(\boldsymbol{a}^0_{\backslash k}) \neq \omega_k(\boldsymbol{a}^1_{\backslash k}). \tag{4}$$

Notice $\boldsymbol{a}^0_{\backslash k}$, $\boldsymbol{a}^1_{\backslash k}$ might differ for just one component. In this case, there is no global: every party is in the causal past of another party. The most famous and simplest example of a non-causal process function of this kind is the AF/BW process, also known as the Lugano process, defined by the following set of equations

$$x_1 = a_3(a_2 \oplus 1), \quad x_2 = a_1(a_3 \oplus 1), \quad x_1 = a_2(a_1 \oplus 1). \tag{5}$$

The input of each party depends on the two outputs of the other parties. However, there can be process functions that meet condition 1 but fail to satisfy condition 2. Even though such functions allow for causal communication among a subset $\mathcal{K}$ of the parties, there is a $n - |\mathcal{K}|$-partite process function that violates condition 1. In other words, there is a configuration of the parties in $\mathcal{K}$ such that the remaining parties communicate without a definite causal order. A simple example of this case is the following 4-partite process. A party in the global past, whose output establishes whether the other parties communicate in a fixed causal order or via the AF/BW process:

$$\begin{aligned} x_1 =& (1 - a_4) + a_4 a_3 (a_2 \oplus 1), \\ x_2 =& (1 - a_4) + a_4 a_1 (a_3 \oplus 1), \\ x_3 =& (1 - a_4) + a_4 a_2 (a_1 \oplus 1), \\ x_4 =& 0. \end{aligned} \tag{6}$$

This obviously satisfies condition 1, thanks to party 4. At the same time, there is a configuration of this party, $a_4 = 1$, such that the output reduced process function $\boldsymbol{\omega}^{a_4=1}(a_1, a_2, a_3)$ is non-causal.

The AF/BW process is the prototypical example of a non-causal process function. It is known to lead to the maximal violation the following causal inequality

$$\begin{aligned} \frac{1}{8}\, [&P(000|000) + P(100|001) + P(001|010) \\ &+ P(001|011) + P(010|100) + P(100|101) \\ &+ P(010|110) + P(000|111)] \leq \frac{3}{4}. \end{aligned} \tag{7}$$

Notice that each probability term $P(\boldsymbol{o}|\boldsymbol{i})$ appearing in the above equation is in correspondence with an event $\boldsymbol{x} = \boldsymbol{\omega}(\boldsymbol{a})$ of the AF/BW process. This observation let us rewrite the above inequality in the following form

$$\frac{1}{8} \sum_{\boldsymbol{i}} P(\boldsymbol{o} = \boldsymbol{\omega}(\boldsymbol{i})|\boldsymbol{i}) \leq \frac{3}{4} \tag{8}$$

This formulation of the inequality let us reinterpret it as a causal game dubbed "guess the process function" game. When the classical communication is governed by the AF/BW process, the parties can win the game with probability one using the following local strategy: each party feeds its external input i_k into the environment, $a_k = i_k$, and gives as output the value received from it, $o_k = x_k$. This amounts to choose the local interventions $P(o_k, a_k|i_k, x_k) = \delta_{o_k, x_k}\delta_{a_k, i_k}$.

This can be generalized to any non-causal process function $\boldsymbol{\omega}^{NC}$. When the parties implement the local strategy described above, the obtained correlations maximally violate the following causal inequality

$$\frac{1}{\prod_k |\mathcal{I}_k|} \sum_{\boldsymbol{i}} P(\boldsymbol{o} = \boldsymbol{\omega}^{NC}(\boldsymbol{i})|\boldsymbol{i}) \leq \beta^C \tag{9}$$

where β_c denotes the bound achievable with correlations obtained via causal communication. When $\forall k$, $|\mathcal{I}_k| = d$, $\beta^C = 1 - \frac{1}{d^n}$ [5]. In the general case, the bound is given by [10]

$$\beta^C = 1 - \frac{1}{\prod_k |\mathcal{I}_k|} \tag{10}$$

B. Relation with complete product bases

In [10], the authors establish a strong relationship between a certain class of orthonormal product bases, called unambiguous, and process functions. We briefly review their construction, and we connect it to d-colourability, a property that the set of local bases must satisfy. Given a qudit system, or more generally, the tensor product of n qudit systems $\mathcal{H}_1 \otimes \cdots \otimes \mathcal{H}_n$, each of dimension d_k, a complete product basis is a set of vectors $\mathcal{S}$ of the form

$$\mathcal{S} = \{|\psi^i\rangle = \bigotimes_{k=1}^{n} |\psi^i_k\rangle\}_{i=1}^{|S|} \tag{11}$$

where $|S| = \prod_k d_k$. For each party k, we can group the local vectors into a set $\mathcal{S}^{(k)} = \{|\psi^i_k\rangle\}_{i=1}^{s_k}$ of $s_k \leq |\mathcal{S}|$ distinct elements (so $S^{(k)}$ does not contain repetitions of the same element). This set can be further divided into $|\mathcal{X}_k|$ subsets $\mathcal{S}^{(k)}_{x_k}$ such that

$$\mathcal{S}^{(k)} = \bigcup_{x_k=0}^{|\mathcal{X}_k|-1} \mathcal{S}^{(k)}_{x_k}$$

and all vectors in a given subset are mutually orthogonal. Since $\mathcal{S}$ is a basis, we have that each subset $\mathcal{S}^{(k)}_{x_k}$ contains exactly d_k (local dimension) elements and is itself a local basis. At this point, for any k we can interpret the $\mathcal{X}_k$ as the set of local settings, and for each setting $\mathcal{S}^{(k)}_{x_k}$, a local vector in $\mathcal{S}^{(k)}_{x_k}$ is associated with an outcome a_k corresponding to the position of the vector in the basis. For each party, we denote by $\mathcal{B}_k$ the computational basis.

We can consider two different paradigmatic examples. The first is three-qubit SHIFT basis

$$\begin{aligned}\mathcal{S} = \{\, &|000\rangle\,, |{+}01\rangle\,, |01{+}\rangle\,, |01{-}\rangle\,,\\ &|1+0\rangle\,, |{-}01\rangle\,, |1-0\rangle\,, |111\rangle\}\end{aligned} \tag{12}$$

For each party $k \in \{0,1,2\}$, the local sets are given by

$$\mathcal{S}^{(k)} = \{|0\rangle\,, |1\rangle\,, |+\rangle\,, |-\rangle\}, \tag{13}$$

and the local settings are

$$\mathcal{S}_0^{(k)} = \{|0\rangle\,, |1\rangle\}, \quad \mathcal{S}_1^{(k)} = \{|+\rangle\,, |-\rangle\} \tag{14}$$

In this case, every vector $\left|\psi^i\right\rangle \in \mathcal{S}$ can be assigned to a unique setting, and it is orthogonal to the vectors that belong to the same setting only. Thus, the map that assigns each outcome string $\boldsymbol{a} = (a_1, a_2, a_3)$ with the corresponding setting $\boldsymbol{x} = (x_1, x_2, x_3)$ clearly is a function, i.e., we do not have that one string of outcomes associated with two different settings. The property for which every local vector belongs to one and only one setting, satisfied by the SHIFT basis, is called *unambiguity* [3, 4, 10], and it is defined as follows

Definition 1. *A complete joint product basis $\mathcal{S}$ is said to be unambiguous if, for every party k, two local vectors are orthogonal if and only if they belong to the same setting, i.e.,*

$$\langle\psi_k^{j'}|\psi_k^{j}\rangle = 0 \iff x_k^{j'} = x_k^{j}. \tag{15}$$

This property, together with the fact that $\mathcal{S}$ is a complete orthonormal basis, ensures that the mapping $\boldsymbol{\omega} : \boldsymbol{\mathcal{A}} \to \mathcal{X}$ is a well defined quasi-process functions. Actually, it has been shown that unambiguity [10] not only implies that $\boldsymbol{\omega}$ is a process function, but that is also unique. Viceversa, every process function can be encoded onto an unambiguous product basis by suitably choosing the local unitaries $U_k^{x_k}$. In summary, for every unambiguous product basis $\mathcal{S}$ there exists a unique process function $\boldsymbol{\omega}$ and a set of local unitaries $\{U_k^{x_k}\}_{k=1}^n$ such that

$$\mathcal{S} = \{\bigotimes_k U_k^{\omega_k(\boldsymbol{a}_{\setminus k})}\,|\boldsymbol{a}\rangle\,|\boldsymbol{a} \in \boldsymbol{\mathcal{A}}\}. \tag{16}$$

A striking consequence of this relation between orthonormal product bases and process functions is a tight relation between the QNLWE property and causality: every unambiguous product basis that is QNLWE is associated with a unique non-causal process function. Viceversa, every non-causal process function can be encoded on an unambiguous QNLWE basis.

While multiqubit product bases are all unambiguous, there are examples of ambiguous bases. The simplest example is provided by the Domino basis [REF], which is the following set of nine states forming a basis of the the two-qutrit system:

$$\begin{aligned}\mathcal{S}_{\mathrm{D}} =\{&|0+\rangle\,, |0-\rangle\,, |2+'\rangle\,, |2-'\rangle\,,\\ &|+'0\rangle\,, |-'0\rangle\,, |+2\rangle\,, |-2\rangle\,, |11\rangle\},\end{aligned} \tag{17}$$

where $|\pm\rangle = \frac{1}{\sqrt{2}}(|0\rangle \pm |1\rangle)$ and $|\pm'\rangle = \frac{1}{\sqrt{2}}(|1\rangle \pm |2\rangle)$ For both parties, the local sets $\mathcal{S}^{(k)}$ are given by

$$\mathcal{S}^{(k)} = \{|0\rangle\,, |1\rangle\,, |2\rangle\,, |+\rangle\,, |-\rangle\,, |+'\rangle, |-'\rangle\} \tag{18}$$

and from this set the following bases can be drawn:

$$\begin{aligned}\mathcal{S}_0^{(k)} &= \{|0\rangle\,, |1\rangle\,, |2\rangle\}\\ \mathcal{S}_1^{(k)} &= \{|0\rangle\,, |+'\rangle\,, |-'\rangle\}\\ \mathcal{S}_2^{(k)} &= \{|+\rangle\,, |-\rangle\,, |2\rangle\}.\end{aligned}$$

For the Domino basis, the vectors $|0\rangle$ and $|2\rangle$ appear in two different settings for both parties, making this basis ambiguous. This does not forbid the possibility of writing $\mathcal{S}_{\mathrm{D}}$ in terms of a quasi-process function, as explicitly shown in [10].

A simple procedure to write down explicitly the quasi-process function can be devised. For each vector, we list all possible settings in the following table

$\left\vert\psi^i\right\rangle$	$(\boldsymbol{a}^i\vert\boldsymbol{x}^i)$
$\vert 0+\rangle$	(02) or (12)
$\vert 0-\rangle$	(02) or (12)
$\vert 2+'\rangle$	(01) or (21)
$\vert 2-'\rangle$	(01) or (21)
$\vert +'0\rangle$	(10) or (11)
$\vert -'0\rangle$	(10) or (11)
$\vert +2\rangle$	(20) or (22)
$\vert -2\rangle$	(20) or (22)
$\vert 11\rangle$	(00)

Therefore, in the first place, for each party we have to assign a setting to each vector. There are two vectors, $|0\rangle$ and $|2\rangle$, which can be assigned to two different settings. We can fix the settings as we wish, so let us suppose that we assign both to the setting $x_k = 0$. Now, notice that for each party, the local set $S^{(k)}$ is such that there exists a function $f_k : S^{(k)} \to \{0,1,2\}$ with the property that for every pair of orthogonal states $|\psi\rangle\,, |\phi\rangle \in S^{(k)}$, $f_k(|\psi\rangle) \neq f_k(|\phi\rangle)$. We can choose the labeling functions f_k in the following way:

$$\begin{aligned}&f_k(|0\rangle) = f_k(|+\rangle) = 0,\\ &f_k(|1\rangle) = f_k(|-\rangle) = f_k(|+'\rangle) = 1,\\ &f_k(|2\rangle) = f_k(|-'\rangle) = 2.\end{aligned}$$

These functions allow us to fix a convention for the outcomes (a_1^i, a_2^i) associated with each vector $\left|\psi^i\right\rangle = \left|\psi_1^i\right\rangle \otimes \left|\psi_2^i\right\rangle$ in $\mathcal{S}_D$ set

$$(a_1, a_2) := (f_1(\left|\psi_1^i\right\rangle), f_2(\left|\psi_2^i\right\rangle)) \tag{19}$$

Given the property of f it is immediate to see that for every pair (a_1, a_2) there exists a unique vector in $\mathcal{S}_D$ such that the above equation holds. Therefore, we have a map $\boldsymbol{\omega} : \boldsymbol{\mathcal{A}} \to \boldsymbol{\mathcal{X}}$ where each outcome (a_1, a_2) associated with a distinct vector is mapped to the unique setting representing that vector. Notice that, depending on the initial choice of the settings, different $\boldsymbol{\omega}$ may be associated with $\mathcal{S}_D$.

Finally, using the following unitaries defined using the functions f_k

$$\begin{aligned}
U_k^0 &= |0\rangle \langle f(0)| + |1\rangle \langle f(1)| + |2\rangle \langle f(2)| = \\
&= |0\rangle \langle 0| + |1\rangle \langle 1| + |2\rangle \langle 2| = I \\
U_k^1 &= |0\rangle \langle f(0)| + |+'\rangle\langle f(+')| + |-'\rangle\langle f(-')| \\
&= |0\rangle \langle 0| + |+'\rangle \langle 1| + |-'\rangle \langle 2| \\
U_k^2 &= |+\rangle \langle f(+)| + |-\rangle \langle f(-)| + |2\rangle \langle f(2)| \\
&= |+\rangle \langle 0| + |-\rangle \langle 1| + |2\rangle \langle 2| \,.
\end{aligned} \tag{20}$$

one can easily check that the basis $\mathcal{S}_{\mathrm{DOM}}$ is obtained using the formula $\otimes_k U_k^{x_k} |\boldsymbol{a}\rangle$.

The recipe described above for the Domino basis $\mathcal{S}_D$ hinges on the existence of the function f_k. That property is equivalent to say that the orthogonality graph of each local set is 3-colourable. We can formalize this property to define a class of product basis

Definition 2 (Local colourability)**.** *Consider a complete product basis $\mathcal{S}$. a local set $S^{(k)}$ is said to be d_k-colourable if for every party k there exists a function*

$$f_k : \mathcal{S}^{(k)} \to \{0, \ldots d-1\} \tag{21}$$

such that whenever $|\phi\rangle, |\psi\rangle \in \mathcal{S}$ are orthogonal $f_k(|\phi\rangle) \neq f_k(|\psi\rangle)$. Then $\mathcal{S}$ is said to be locally colourable if for every k $\mathcal{S}^{(k)}$ is d_k-colourable.

As already stated, this definition can be phrased in graph-theoretic terms: $\mathcal{S}$ is locally colourable if the orthogonality graph of each local set $S^{(k)}$ is d-colourable. On the one hand, this condition is implied by the aforementioned unambiguity condition. Indeed, the orthogonality graph of unambiguous local setting correspond to the disjoint union of $|\mathcal{X}_k|$ d_k-colourable graphs. On the other hand this represents a weaker constraint since it applies also to ambiguous bases, such as the Domino basis [6]. However, there might be some cases in which this assumption is violated, since one can build local sets that are not d-colourable. A counterexample for $d = 3$ can be easily built using the thirteen states proposed in [25]. Effectively, the constraint of local d-colourability amounts to a noncontextual assignment of "colours" to the local sets. In the case of $d = 2$, such a colouring always exists: this fact underlies the Kochen-Specker hidden-variable model for a qubit [16, 24] and it explains why all unentangled multiqubit bases are "unambiguous" [10], defining process functions.

The recipe described above for in the special case of the Domino basis can be applied to any locally d-colourable product basis $\mathcal{S}$.

Proposition 1. *For every locally colourable product basis $\mathcal{S}$ there always exists a quasi-process function*

$$\boldsymbol{\omega} : \bigtimes_{k=1}^{n} \{0, \ldots, d_k - 1\} \to \boldsymbol{\mathcal{X}} \tag{22}$$

such that

$$\mathcal{S} = \{\bigotimes_{k=1}^{n} U_k^{\omega_k(\boldsymbol{a})} |\boldsymbol{a}\rangle \,|\boldsymbol{a} \in \bigtimes_{k=1}^{n} \{0, \ldots, d_k - 1\}\} \tag{23}$$

Proof. We use the notation introduced above. For each party k, $\mathcal{S}^{(k)}$ denotes the set of local vectors, and $\mathcal{S}_{x_k}^{(k)}$ denotes the local settings. If $\mathcal{S}$ is unambiguous, we already know that such an $\boldsymbol{\omega}$ exists [10], is unique and is a process function. So let us consider the ambiguous case. For every $\left|\psi^i\right\rangle$ there are multiple settings $\boldsymbol{x}^i = (x_1, \ldots, x_n)$ which might be associated with $\left|\psi^i\right\rangle$. For each vector, pick one particular setting $\boldsymbol{x}^i$ and call this function $h : \mathcal{S} \to \bigtimes_k \mathcal{X}_k$. Now, by local colourability, for every k there exists a family of functions

$$f_k : \mathcal{S}^{(k)} \to \{0, \ldots d-1\} \tag{24}$$

such that $f_k(\left|\psi_k^i\right\rangle) \neq f_k(|\psi_k^{i'}\rangle)$ whenever $\langle \psi_k^i | \psi_k^{i'} \rangle = 0$. For each vector $\left|\psi^i\right\rangle \in \mathcal{S}$, we use this family to define the outcomes via the mapping $\boldsymbol{f} : \mathcal{S} \to \bigtimes_{k=1}^{n} \{0, \ldots, d_k - 1\}$ as follows

$$\boldsymbol{a^i} = (a_1^i, \ldots, a_n^i) := (f_1(\left|\psi_1^i\right\rangle), \ldots, f_n(\left|\psi_n^i\right\rangle)) \tag{25}$$

Notice that this mapping from $\mathcal{S}$ to $\mathcal{A} \subseteq \bigtimes_{k=1}^{n} \{0, \ldots, d_k - 1\}$. It is easily seen to be injective. Let $\left|\psi^i\right\rangle \neq |\psi^{i'}\rangle$. Then $\langle \psi_k^i | \psi_k^{i'} \rangle = 0$ for some k, which implies

$$a_k^i = f_k(\left|\psi_k^i\right\rangle) \neq f_k(|\psi_k^{i'}\rangle) = a_k^{i'}.$$

Therefore $|\mathcal{A}| = \prod_{k=1}^{n} d_k$, whence the mapping is also surjective, thus invertible. The quasi-process function is then defined as $\boldsymbol{\omega} := \boldsymbol{h} \circ f^{-1}$.

A crucial point is that we can exploit the functions (24) to explicitly write down the local unitaries connecting the computational basis $\mathcal{B}_k$ to the settings $\mathcal{S}^{(k)}_{x_k}$. We have

$$U_k^{x_k} = \sum_{i_{x_k}} |\psi_k^{i_{x_k}}\rangle\langle f_k(\psi_k^{i_{x_k}})| \tag{26}$$

where i_{x_k} runs over the elements in $\mathcal{S}^{(k)}$ such that $|\psi_k^{i_{x_k}}\rangle \in \mathcal{S}^{(k)}_{x_k}$. Given $|\psi_k^i\rangle$, for every $x_k \neq x'_k$ such that $|\psi_k^i\rangle \in \mathcal{S}^{(k)}_{x_k}, \mathcal{S}^{(k)}_{x'_k}$ we always have

$$|\psi_k^i\rangle = U_k^{x_k}|f_k(|\psi_k^i\rangle)\rangle = U_k^{x'_k}|f_k(|\psi_k^i\rangle)\rangle, \tag{27}$$

namely, the vector $|\psi_k^i\rangle$ is obtained from the same vector in the computational basis, whichever unitary is used. We then have

$$\begin{aligned}\bigotimes_k U_k^{\omega_k(\boldsymbol{a})}|\boldsymbol{a}\rangle &= \bigotimes_k U_k^{\omega_k(\boldsymbol{f}(|\psi^i\rangle))}|\boldsymbol{f}(\psi^i)\rangle \\ &= \bigotimes_k U_k^{x_k}|f_k(\psi_k^i)\rangle = \bigotimes_k |\psi_k^i\rangle = |\psi^i\rangle\end{aligned} \tag{28}$$

and this concludes the proof. □

It is worth emphasizing that defining an invertible function $\boldsymbol{f} : \mathcal{S} \to \bigtimes_{k=1}^n \{0, \dots, d_k - 1\}$ is always possible, since $\mathcal{S}$ is an orthonormal basis and hence has cardinality exactly $\prod_k d_k$. Therefore, there is no obstruction to associating each vector of the basis with a particular event, i.e., defining a function from outcomes to settings. However, one might fail to find a way to write the elements of $\mathcal{S}$ in terms of suitable unitaries. This is where the local noncontextual assignments have been used in the derivation, and this is reflected in particular in equation (27). We finally remark that locally colourable QNLWE basis implies that there does not exist a causal process function. Trivially, if $\boldsymbol{\omega}$ is causal, the basis would trivially LOCC discriminable with one round of classical communication.

III. CAUSAL INEQUALITIES AS WITNESSES OF NON-STABILIZERNESS

In a recent work [18] it has been shown that the SHIFT basis, despite being made of stabilizer product states, cannot be perfectly discriminated using stabilizer operations. Such operation consists of the preparation of some ancillary qubits, Clifford operations and measurements of Pauli operators. This result has been proved using a subset of the SHIFT basis, which we know it is obtainable from a non-causal process function in the way explained in section II. In this section, we prove that in the multiqubit scenario, every product stabilizer basis whose associated process function is non-causal exhibits non-stabilizerness without magic.

A. Stabilizer formalism

We begin by introducing the Pauli group and the stabilizer formalism [13, 15, 22] in dimension d, assuming d prime throughout the rest of the work. The Pauli group is defined in terms of the shift and boost operators X and Z respectively. Their action on the computational basis is given by

$$\begin{aligned} X\,|a\rangle &= |a \oplus 1\rangle \\ Z\,|a\rangle &= \omega^a\,|a\rangle \end{aligned} \tag{29}$$

where $\omega = e^{\frac{2\pi i}{d}}$. Both Z and X are order d, namely $Z^d = X^d = I$, and they satisfy the following commutation relation

$$XZ = \omega^{-1} ZX. \tag{30}$$

The Pauli group is then generated by X and Z

$$\mathcal{P}_d = \{\omega^c X^a Z^b | a, b, c \in \mathbb{Z}_d\} \tag{31}$$

By the primality of d, every $Q \in \mathcal{P}_d$ has order d. This implies that each $I \neq Q \in \mathcal{P}_d$ has d different non-degenerate eigenvalues $\{\omega^j\}_{j\in\mathbb{Z}_d}$.

The n-qudit Pauli group $\mathcal{P}_d^n$. is generated by the tensor product of single system Pauli operators. Each $Q \in \mathcal{P}_d^n$ is of the following form

$$Q = \omega^c X^{a_1} Z^{b_1} \otimes \cdots \otimes X^{a_n} Z^{b_n}, \quad c, a_i, b_i \in \mathbb{Z}_d \tag{32}$$

Also Q has order d and eigenvalues that are the d-th root of the unity. For n qudit a stabilizer state is defined as a vector that is left invariant by a maximal abelian subgroup of $\mathcal{P}_d^n$, called the stabilizer subgroup. Every stabilizer subgroup has order n and completely determines the state. In particular, notice that every product stabilizer state is uniquely determined by local generators, i.e., Pauli operators that are non trivial only on a single factor. As an example, consider the state $|{+}01\rangle$ from the SHIFT basis. The set of generators is given by $\{X_1, Z_2, -Z_3\}$. Here we have used the shorthand notation P_k to denote the operator that acts non trivially only on the k-th factor. We will use this notational convention hereafter. In prime dimension, for a single system, the number of stabilizer eigenbases equals $d+1$ and they are mutually unbiased. Each basis is the set of common eigenvectors of a maximally abelian subgroup of $\mathcal{P}_d$. The

Clifford group is defined as the normalizer of $\mathcal{P}_d^n$,i.e., it is the set of unitary operators U such that $UPU^\dagger \in \mathcal{P}_d^n$ whenever $P \in \mathcal{P}_d^n$.

A measurement in the stabilizer formalism is given by a projective measurement in the spectral decomposition of some Pauli operator. Suppose that we want to measure a Pauli operator Q. If we do this on a state stabilized by a subgroup whose generators anticommute with Q, the outcome probability distribution is uniform, since $\langle\psi| Q |\psi\rangle$. The projector onto the ω^j-eigenspace is given by. $\Pi_j = \frac{1}{d}\sum_{t=0}^{d-1} \omega^{-jt} Q^t$, which readily implies $p(j) = \frac{1}{d}$.

An adaptive stabilizer discrimination protocol consists of: i) measurements of Pauli operators, ii) Clifford operations, iii) preparation of ℓ ancillary qubits in the state $|0\rangle$. Operations i) and ii) can be performed adaptively, depending on the outcomes of previous measurements. The superoperator associated with such a protocol is given by

$$\mathcal{D}(\cdot) = \sum_{\boldsymbol{m}} C_n \Pi_{P_k}^{m_k} \dots C_1 \Pi_{P_1}^{m_1} \cdot \Pi_{P_1}^{m_1} C_1^\dagger \dots \Pi_{P_k}^{m_k} C_k^\dagger \quad (33)$$

where each term $C_\ell \Pi_{P_\ell}^{m_\ell}$ generally depends on the previous $\ell - 1$ outcomes $m_1, \dots, m_{\ell-1}$. Since the C operations are Clifford and each Π_P^m is a Pauli measurements $U\Pi_P^m U^\dagger = \Pi_{P'}^m$ for a suitable Pauli operator P'. Therefore, all the Clifford unitaries can be moved at the end, and the protocol is essentially given by a sequence of Pauli measurements

$$\Pi_{Pk}^{m_k} \circ \dots \circ \Pi_{P1}^{m_1} \quad (34)$$

with k arbitrarily large. Thus, non-stabilizerness without magic can be defined as follows [18]

Definition 3. *A set of product stabilizer states $\mathcal{S}$ is said to exhibit non-stabilizerness without magic if it cannot be perfectly discriminated through an adaptive stabilizer discrimination protocol.*

A simple example of a basis that can be perfectly discriminated via stabilizer operations is the two-qubit basis $\{|00\rangle, |01\rangle, |1+\rangle, |1-\rangle\}$. A trivial LOCC protocol consists of Alice measuring her qubits and sending the result to Bob: if she obtains 0, then Bob measures in the Z basis, otherwise he measures in the X basis. Notice that, even though the measurement protocol is a legitimate stabilizer protocol, the unitary implementing the measurement in a circuital representation is a controlled Hadamard, which is a non-Clifford unitary.

In the three-qubit scenario, the SHIFT basis in equation (12) constitutes an example of a set exhibiting non-stabilizerness without magic. We reproduce the set here for the reader's convenience

$$\begin{aligned}\mathcal{S} = \{\, &|000\rangle, |{+}01\rangle, |01{+}\rangle, |01{-}\rangle, \\ &|1+0\rangle, |{-}01\rangle, |1-0\rangle, |111\rangle\}\end{aligned} \quad (35)$$

This basis is made of product stabilizer states. However, it is clear that every non-trivial Pauli measurement described by some operator P cannot discriminate all states. For example, suppose that we measure Z_1 first. Then there are two vectors, $|{+}01\rangle$ and $|{-}01\rangle$ which are not perfectly discriminated by this measurement. Indeed, such vectors have, respectively, X_1 and $-X_1$ among the stabilizer generators, and Z_1 anticommutes with both. Thus, the outcome probabilities are both $\frac{1}{2}$. Moreover, it is not difficult to realize that the two post-measurement states are orthogonal. The same conclusions is reached considering every possible pauli measurement $P = P_1 \otimes P_2 \otimes P_3$ that can be performed. Using local measurement, this would be obvious since the protocol would correspond to a LOCC protocol, and this basis is known to exhibit QNLWE.

B. Main result

From the discussion in section II, we know that every unambiguous QNLWE basis $\mathcal{S}$ can be associated with a non-causal process function [10] through the following formula

$$\mathcal{S} = \{\otimes_k U_k^{\omega_k(\boldsymbol{a}_{\setminus k})} |\boldsymbol{a}\rangle \,|\, \boldsymbol{a} \in \{0,1\}^n\}. \quad (36)$$

In the multiqubit case, every product basis is unambiguous and can therefore be expressed as above. The SHIFT basis provides an example of a basis that can be realized in this way. Indeed, consider the unitaries $U^0 = \mathbb{1}$ and $U^1 = H$ for all parties, where H denotes the Hadamard gate, and let $\boldsymbol{\omega} : 0,1^3 \to 0,1^3$ be the Lugano process (equation (5)). One can then easily verify that the SHIFT basis is obtained by applying equation (36). Since the Lugano process is a non-causal tripartite process, a natural question arises: do all stabilizer product bases arising from a non-causal process exhibit non-stabilizerness without magic, as the SHIFT basis does? In this section, we focus on the multiqubit scenario and answer this question positively, by proving the following theorem

Theorem 2. *Every multiqubit stabilizer product basis $\mathcal{S}$ associated with a non-causal process function of the following form*

$$\mathcal{S} = \{\otimes_k H^{\omega_k(\boldsymbol{a}_{\setminus k})} |\boldsymbol{a}\rangle \,|\, \boldsymbol{a} \in \{0,1\}^n\} \quad (37)$$

exhibits non-stabilizerness without magic.

The proof is given in the ancilla-free case. Indeed, as proved in [18] (see Lemma 1 in Appendix A) adding ancillary qubits prepared in stabilizer states do not give any advantage in the discrimination problem. The strategy of the proof is the following: i) for every non trivial measurement Π^P, there exists a state $|\psi_1\rangle$ which cannot be perfectly discriminated after the measurement. This is always guaranteed by the non-causality of the process function; ii) there exists another state $|\psi_2\rangle$ that is no longer orthogonal to $|\psi_1\rangle$. Therefore these two states are already no longer discriminable after the first measurement. The full proof is provided in appendix A. Notice that the statement involves only rotations in the X eigenbasis, but the proof strategy can be adapted to cover cases in which rotations in the Y eigenbasis are also considered.

We notice that the reversed implication is trivially true: bases exhibiting non-stabilizerness without magic are associated with non-causal process functions. By contrapositive, if the process function is causal, then there exists a one-round LOCC protocol made of stabilizer measurements. Therefore, for stabilizer product basis we have an equivalence between non-locality (in the form of QNLWE), non-causality and non-stabilizerness. We finally notice that every non-causal process function is associated with a causal game whose causal bound is maximally violate, the "guess the process function"[10], as described in section II. As a corollary of our main result, causal inequalities then represent a witness of non-stabilizerness in the multiqubit scenario.

IV. CONCLUSIONS

In this work we established a direct operational link between causal structure and non-stabilizerness in the multiqubit stabilizer setting. By exploiting the correspondence between complete product bases and process functions, we showed that whenever the process function associated with a stabilizer product basis is non-causal, the basis cannot be perfectly discriminated using adaptive stabilizer protocols. Therefore, we have extended the result of [18] to every product stabilizer basis other than the SHIFT basis. In this way, causal inequalities acquire a new interpretation: their violation certifies an obstruction to discrimination within the stabilizer formalism, even though the states involved contain no magic. Combined with previous results connecting non-causal process functions and QNLWE, our findings seems to reveal a structural equivalence, at least in the multiqubit case, between three phenomena: QNLWE, non-causality of the associated process function, and non-stabilizerness without magic.

Although our work already reveals an interplay between the discrimination problem of stabilizer product bases and the non-causality of the underlying process function, our analysis has mainly been devoted to the multiqubit case. However, the relation between QNLWE and non-causal process functions has been fully understood for unambiguous product bases [10] in arbitrary dimension. It would be desirable to strengthen our results by proving theorem 2 in the multiqudit scenario as well. In prime dimension, local stabilizer bases are mutually unbiased, and this implies that product bases built out of local stabilizer bases are unambiguous. This ensures the existence of a process function associated with such a basis, which allows one to express the basis in terms of local unitaries as in (36). Thus, in this case one may wonder whether it is possible to extend the proof given for theorem 2 to arbitrary prime dimension. This extension appears to be possible, at least in the ancilla free case, to the multiqudit case. The main problem might be to show that using ancillary qudits does not provide any advantage in the discrimination problem, as shown in the qubit case [18], and we leave this as an open problem.

[1] Scott Aaronson and Daniel Gottesman. Improved simulation of stabilizer circuits. *Phys. Rev. A*, 70:052328, Nov 2004.

[2] Mateus Araújo, Fabio Costa, and Časlav Brukner. Computational advantage from quantum-controlled ordering of gates. *Phys. Rev. Lett.*, 113:250402, Dec 2014.

[3] R. Augusiak, T. Fritz, Ma. Kotowski, Mi. Kotowski, M. Pawłowski, M. Lewenstein, and A. Acín. Tight bell inequalities with no quantum violation from qubit unextendible product bases. *Phys. Rev. A*, 85:042113, Apr 2012.

[4] R. Augusiak, J. Stasińska, C. Hadley, J. K. Korbicz, M. Lewenstein, and A. Acín. Bell inequalities with no quantum violation and unextendable product bases. *Phys. Rev. Lett.*, 107:070401, Aug 2011.

[5] Ämin Baumeler and Stefan Wolf. The space of logically consistent classical processes without causal order. *New Journal of Physics*, 18(1):013036, jan 2016.

[6] Charles H. Bennett, David P. DiVincenzo, Christopher A. Fuchs, Tal Mor, Eric Rains, Peter W. Shor, John A. Smolin, and William K. Wootters. Quantum nonlocality without entanglement. *Phys. Rev. A*, 59:1070–1091, Feb 1999.

[7] Giulio Chiribella. Perfect discrimination of no-signalling channels via quantum superposition of causal structures. *Phys. Rev. A*, 86:040301, Oct 2012.

[8] Giulio Chiribella, Manik Banik, Some Sankar Bhattacharya, Tamal Guha, Mir Alimuddin, Arup Roy, Sutapa Saha, Sristy Agrawal, and Guruprasad Kar. Indefinite causal order enables perfect quantum communication with zero capacity channels. *New Journal of Physics*, 23(3):033039, mar 2021.

[9] Eric Chitambar and Gilad Gour. Quantum resource theories. *Rev. Mod. Phys.*, 91:025001, Apr 2019.

[10] Hippolyte Dourdent, Kyrylo Simonov, Andreas Leitherer, Emanuel-Cristian Boghiu, Ravi Kunjwal, Saronath Halder, Remigiusz Augusiak, and Antonio Acín. Paradox-free classical non-causality and unambiguous non-locality without entanglement are equivalent, 2025.

[11] Daniel Gottesman. Stabilizer codes and quantum error correction, 1997.

[12] Daniel Gottesman. The heisenberg representation of quantum computers, 1998.

[13] D. Gross. Hudson's theorem for finite-dimensional quantum systems. *Journal of Mathematical Physics*, 47(12):122107, 12 2006.

[14] Philippe Allard Guérin, Adrien Feix, Mateus Araújo, and Časlav Brukner. Exponential communication complexity advantage from quantum superposition of the direction of communication. *Phys. Rev. Lett.*, 117:100502, Sep 2016.

[15] Erik Hostens, Jeroen Dehaene, and Bart De Moor. Stabilizer states and clifford operations for systems of arbitrary dimensions and modular arithmetic. *Phys. Rev. A*, 71:042315, Apr 2005.

[16] Simon Kochen and E. P. Specker. The problem of hidden variables in quantum mechanics. *Journal of Mathematics and Mechanics*, 17(1):59–87, 1967.

[17] Ravi Kunjwal and Ämin Baumeler. Trading causal order for locality. *Phys. Rev. Lett.*, 131:120201, Sep 2023.

[18] Hyukjoon Kwon. Nonstabilizerness without magic: Classically simulatable quantum states that are indistinguishable by classically simulatable quantum circuits, 2025.

[19] Simon Milz, Jessica Bavaresco, and Giulio Chiribella. Resource theory of causal connection. *Quantum*, 6:788, August 2022.

[20] Maarten Nest. Classical simulation of quantum computation, the gottesmann-knill theorem, and slightly beyond. *Quantum Information & Computation*, 10:258–271, 03 2010.

[21] Ognyan Oreshkov, Fabio Costa, and Časlav Brukner. Quantum correlations with no causal order. *Nature Communications*, 3:1092 EP –, 10 2012.

[22] J Tolar. On clifford groups in quantum computing. *Journal of Physics: Conference Series*, 1071(1):012022, aug 2018.

[23] Victor Veitch, S A Hamed Mousavian, Daniel Gottesman, and Joseph Emerson. The resource theory of stabilizer quantum computation. *New Journal of Physics*, 16(1):013009, jan 2014.

[24] Victoria J Wright and Ravi Kunjwal. Contextuality in composite systems: the role of entanglement in the Kochen-Specker theorem. *Quantum*, 7:900, January 2023.

[25] Sixia Yu and C. H. Oh. State-independent proof of kochen-specker theorem with 13 rays. *Phys. Rev. Lett.*, 108:030402, Jan 2012.

Appendix A: Proof of Theorem 2

We first state and prove the following lemma

Lemma 1. *Consider a quasi-process function $\boldsymbol{\omega} : \mathcal{A} \to \mathcal{X}$. Consider a bipartition $\mathcal{H}, \mathcal{L} \subseteq \{1, \dots, n\}$ of the parties, $\mathcal{H} \cap \mathcal{L} = \emptyset$ and $\mathcal{H} \cup \mathcal{L} = \emptyset$, for every $\boldsymbol{a}$, set*

$$\begin{aligned} \boldsymbol{a}_{\mathcal{H}} &:= (a_{h_1}, \dots, a_{h_{|\mathcal{H}|}}) \\ \boldsymbol{\omega}_{\mathcal{H}} &:= (\omega_{h_1}, \dots, \omega_{h_{|\mathcal{H}|}}) \end{aligned} \tag{A1}$$

and similarly for ℓ Suppose that there exists $\boldsymbol{a} = (a_1, \dots, a_n)$ such that

$$\begin{aligned} \omega_{\mathcal{H}}(\boldsymbol{a}_{\mathcal{H}} \oplus 1, \boldsymbol{a}_{\mathcal{L}}) &= \omega_{\mathcal{H}}(\boldsymbol{a}_{\mathcal{H}}, \boldsymbol{a}_{\mathcal{L}}) \oplus \mathbf{1}, \\ \omega_{\mathcal{L}}(\boldsymbol{a}_{\mathcal{H}} \oplus 1, \boldsymbol{a}_{\mathcal{L}}) &= \omega_{\mathcal{L}}(\boldsymbol{a}_{\mathcal{H}}, \boldsymbol{a}_{\mathcal{L}}). \end{aligned} \tag{A2}$$

Then $\boldsymbol{\omega}$ is not a process function.

Proof. Set $\boldsymbol{x}_{\mathcal{H}} = \boldsymbol{\omega}_{\mathcal{H}}(\boldsymbol{a})$ and similarly for $\mathcal{L}$. Consider the local interventions defined as

$$\begin{aligned} &f_h(x_h) = a_h, \quad f_h(x_h \oplus 1) = a_h \oplus 1, \quad \forall h \in \mathcal{H}, \\ &f_\ell(x_\ell) = a_\ell = f_\ell(x_\ell \oplus 1), \quad \forall \ell \in \mathcal{L} \end{aligned} \tag{A3}$$

Then, it is immediate to check that

$$(\boldsymbol{x}_{\mathcal{H}}, \boldsymbol{x}_{\mathcal{L}}), \quad (\boldsymbol{x}_{\mathcal{H}} \oplus \mathbf{1}, \boldsymbol{x}_{\mathcal{L}}) \tag{A4}$$

are both fixed points of $\boldsymbol{\omega} \circ \boldsymbol{f}$. Thus $\boldsymbol{\omega}$ is not a process function. □

By contrapositive, the lemma states that for any process function, when the value of the outputs $\boldsymbol{a}_{\mathcal{H}}$ of some parties are flipped, not all the corresponding components of $\boldsymbol{\omega}$ can flip while keeping the components outside of $\mathcal{H}$ fixed.

We now prove theorem 2. Notice that it is pointless to consider measurements P in which some factors correspond to the Y operators. Such a measurement would anticommute with the stabilizer of every vector in $\mathcal{S}$, so that no state would be discriminated by such a measurement. The whole set $\mathcal{S}$ would collapse on one of the 2^{n-1}-dimensional eigenspace of P, which would certainly contain non-orthogonal vectors. We then onsider a measurement P made of tensor product of Pauli operators X and Z only. Let $N_{|\psi\rangle}$ be the number of stabilizer generators of $|\psi\rangle$ which anticommute with P. Now, the process function $\boldsymbol{\omega}$ is non-causal. Let us first assume that condition i) in theorem 1 is violated. Then, for each party k there exists $\boldsymbol{a}^0_{\setminus k}$ and $\boldsymbol{a}^1_{\setminus k}$ such that

$$\omega_k(\boldsymbol{a}^0_{\setminus k}) = 0, \qquad \omega_k(\boldsymbol{a}^1_{\setminus k}) = 1.$$

Thus, both the X and Z eigenbasis appears for each party. Therefore, there always exists a vector $|\psi\rangle$ such that $N_{|\psi\rangle} \neq 0$. For simplicity, let us consider a vector corresponding to the smallest non-zero value of $N_{|\psi\rangle} \equiv N$. Let us write such vector as

$$|\psi\rangle = |\psi_1\rangle \dots |\psi_N\rangle \, |\psi_{N+1}\rangle \dots |\psi_n\rangle \tag{A5}$$

where the parties have been rearranged in such a way that $Pg_\ell = -g_\ell P$ for every $\ell = 1, \dots, N$ and $Pg_\ell = g_\ell P$ for $\ell = N+1 \dots n$. First, notice that the anticommutation of P with a stabilizer generator of $|\psi\rangle$ implies that the outcome probabilities are the same $p(+\,|\psi\rangle) = p(-\,|\psi\rangle) = \frac{1}{2}$. Therefore, this state is not discriminated by the measurement. Let $|a_1\rangle \dots |a_n\rangle$ be the vector of the computational basis mapped to $|\psi\rangle$.

$$\begin{aligned} |\psi\rangle &= \bigotimes_\ell \left(H^{\omega_\ell(\boldsymbol{a}_{\setminus \ell})} \, |a_\ell\rangle \right) = \\ &= \left[\bigotimes_\ell \left(H^{\omega_\ell(\boldsymbol{a}_{\setminus \ell})} \, |a_\ell\rangle \right) \right] \otimes \left[\bigotimes_\ell \left(H^{\omega_\ell(\boldsymbol{a}_{\setminus \ell})} \, |a_\ell\rangle \right) \right] \end{aligned} \tag{A6}$$

We now consider the vector $|\psi'\rangle$ obtained by replacing $a_\ell \to a_\ell \oplus 1$ for every $\ell = 1, \ldots, N$

$$|\tilde{\psi}\rangle = \bigotimes_\ell \left(H^{\omega_\ell(\tilde{\boldsymbol{a}}_{\setminus \ell})} |\tilde{a}_\ell\rangle \right) = \\ = \left[\bigotimes_\ell \left(H^{\omega_\ell(\tilde{\boldsymbol{a}}_{\setminus \ell})} \,|a_\ell \oplus 1\rangle \right) \right] \otimes \left[\bigotimes_\ell \left(H^{\omega_\ell(\tilde{\boldsymbol{a}}_{\setminus \ell})} \,|a_\ell\rangle \right) \right]. \tag{A7}$$

Where $\tilde{\boldsymbol{a}}$ By lemma 1, it is not possible that Notice that, as a consequence of this replacement, there will be some indices ℓ for which the ω_ℓ has changed value. Since we considered a $|\psi\rangle$ corresponding to the smallest non-vanishing N, there will be at least N indices such that the stabilizer generators of the new vector anticommute with P: $\tilde{g}_\ell P = -P\tilde{g}_\ell$. In particular, there are k indices $\ell \in 1, \ldots, N$ such that $\tilde{g}_\ell P = P\tilde{g}_\ell$. Moreover, there are k' indices $\ell \in N+1, \ldots, n$ such that now $\tilde{g}_\ell P = -P\tilde{g}_\ell$. By suitably rearranging the qubits we can write the vectors $|\psi\rangle$ and $|\tilde{\psi}\rangle$ as follows

$$\begin{aligned} |\tilde{\psi}\rangle &= |\tilde{\psi}_{1\ldots k}\rangle |\tilde{\psi}_{k+1\ldots N}\rangle |\tilde{\psi}_{N+1\ldots N+k'}\rangle |\psi_{N+k'+1\ldots n}\rangle \\ |\psi\rangle &= |\psi_{1\ldots k}\rangle |\psi_{k+1\ldots N}\rangle |\psi_{N+1\ldots N+k'}\rangle |\psi_{N+k'+1\ldots n}\rangle \end{aligned} \tag{A8}$$

where $|\phi_{i_1\ldots i_h}\rangle = |\phi_{i_1}\rangle \ldots |\phi_{i_h}\rangle$. We have partitioned the vector so that

1. $\tilde{g}_\ell P = P\tilde{g}_\ell$ and $g_\ell P = -Pg_\ell$ for every $\ell \in \{1, \ldots, k\}$, implying that $|\tilde{\psi}_{1\ldots k}\rangle$ is an eigenvector of $P_1 \otimes \cdots \otimes P_k$;
2. $\tilde{g}_\ell P = -P\tilde{g}_\ell$ and $g_\ell P = -Pg_\ell$ for every $\ell \in \{k+1, \ldots, N\}$;
3. $\tilde{g}_\ell P = -P\tilde{g}_\ell$ and $g_\ell P = Pg_\ell$ for every $\ell \in \{N+1, \ldots, N+k'\}$, implying that $|\psi_{N+1\ldots N+k'}\rangle$ is an eigenvector of $P_{N+1} \otimes \cdots \otimes P_{N+k'}$;
4. $\tilde{g}_\ell P = P\tilde{g}_\ell$ and $g_\ell P = Pg_\ell$ for every $\ell \in \{N+k'+1, \ldots, n\}$, implying that $|\psi_{N+k'+1\ldots n}\rangle$ is an eigenvector of $P_{N+k'+1} \otimes \cdots \otimes P_n$.

We now see that the post-measurement states $\frac{1}{\sqrt{p_\pm}}\Pi^P_\pm|\psi\rangle$ and $\frac{1}{\sqrt{p_\pm}}\Pi^P_\pm|\tilde{\psi}\rangle$ are non-orthogonal. Here $\Pi^P_\pm$. Here, $\Pi^P_\pm$ denotes the projector onto the two eigenspaces of the operator P. We write this operator as follows

$$P = P_{1\ldots k} \otimes P_{k+1\ldots N} \otimes P_{N+1\ldots N+k'} \otimes P_{N+k'+1\ldots n} \tag{A9}$$

where $P_{i_1\ldots i_h} = P_{i_1} \otimes \cdots \otimes P_{i_h}$. Given this partitioning, the projector operators on the eigenspaces corresponding to the eigenvalues $+1$ and -1 are respectively given by

$$\begin{aligned} \Pi^P_+ = \; & \big(\Pi^+ \otimes \Pi^+ \otimes \Pi^+ + \Pi^+ \otimes \Pi^- \otimes \Pi^- + \\ & \Pi^- \otimes \Pi^+ \otimes \Pi^- + \Pi^- \otimes \Pi^- \otimes \Pi^+\big) \otimes \Pi^+ + \\ & \big(\Pi^+ \otimes \Pi^+ \otimes \Pi^- + \Pi^+ \otimes \Pi^- \otimes \Pi^+ + \\ & \Pi^- \otimes \Pi^- \otimes \Pi^+ + \Pi^- \otimes \Pi^- \otimes \Pi^-\big) \otimes \Pi^-, \end{aligned} \tag{A10}$$

for $+1$ and

$$\begin{aligned} \Pi^P_- = \; & \big(\Pi^+ \otimes \Pi^+ \otimes \Pi^+ + \Pi^+ \otimes \Pi^- \otimes \Pi^- + \\ & \Pi^- \otimes \Pi^+ \otimes \Pi^- + \Pi^- \otimes \Pi^- \otimes \Pi^+\big) \otimes \Pi^- + \\ & \big(\Pi^+ \otimes \Pi^+ \otimes \Pi^- + \Pi^+ \otimes \Pi^- \otimes \Pi^+ + \\ & \Pi^- \otimes \Pi^- \otimes \Pi^+ + \Pi^- \otimes \Pi^- \otimes \Pi^-\big) \otimes \Pi^+, \end{aligned} \tag{A11}$$

for -1. Suppose that the measurement result is $+1$, and the other case is treated similarly. In such a case, the overlap between the post-measurement states is given by

$$\begin{aligned} & \langle\psi_{1\ldots k}|\langle\psi_{k+1\ldots N}|\langle\psi_{N+1\ldots N+k'}|\langle\psi_{N+k'+1\ldots n}| \\ & \quad \big[\big(\Pi^+ \otimes \Pi^+ \otimes \Pi^+ + \Pi^+ \otimes \Pi^- \otimes \Pi^- \\ & \quad \Pi^- \otimes \Pi^+ \otimes \Pi^- + \Pi^- \otimes \Pi^- \otimes \Pi^+\big) \otimes \Pi^+ \\ & \quad \big(\Pi^+ \otimes \Pi^+ \otimes \Pi^- + \Pi^+ \otimes \Pi^- \otimes \Pi^+ \\ & \quad \Pi^- \otimes \Pi^- \otimes \Pi^+ + \Pi^- \otimes \Pi^- \otimes \Pi^-\big) \otimes \Pi^-\big] \\ & \quad |\tilde{\psi}\rangle = |\tilde{\psi}_{1\ldots k}\rangle|\tilde{\psi}_{k+1\ldots N}\rangle|\tilde{\psi}_{N+1\ldots N+k'}\rangle|\psi_{N+k'+1\ldots n}\rangle \end{aligned} \tag{A12}$$

Recall now that $|\tilde{\psi}_{1\ldots k}\rangle$, $|\psi_{N+1\ldots N+k'}\rangle$ and $|\psi_{N+k'\ldots n}\rangle$ are eigenvectors of $P_{1\ldots k}$, $P_{N+1\ldots N+k'}$ and $P_{N+k'\ldots n}$ respectively. Therefore, depending on the respective eigenvalues, only one term of the projector contributes to the sum inside the squared brackets. This implies that the overlap is proportional to $\langle\psi_{k+1\ldots N}|\tilde{\Pi}^\pm|\tilde{\psi}_{k+1\ldots N}\rangle$, where

$$\begin{aligned} |\psi_{k+1\ldots N}\rangle &= \bigotimes_\ell \left(H^{\omega_\ell(\boldsymbol{a}_{\setminus \ell})} \,|a_\ell\rangle \right), \\ |\tilde{\psi}_{k+1\ldots N}\rangle &= \bigotimes_\ell \left(H^{\omega_\ell(\tilde{\boldsymbol{a}}_{\setminus \ell})} \,|a_\ell \oplus 1\rangle \right), \\ \tilde{\Pi}^\pm &= \sum_{\substack{s_{k+1},\ldots,s_N \\ s_{k+1}\cdot\cdots\cdot s_N = \pm}} \bigotimes_{\ell=k+1}^{N} \Pi^{s_\ell}. \end{aligned} \tag{A13}$$

$\tilde{\Pi}^\pm$ are the projectors onto the eigenspaces of $P_{k+1,\ldots,N}$. For each factor in the projector, Π^{s_ℓ} is a single qubit projector onto the Z or the X eigenbasis, depending on whether the corresponding operator in P is $P_\ell = Z$ or $P_\ell = X$. For each ℓ, if $P_\ell = X$, $\omega_\ell(\boldsymbol{a}_{\setminus \ell}) = \omega_\ell(\tilde{\boldsymbol{a}}_{\setminus \ell}) = 0$ and viceversa, if $P_\ell = Z$, then $\omega_\ell(\boldsymbol{a}_{\setminus \ell}) = \omega_\ell(\tilde{\boldsymbol{a}}_{\setminus \ell}) = 1$, since by hypothesis, they must be in a basis rotated with respect to the one defined by P_ℓ. The overlap is then proportional to the following sum:

$$\langle\psi_{k+1\ldots N}|\tilde{\Pi}^\pm|\tilde{\psi}_{k+1\ldots N}\rangle = \sum_{\substack{s_{k+1},\ldots,s_N \\ s_{k+1}\cdot\cdots\cdot s_N = \pm}} \prod_{\ell=k+1}^{N} \langle\psi_\ell|\Pi^{s_\ell}|\tilde{\psi}_\ell\rangle \tag{A14}$$

We now see that this sum is non-vanishing. Every term in each product is explicitly given by

$$\langle a_\ell| H^{\omega_\ell(\boldsymbol{a}_{\setminus \ell})} \Pi^{s_\ell} H^{\omega_\ell(\tilde{\boldsymbol{a}}_{\setminus \ell})} |a_\ell \oplus 1\rangle. \tag{A15}$$

We observe that each of these terms is $\frac{1}{2}$ or $-\frac{1}{2}$, depending on whether $s_\ell = +1$ or -1. Indeed, in the latter case, the single qubit projector can be either $\Pi^- = |-\rangle\langle -|$ or $\Pi^- = |1\rangle\langle 1|$, depending on P_ℓ. Therefore, the possible values for the matrix elements are given by

$$\begin{aligned}\langle 0|-\rangle\langle -|1\rangle &= -\frac{1}{2} = \langle 1|-\rangle\langle -|0\rangle,\\ \langle +|1\rangle\langle 1|-\rangle &= -\frac{1}{2} = \langle -|1\rangle\langle 1|+\rangle.\end{aligned} \tag{A16}$$

Now, if the the global projector $\tilde{\Pi}^-$ is the one that contributes in equation (A12), the product $s_1 \cdot \cdots \cdot s_n$ must yield -1 for each term $\prod_{\ell=k+1}^{N} \langle \psi_\ell | \Pi^{s_\ell} | \tilde{\psi}_\ell \rangle$ in the sum of equation (A14). Therefore, each term contains a product of an odd number of matrix elements with $\Pi^s_\ell = \Pi^-$, whence all the terms in the sum are negative and the whole sum is non-vanishing. The same reasoning applies when the measurement outcome is $+1$. We then have found two vectors that are not discriminated by P and such that they become non orthogonal after the measurement.

Now, suppose that ω is non-causal but condition i) is met. Thus, there must exist a party k_1 and b_{k_1} such that the $n-1$-partite process function $\omega^{b_{k_1}} : \mathcal{A}_{\setminus k_1} \to \mathcal{X}_{\setminus k_1}$ is non-causal. If we iterate this, there must exist a subset $\{1, \ldots, n\} \setminus \mathcal{K}$ of the parties such that the output-reduced process function

$$\omega^{\boldsymbol{b}_{\mathcal{K}}} := \omega_{\setminus \mathcal{K}}(\boldsymbol{a}_{\setminus \mathcal{K}}, \boldsymbol{a}_{\mathcal{K}}).$$

violate condition i). The set $\mathcal{K} = \{k_1, k_2, \ldots, k_m\}$ is the set of parties who communicate in a defined causal order: k_1 is in the global past, k_2 can receive a signal from k_1 only, and so on. The corresponding values $\boldsymbol{b}_{\mathcal{K}} = (b_{k_1}, \ldots, b_{k_m})$ are chosen in such a way that $\omega^{\boldsymbol{b}_{\mathcal{K}}}$ violates condition i). Therefore, for the subset of states identified by $\boldsymbol{b}_{\mathcal{K}}$, the same proof of the previous case applies.