Jamming Extensions of Quantum Correlations Lead to Hidden Superluminal Signaling

Ravishankar Ramanathan,[1,*] Xie Sicheng,[1] Michał Eckstein,[2] and Paweł Horodecki[3,4]

[1] *School of Computing and Data Science, The University of Hong Kong, Pokfulam Road, Hong Kong*
[2] *Institute of Theoretical Physics, Faculty of Physics, Astronomy and Applied Computer Science, Jagiellonian University, ul. Łojasiewicza 11, 30-348 Kraków, Poland*
[3] *International Centre for Theory of Quantum Technologies (ICTQT), University of Gdańsk, Jana Bażyńskiego 8, 80-309 Gdańsk*
[4] *Faculty of Applied Physics and Mathematics, Gdańsk University of Technology, Gabriela Narutowicza 11/12, 80-233 Gdańsk, Poland*

Relativistic causality in certain spacetime configurations permits jamming - superluminal causal influences that nevertheless do not enable superluminal signaling. We fully characterize the correlations compatible with relativistic causality (RC) by extending the operator framework of PRL 104, 140404. When the underlying state is required to be quantum, we prove that any nontrivial jamming - whether state-independent or state-dependent - necessarily leads to hidden superluminal signaling. Thus, the only consistent possibilities are standard quantum correlations without jamming or the full relativistically causal correlation set, with intermediate jamming extensions of quantum correlations leading to signaling. As an application of our characterization, we investigate the security of device-independent (DI) cryptographic primitives against adversaries constrained only by relativistic causality. We construct explicit RC attacks that break DI bit commitment and DI secret sharing in jamming geometries, even when these protocols remain secure against no-signaling adversaries. Our results show that security against relativistic adversaries requires behaviors to be specified together with the spacetime locations of their measurement events; input-output statistics alone are insufficient.

Introduction.- Relativistic causality and the no-signaling principle are often treated as interchangeable constraints on physical correlations. However, it has been recently discovered that in multipartite Bell scenarios, they are not equivalent [3–5]. Depending on the spacetime locations of the measurement events, relativistic causality only enforces a subset of the usual no-signaling equalities. In particular in certain configurations, this permits jamming - a form of superluminal causal influence that alters the joint correlations of a subset of parties while leaving individual marginals untouched, thereby preserving the relativistic causal constraint of no superluminal signaling [3–5].

This leads to a fundamental question: can quantum correlations exploit the causal loophole left open under relativistic causality? In other words, do there exist extensions of quantum correlations (intermediate between the quantum set and the full relativistic causal set) that allow nontrivial jamming while remaining compatible with quantum kinematics and the impossibility of superluminal signaling. In this paper, we answer this question in the negative. We first extend the unified operator framework for multipartite correlations (including no-signaling, quantum and classical correlations) by Acin et al. in [1] and characterize the maximal set of correlations compatible with relativistic causality. This $\mathcal{RC}$ set is expressed in this formalism using a Hermitian operator Ω and locally identity-preserving jamming maps Φ, is operationally closed and free of any signaling pathologies. As such, it is a well-defined object (formally a convex polytope) consistent with relativistic causality and worthy of study in a similar fashion to the no-signaling polytope [2, 8, 18]. We then show that when the underlying state is required to be a valid density operator (i.e., when Ω is required to be a positive semidefinite ρ), any attempt to introduce nontrivial jamming maps (in either a state-independent or state-dependent manner) necessarily generates hidden superluminal signaling. This argument is an analog of the classic result by Gisin [30, 32, 33] on the incompatibility of state-dependent maps with no-signaling. Specifically here by hidden superluminal signaling, we mean that a remote party can choose between two ensemble decompositions of the same density operator, and that this choice produces distinguishable output states after a state-dependent jamming evolution. The remote measurement choice can therefore be inferred outside its future light cone, even though the original input-output statistics satisfy the relativistic causality marginal constraints, giving rise to superluminal signaling. The remaining consistent possibilities are therefore the standard set of quantum correlations without jamming or the full relativistically causal correlation set, with intermediate sets leading to signaling.

The set of relativistically causal correlations has cryptographic consequences. Device-Independent (DI) protocols aim to achieve security solely based on the input-output behavior of devices without placing on any trust on their inner workings or quantum characterization. The ultimate security goal here is to achieve security against adversaries only limited by the principle of no-superluminal-signaling. One primary motivation for studying the set $\mathcal{RC}$ has been its implications [4, 7] for the security of DI randomness generation [38–40] and key distribution [36], where attacks have been found in

* ravi@cs.hku.hk

$\mathcal{RC}$ theories exploiting a lack of monogamy of nonlocality. In this work, we investigate the implications for the relativistic security of a different class of primitives, namely mistrustful cryptographic protocols, in which one or more protocol participants may themselves actively deviate from the prescribed strategy. DI protocols have been established for primitives like bit commitment (by Fehr and Fillinger [15]) and secret sharing (by Moreno et al. [14]), and proven secure against no-signaling adversaries under the assumption of independent and identically distributed (i.i.d.) behavior over multiple rounds. We show that these protocols can be rendered insecure in measurement configurations that admit the possibility of jamming. Specifically, we construct explicit relativistic attacks on these protocols, demonstrating that $\mathcal{RC}$ correlations allow active cheating in dishonest cryptography rather than just passive leakage to an external adversary. Notably, these results are not consequences merely of the aforementioned lack of monogamy of nonlocality. These attacks reinforce the notion that input-output statistics alone are insufficient for security against relativistic adversaries, the behaviors must also carry the spacetime labels of their constituent measurement events.

A Unified Framework for Relativistic Causal, No-Signaling, Quantum and Classical Correlations.-

We begin by recalling the set of correlations of interest. For convenience and to avoid cluttering notation, our exposition throughout the paper will mainly focus on the three-player scenario with the generalization to the arbitrary n-player scenario being deferred to Appendix A. Accordingly, we consider a tripartite Bell experiment with three players Alice, Bob and Charlie who, at spacelike separated points (t_A, p_A), (t_B, p_B) and (t_C, p_C) freely and randomly choose measurements labeled by x, y, z with $x \in [m_A], y \in [m_B], z \in [m_C]$ and obtain outcomes $a \in [r_A], b \in [r_B], c \in [r_C]$ respectively, where we adopt the notation $[n] := \{1, 2, \ldots, n\}$. The experiment is characterized by the set of joint probability distributions $P(a, b, c|x, y, z)$ that together constitute the behavior/correlation $\{P(a, b, c|x, y, z)\}$. The set of no-signaling correlations [2, 8, 18] are those that obey the conditions that the marginal distributions seen by any subset of players is independent of the inputs chosen by the complementary set of players. Accordingly, we denote the set $\mathcal{NS}$ as those behaviors $\{P(a, b, c|x, y, z)\}$ for which the Alice-Bob marginal $P(a, b|x, y)$ is independent of z for all a, b, x, y, z, and likewise for all Alice-Charlie and Bob-Charlie marginal probabilities. This set of behaviors $\mathcal{NS}$ constitutes a convex polytope of dimension $dim(\mathcal{NS}) = \prod_{K=A,B,C}[1 + m_K(r_K - 1)] - 1$.

As shown in [3–5, 34], the relativistic causality conditions that enforce no superluminal signaling are not precisely captured by the above no-signaling constraints in general multipartite scenarios. In particular, in certain spacetime configurations of the measurement events that are termed jamming configurations, only a subset of the no-signaling conditions need to be considered. In a jamming configuration such as shown in Fig. 1, where the intersection of the future light-cones of Alice's and Charlie's output events lies inside the future light-cone of Bob's input event, relativistic causality permits the joint correlations of Alice and Charlie to depend on Bob's input, provided their single-party marginals remain independent of that input. While detailed derivations and explanations can be found in [3–5, 34], intuitively the reason is clear. The joint correlations of Alice and Charlie's outcomes are only accessible at a spacetime point in the intersection of their future light-cones. As such, any superluminal causal influence that jams their correlations still does not lead to superluminal signaling - no faster-than-light transmission of information takes place, and the resulting set of correlations remains compatible with relativistic causality. This larger set of behaviors $\{P(a, b, c|x, y, z)\}$ compatible with this weaker set of conditions is termed the set of relativistically causal correlation $\mathcal{RC}$. The set $\mathcal{RC}$ is again a convex polytope, this time of dimension $dim(\mathcal{RC}) = dim(\mathcal{NS}) + (m_B - 1)m_A(r_A - 1)m_C(r_C - 1)$ [4, 7].

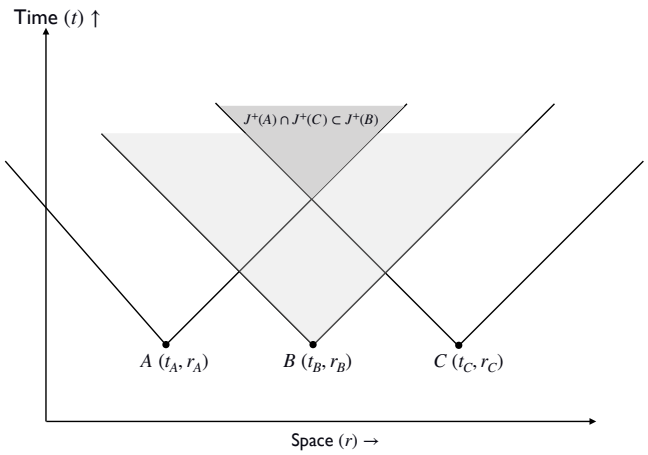

FIG. 1. The Jamming spacetime configuration for three spacelike separated players at locations $A(t_A, r_A)$, $B(t_B, r_B)$ and $C(t_C, r_C)$. In this configuration, the intersection of the future light cones of A and C, denoted $J^+(A) \cap J^+(C)$ is entirely contained within the future light cone of B denoted $J^+(B)$.

In [1], Acín et al. formulated a unified framework of no-signaling, quantum and classical correlations in terms of an elegant operator representation. Our first result is an extension of this framework to include the set of relativistically causal correlations. In particular, we characterize the set of $\mathcal{RC}$ correlations Thm. 1, with a detailed proof provided in Appendix A.

Theorem 1. *An arbitrary tripartite behavior* $\{P(a, b, c|x, y, z)\}$ *satisfies the Relativistic Causality* $\mathcal{RC}$ *constraints for a jamming space-time configuration (where the intersection of Alice and Charlie's future light*

cones is contained within Bob's future light cone) if and only if there exist:

- *Local Hilbert spaces* $\mathcal{H}_A, \mathcal{H}_B, \mathcal{H}_C$ *and sets of local quantum measurements* $\{M_a^x\}, \{M_b^y\}, \{M_c^z\}$ *(that sum to their respective local identities, e.g.,* $\sum_a M_a^x = \mathbb{I}_A$*),*
- *a Hermitian operator* $\Omega_{ABC} \in \mathcal{L}(\mathcal{H}_A \otimes \mathcal{H}_B \otimes \mathcal{H}_C)$ *of unit trace* $(tr(\Omega_{ABC}) = 1)$*, and*
- *a family of linear, trace-preserving maps* $\Phi_y : \mathcal{L}(\mathcal{H}_A \otimes \mathcal{H}_C) \to \mathcal{L}(\mathcal{H}_A \otimes \mathcal{H}_C)$*, parameterized by Bob's setting* y*,*

such that:

$$P(a,b,c|x,y,z) = tr\Big[\Omega_{ABC}\big(\Phi_y(M_a^x \otimes M_c^z) \otimes M_b^y\big)\Big], \quad (1)$$

for all a, b, c, x, y, z*. And the map* Φ_y *for any* y *is locally identity-preserving, i.e., its action on any local observable* $X_A \in \mathcal{L}(\mathcal{H}_A)$ *and* $X_C \in \mathcal{L}(\mathcal{H}_C)$ *satisfies:*

$$\begin{aligned} \Phi_y(X_A \otimes \mathbb{I}_C) &= X_A \otimes \mathbb{I}_C, \\ \Phi_y(\mathbb{I}_A \otimes X_C) &= \mathbb{I}_A \otimes X_C. \end{aligned} \quad (2)$$

Here $\mathcal{L}(\mathcal{H}_A \otimes H)$ denotes the set of all linear operators from the composite Alice-Charlie Hilbert space to itself. When all maps Φ_y are the identity ($\Phi_y = \mathbb{I}_{AC}$), this expression recovers the operator representation of no-signaling correlations from [1]. To emphasize the power of the operator framework and to show the distinction between $\mathcal{RC}$ and $\mathcal{NS}$, we derive in Appendix A the operator representation of an example correlation that belongs to $\mathcal{RC}$ but is outside the set $\mathcal{NS}$. We end this section by emphasizing that the set $\mathcal{RC}$ is convex, free of signaling pathologies, and operationally constitutes the natural relativistically causal generalization of the no-signaling polytope to arbitrary spacetime configurations of measurement events, and that Thm. 1 characterizes this set in an elegant and unified operator framework.

Jamming extensions of quantum correlations.- The physical significance of jamming has been subject to debate recently. In [5] it was argued on the basis of certain fine-tuned configurations that even the subset of relativistically causal constraints that we have outlined above might not rule out causal loops in $1+1$ Minkowski spacetime, while in [21] it was argued on the basis of certain entropic monogamy relations obeyed in such theories that they might lead to superluminal signaling, see [34] as well for clarifications and counter arguments. In this section, we consider the more natural question - if relativistic causality permits jamming, is it possible to extend quantum correlations to incorporate jamming, and are such extensions physical and operationally meaningful? In this context, we remark that besides the no-signaling set which has been studied for its physicality in terms of communication complexity [13], information causality [11] and other reasonable information-theoretic principles [12], so have intermediate sets such as the 'almost quantum' set [9]. As such, it is of significant interest to identify whether a similar intermediate set, extending quantum correlations to incorporate jamming while respecting relativistic causality, is physically plausible and to identify information-theoretic principles to rule out jamming.

Accordingly, we consider the situation wherein the physical state of the system Ω_{ABC} in Thm. 1 is required to be a positive semidefinite Hermitian operator of unit trace $\rho_{ABC} \in \mathcal{L}(\mathcal{H}_A \otimes \mathcal{H}_B \otimes \mathcal{H}_C)$ ($\rho_{ABC} \geq 0$ and $\mathrm{tr}(\rho_{ABC}) = 1$). That is, we consider the set of correlations obtained when in our framework we restrict ourselves to the set of states from standard quantum theory while still allowing for the jamming maps. For technical reasons which we will elaborate on later, we also restrict the local Hilbert spaces $\mathcal{H}_A, \mathcal{H}_B, \mathcal{H}_C$ to be finite-dimensional. We then obtain an extension of quantum theory where the joint probability of outcomes given settings is given by: $P(a,b,c|x,y,z) = \mathrm{tr}[\rho_{ABC}(\Phi_y(M_a^x \otimes M_c^z) \otimes M_b^y)]$, which is identical to quantum theory except for the addition of the jamming map Φ_y. Now to ensure non-negativity of the probabilities $p(a,b,c|x,y,z)$, we must impose that the maps obey $\Big((\Phi_y)_{AC}^\dagger \otimes \mathbb{I}_B\Big)(\rho_{ABC}) \succeq 0$, i.e., $\Phi_y^\dagger(\rho_{AC}) \succeq 0$. We term the resulting correlations, Relativistically Causal Quantum ($\mathcal{RCQ}$) correlations. Formally, we have the following.

Definition 1. *A tripartite behavior* $P(a,b,c|x,y,z)$ *belongs to the set of relativistically-causal quantum (*$\mathcal{RCQ}$*) correlations for a jamming space-time configuration if and only if there exist:*

- *local finite-dimensional Hilbert spaces* $\mathcal{H}_A, \mathcal{H}_B, \mathcal{H}_C$ *and sets of local quantum measurements* $\{M_a^x\}, \{M_b^y\}, \{M_c^z\}$ *(that sum to their respective local identities, e.g.,* $\sum_b M_b^y = \mathbb{I}_B$*),*
- *a positive semidefinite Hermitian operator* $\rho_{ABC} \geq 0$ *in* $\mathcal{L}(\mathcal{H}_A \otimes \mathcal{H}_B \otimes \mathcal{H}_C)$ *of unit trace* $(tr(\rho_{ABC}) = 1)$*,*
- *a family of linear maps* $\Phi_y : \mathcal{L}(\mathcal{H}_A \otimes \mathcal{H}_C) \to \mathcal{L}(\mathcal{H}_A \otimes \mathcal{H}_C)$*, parameterized by Bob's setting* y*, that are trace-preserving* $\Phi_y^\dagger(\mathbb{I}_A \otimes \mathbb{I}_C) = \mathbb{I}_A \otimes \mathbb{I}_C$*, and preserve local identities, i.e., their action on any local observable* $X_A \in \mathcal{L}(\mathcal{H}_A)$ *and* $X_C \in \mathcal{L}(\mathcal{H}_C)$ *satisfies:*

$$\Phi_y(X_A \otimes \mathbb{I}_C) = X_A \otimes \mathbb{I}_C,$$

$$\Phi_y(\mathbb{I}_A \otimes X_C) = \mathbb{I}_A \otimes X_C,$$

and furthermore, the maps Φ_y *preserve positivity of the initial state, i.e.,* $\Phi_y^\dagger(\rho_{AC}) \succeq 0$*,*

such that:

$$P(a,b,c|x,y,z) = tr\Big[\rho_{ABC}\big(\Phi_y(M_a^x \otimes M_c^z) \otimes M_b^y\big)\Big]. \quad (3)$$

The set of correlations $\mathcal{RCQ}$ in Def. 1 is a nontrivial extension of quantum correlations in that there exist RCQ correlations outside the standard quantum set. In Appendix B, we present a paradigmatic example of a correlation in $\mathcal{RCQ}$ (together with the corresponding state ρ_{ABC} and jamming maps Φ_y) that is not in the quantum set. Specifically, in the example Bob jams the correlations between Alice-Charlie, switching from a maximally nonlocal correlation for $y = 0$ (violating the CHSH inequality up to its quantum maximum value) to a fully mixed behavior for $y = 1$.

It is important to note that the distinction with standard quantum theory arises here because of the action of jamming maps Φ_y which are linear trace-preserving but not necessarily CPTP or even just positive. The natural question arises on the physicality of such jamming extensions of quantum theory. Our first result in this section that while the jamming maps (or more precisely their adjoint maps) are required to preserve positivity of the initial state ρ_{AC}, relaxing them to be positive on all states exactly recovers quantum theory. We formalize this via the following theorem.

Theorem 2. *Let the linear trace-preserving maps $\Phi_y : \mathcal{L}(\mathcal{H}_A \otimes \mathcal{H}_C) \to \mathcal{L}(\mathcal{H}_A \otimes \mathcal{H}_C)$ satisfy the Relativistic Causality constraints in (A2). If Φ_y are additionally required to be positive maps for all y, then $\Phi_y \; \forall y$ is strictly the identity map on the joint space $\mathcal{L}(\mathcal{H}_A \otimes \mathcal{H}_C)$, and the behaviors reduce to those obtained in quantum theory.*

Proof. First observe that since Φ_y is a Positive map, its adjoint $\Phi_y^\dagger$ is also a Positive map. This ensures that for any positive semidefinite operator ρ_{AC}, the mapped operator $\Phi_y^\dagger(\rho_{AC})$ remains positive semidefinite.

Now, the Relativistic Causality constraints require Φ_y to be locally identity-preserving as per (A2). This gives

$$\begin{aligned} \mathrm{tr}[\Phi_y^\dagger(\rho_{AC})(X_A \otimes \mathbb{I}_C)] &= \mathrm{tr}[\rho_{AC}\Phi_y(X_A \otimes \mathbb{I}_C)] \\ &= \mathrm{tr}[\rho_{AC}(X_A \otimes \mathbb{I}_C)], \end{aligned} \tag{4}$$

showing that $\mathrm{tr}_C[\Phi_y^\dagger(\rho_{AC})] = \mathrm{tr}_C[\rho_{AC}]$. By similar reasoning for X_C, we also have $\mathrm{tr}_A[\Phi_y^\dagger(\rho_{AC})] = \mathrm{tr}_A[\rho_{AC}]$. Thus, the adjoint map $\Phi_y^\dagger$ perfectly preserves the marginal reduced density matrices of any bipartite input operator.

Now, consider the action of $\Phi_y^\dagger$ on an arbitrary pure product state $\Pi_{AC} = |\psi\rangle\langle\psi|_A \otimes |\phi\rangle\langle\phi|_C$. Since $\Phi_y^\dagger$ is a positive map, the resulting operator $\tilde{\rho}_{AC} = \Phi_y^\dagger(\Pi_{AC})$ is positive semidefinite. Applying the marginal preservation property, we obtain that $\mathrm{tr}_C[\tilde{\rho}_{AC}] = |\psi\rangle\langle\psi|_A$. Since the bipartite state has a perfectly pure marginal, the joint state must factor as a tensor product and the output must take the form $\tilde{\rho}_{AC} = |\psi\rangle\langle\psi|_A \otimes \sigma_C$, for some positive semidefinite σ_C. Similarly applying the marginal preservation property for Alice's subsystem, we obtain that $\mathrm{tr}_A[\tilde{\rho}_{AC}] = \sigma_C = |\phi\rangle\langle\phi|_C$. We therefore see that the map returns the pure product state as

$$\Phi_y^\dagger(|\psi\rangle\langle\psi|_A \otimes |\phi\rangle\langle\phi|_C) = |\psi\rangle\langle\psi|_A \otimes |\phi\rangle\langle\phi|_C.$$

Now, recall that the set of all pure product states forms a basis that spans the entire space $\mathcal{L}(\mathcal{H}_A \otimes \mathcal{H}_C)$. Here we explicitly use the assumption of finite-dimensional Hilbert spaces to guarantee that pure-state projectors ($|\psi\rangle\langle\psi|$) span the real space of Hermitian operators on $\mathcal{H}_A$ and likewise on $\mathcal{H}_C$. Consequently the set of all operators of the form $X_A \otimes Y_C$ with $X_A = X_A^\dagger$ and $Y_C = Y_C^\dagger$ is spanned by elements of the form $|\psi\rangle\langle\psi| \otimes |\phi\rangle\langle\phi|$. Since the linear map $\Phi_y^\dagger$ acts as the identity on a complete spanning set, it follows by linearity that it must act as the identity on all operators, giving :

$$\Phi_y^\dagger \equiv \mathbb{I}_{AC}.$$

Substituting this back into the probability distribution yields:

$$P(a, b, c|x, y, z) = \mathrm{tr}[\rho_{ABC}(M_a^x \otimes M_c^z \otimes M_b^y)],$$

the usual Born rule of quantum theory. □

As we have seen, Thm. 2 rules out any family of jamming maps $\{\Phi_y\}$ that preserve positivity on every quantum state. We are therefore left with the weaker requirement that the maps Φ_y (more precisely their adjoints) need only preserve positivity on the particular state ρ_{AC} shared by Alice and Charlie. We now show that even this state-dependent version is unphysical - it reintroduces hidden superluminal signaling, but via a different mechanism analogous to that originally identified by Gisin [30, 32] for nonlinear quantum mechanics.

Let us make this precise by presenting the physical argument here, with a formal Thm. B.1 and proof provided in App. B. Accordingly, suppose that for every state ρ_{AC} of Alice and Charlie's systems, and every input y, there exists a linear, trace-preserving, locally identity-preserving map $\Phi_y^{\rho_{AC}}$ such that $\left(\Phi_y^{\rho_{AC}}\right)^\dagger(\rho_{AC}) \geq 0$. Now suppose that the state ρ_{AC} admits two different ensemble decompositions:

$$\begin{aligned} \rho_{AC} &= p\rho^0{}_{AC} + (1-p)\rho^1_{AC}, \\ &= q\sigma^0_{AC} + (1-q)\sigma^1_{AC}, \end{aligned} \tag{5}$$

for some $p, q \in (0, 1)$. A remote party David (D) who holds a purification of ρ_{AC} can choose which ensemble is realized by performing a measurement indexed by $w \in \{0, 1\}$ on the purifying system. When the state-dependent map is applied after the ensemble is prepared, the final states seen by Alice and Charlie are $p\left(\Phi_y^{\rho^0_{AC}}\right)^\dagger(\rho^0_{AC}) + (1-p)\left(\Phi_y^{\rho^1_{AC}}\right)^\dagger(\rho^1_{AC})$ when $w = 0$, versus $q\left(\Phi_y^{\sigma^0_{AC}}\right)^\dagger(\sigma^0_{AC}) + (1-q)\left(\Phi_y^{\sigma^1_{AC}}\right)^\dagger(\sigma^1_{AC})$ when $w = 1$. In case these two mixtures are different, Alice and Charlie can distinguish which ensemble was prepared and thereby learn David's measurement choice w. Since David can be space-like separated from Alice and Charlie, this constitutes superluminal signaling.

The only way to prevent such signaling is to demand that the final state seen by Alice and Charlie depend

only on the average state ρ_{AC}, and not on the specific ensemble used to prepare it. In other words, for every input y and every convex decomposition of ρ_{AC} as $\rho_{AC} = p\rho^0_{AC} + (1-p)\rho^1_{AC}$, one must have

$$\left(\Phi_y^{p\rho^0_{AC}+(1-p)\rho^1_{AC}}\right)^\dagger \left(p\rho^0_{AC} + (1-p)\rho^1_{AC}\right) = p\left(\Phi_y^{\rho^0_{AC}}\right)^\dagger (\rho^0_{AC}) + (1-p)\left(\Phi_y^{\rho^1_{AC}}\right)^\dagger (\rho^1_{AC}). \quad (6)$$

Note that the same equality must hold for every ensemble. Eq. 6 is a strong consistency condition that relates the maps belonging to different states. Combined with the already assumed properties of Φ_y - namely, linearity, trace preservation, and the requirement that the map act as the identity on every local subsystem - this condition forces a drastic simplification. As shown formally in Lemma B.1 and Thm. B.1 in Appendix B, the effective evolution is then necessarily the identity on every input state, i.e.,

$$\left(\Phi_y^{\rho_{AC}}\right)^\dagger (\rho_{AC}) = \rho_{AC}, \quad (7)$$

for every ρ_{AC} and every y, and no nontrivial jamming is possible. Therefore, we conclude that no nontrivial jamming extension of quantum correlations is possible - any attempt to introduce nontrivial jamming maps in either state-independent or state-dependent fashion leads to hidden superluminal signaling.

Implications for security of Device-Independent Protocols against Relativistic adversaries.- In a device-independent (DI) protocol, no trust is placed on an internal model for the devices, instead the devices receive inputs and return outputs and security is inferred from the observed conditional distributions $\{P(a,b,c|x,y,z)\}$. DI protocols for cryptographic primitives such as key distribution [36], random number generation [38–40], bit commitment [15] and secret sharing [14] have been constructed, and the ultimate notion of security considered thus far has been that of a no-signaling (NS) adversary, where the set of behaviors is constrained to lie in the $\mathcal{NS}$ polytope. On the other hand, as we have seen, relativistic causality imposes only those marginal constraints that are compatible with the spacetime locations of the corresponding measurement events - such additional freedom can be exploited by a relativistic adversary. Some implications on the monogamy of nonlocality and DI protocols for random number generation and key distribution were considered in [4, 7]. On the other hand, the question was left open for a different class of primitives termed mistrustful cryptographic protocols in which one or more protocol participants themselves may be dishonest. We illustrate this point with two mistrustful primitives which have been shown to be secure against no-signaling adversaries.

Consider first the three-prover bit-commitment protocol of Fehr and Fillinger [15]. Here, a classical verifier interacts with three separated devices A, B and C. The committing party coordinates these devices and uses them to commit to a bit $s \in \{0,1\}$. During the commitment phase, the verifier sends questions to the devices and records their answers. During the opening phase, the committing party announces s together with auxiliary data, and the verifier sends additional test questions to randomly selected devices - the opening is accepted only if the answers satisfy the prescribed consistency checks. The protocol is hiding when the transcript does not reveal s, while it is binding when a dishonest committing party cannot produce accepting openings for both $s = 0$ and $s = 1$ with high probability. It is important to note that the three-prover structure is crucial here - separation prevents the devices from coordinating their answers after receiving the verifier's questions. In a jamming configuration however, one of the devices can influence the joint correlations of the other two without changing either of their marginals. As we show rigorously in Appendix C 1, this extra freedom can be used to coordinate valid opening strings for both committed bits, thereby allowing the committing party to bypass the binding requirement. Specifically, the jamming device's input selects which pair of strings is jointly realized while the single-device statistics and hence the verifier's observed local tests remain unchanged. This allows the adversary to achieve the sum of success probabilities for opening to 0 and 1 in $\mathcal{RC}$ theories as

$$p_0^{RC} + p_1^{RC} \geq 2 - 2\epsilon - 2\delta, \quad (8)$$

where ϵ and δ quantify the protocol's hiding and soundness errors. In the ideal limit with $\epsilon, \delta \to 0$, the right-hand side approaches 2 violating the security bound proven against no-signaling adversaries of $p_0^{NS} + p_1^{NS} \leq 1 + \eta$ with $\eta < 1$ in [15].

An alternative jamming mechanism also breaks DI secret sharing [14]. In the secret sharing protocol, the honest devices produce correlations certifying that a designated group must collaborate in order to reconstruct a secret, while an external devices should have negligible guessing probability of the secret. Here, an RC jammer is able to alter the relevant joint Alice-Charlie correlations even while maintaining the tested statistics, and is able to supply an output that determines the secret perfectly leading to a guessing probability $P_{guess} = 1$. These examples go beyond the lack of monogamy in RC theories noted in [4, 7] and show that security against no-signaling devices is not sufficient to imply security against relativistic adversaries. Notably, the considered behavior must be specified together with the spacetime locations of its measurement events, the input-output statistics alone are insufficient.

Discussion.- Relativistic causality in jamming configurations allow for correlations beyond the standard no-signaling set termed $\mathcal{RC}$ correlations. We proved that the analogous extension that keeps a quantum state and allows such jamming necessarily makes the effective dynamics state-dependent. Under the requirements of ensemble equivalence and no-signaling for all preparations, this state dependence reintroduces superluminal signal-

ing. Consequently the remaining consistent possibilities are ordinary quantum theory (without any jamming) or a fully post-quantum $\mathcal{RC}$ theory built on arbitrary Hermitian operators. The latter is completely characterized by a simple operator representation that generalizes the no-signaling framework of Acin et al. [1].

The relativistic causality framework makes us take geometry-dependent causal constraints seriously. The full no-signaling assumption is an over-idealisation once the actual spacetime locations of the parties are taken into account. Specifically, in DI protocols aiming to be secure against relativistic adversaries, we have shown that the relativistic causal correlations enable perfect attacks on DI bit commitment and secret sharing that bypasses all no-signaling attacks. It reinforces the idea [4, 7] that the standard no-signaling adversary model is insufficient once the causal structure of the experiment is taken into account, and the $\mathcal{RC}$ set of Thm. 1 is the correct replacement.

Other approaches have studied the physicality of jamming mechanisms. For instance, [21] studies the constraints on relativistic causal correlations via entropic monogamy and concludes (by alternative means) that it is likely that physical jamming mechanisms must signal (see also [34] for a discussion and counter-arguments). Our second result is a complementary no-go result - any attempt to realize jamming correlations while retaining a quantum state forces state-dependent dynamics that reintroduce superluminal signaling under standard operational assumptions. It has also been noted in axiomatic reconstructions of quantum theory [41] that combining standard density matrices with non-causal evolutions leads to paradoxes. In this context, our second result might be viewed as a rigorous no-go theorem proving exactly why this is the case for relativistically causal jamming scenarios. In general, a theory in which effects such as jamming occur by point-to-region influences, may eventually turn out to be incompatible with stability, renormalisability or a continuum limit, but that incompatibility has to be clearly demonstrated and understood rather than assumed a priori. As such, it is important to understand the physical and information-theoretical principles that rule out jamming mechanisms and their consequences for security of cryptographic protocols that aim for security against relativistic adversaries.

Acknowledgments.- We acknowledge support from the General Research Fund (GRF) Grant No. 17307925, and the Research Impact Fund (RIF) Grant No. R7035-21.

[1] A. Acín, R. Augusiak, D. Cavalcanti, C. Hadley, J. K. Korbicz, M. Lewenstein, Ll. Masanes, and M. Piani. *Unified Framework for Correlations in Terms of Local Quantum Observables*. Phys. Rev. Lett. 104, 140404 (2010).

[2] Ll. Masanes, A. Acín, N. Gisin. *General properties of Nonsignaling Theories*. Phys. Rev. A. 73, 012112 (2006).

[3] J. Grunhaus, S. Popescu, and D. Rohrlich. *Jamming nonlocal quantum correlations*. Phys. Rev. A 53, 3781–3784 (1996).

[4] P. Horodecki, and R. Ramanathan. *Relativistic Causality vs. No-Signaling as the limiting paradigm for correlations in physical theories*. Nat. Commun. 10, 1701 (2019).

[5] V. Vilasini, and R. Colbeck. *Impossibility of superluminal signaling in Minkowski spacetime does not rule out causal loops*. Phys. Rev. Lett. 129, 110401 (2022).

[6] A. Fine. *Joint distributions, quantum correlations, and commuting observables*. J. Math. Phys. 23, 1306–1310 (1982).

[7] R. Salazar, M. Kamon, D. Goyeneche, K. Horodecki, D. Saha, R. Ramanathan, and P. Horodecki. *A No-go theorem for device-independent security in relativistic causal theories*. Phys. Rev. Research 3, 033146 (2021).

[8] S. Pironio, J.-D. Bancal, and V. Scarani. *Extremal correlations of the tripartite no-signaling polytope*. J. Phys. A: Math. Theor. 44, 065303 (2011).

[9] M. Navascués, Y. Guryanova, M. J. Hoban, and A. Acín. *Almost quantum correlations*. Nat. Commun. 6, 6288 (2015).

[10] J. Allcock, N. Brunner, N. Linden, S. Popescu, P. Skrzypczyk, and T. Vertesi. *Closed sets of nonlocal correlations*. Phys. Rev. A 80, 062107 (2009).

[11] M. Pawlowski, T. Paterek, D. Kaszlikowski, V. Scarani, A. Winter, and M. Zukowski. *Information causality as a physical principle*. Nature 461, 1101 (2009).

[12] T. Fritz, A. B. Sainz, R. Augusiak, J. B. Brask, R. Chaves, A. Leverrier, A. Acín. *Local orthogonality as a multipartite principle for quantum correlations*. Nat. Commun. 4, 2263 (2013).

[13] G. Brassard, H. Buhrman, N. Linden, A. A. Methot, A. Tapp, and F. Unger. *Limit on nonlocality in any world in which communication complexity is not trivial*. Phys. Rev. Lett. 96, 250401 (2006).

[14] M. G. M. Moreno, S. Brito, R. V. Nery, and R. Chaves. *Device-Independent Secret Sharing and a Stronger Form of Bell Nonlocality*. Phys. Rev. A 101, 052339 (2020).

[15] S. Fehr and M. Fillinger. *Multi-Prover Commitments Against NonSignaling Attacks*. In *Advances in Cryptology - CRYPTO 2015*, Part II, pp. 403–421, Springer (2015).

[16] M.-D. Choi. *A Schwarz inequality for positive linear maps on C^*-algebras*. Illinois Journal of Mathematics, 18, 565-574 (1974).

[17] V. I. Paulsen. *Completely Bounded Maps and Operator Algebras*. Cambridge Studies in Advanced Mathematics 78, Cambridge University Press (2003).

[18] S. Popescu and D. Rohrlich. *Quantum nonlocality as an axiom*. Found. Phys. 24, 379-385 (1994).

[19] S. Popescu. *Nonlocality beyond quantum mechanics*. Nat. Phys. 10, 264 (2014).

[20] J. Barrett. *Information processing in generalized probabilistic theories*. Phys. Rev. A 75, 032304 (2007).

[21] M. Weilenmann. *Monogamy relations for relativistically causal correlations*. Nat. Commun. 16, 269 (2025).

[22] B. Toner. *Monogamy of non-local quantum correlations*. Proc. R. Soc. A 465, 59 (2009).

[23] R. Ramanathan, and P. Mironowicz. *Trade-offs in multiparty Bell inequality violations in qubit networks*. Phys.

Rev. A 98, 022133 (2018).

[24] M. C. Tran, R. Ramanathan, M. McKague, D. Kaszlikowski, and T. Paterek. *Bell monogamy relations in arbitrary qubit networks*. Phys. Rev. A 98, 052325 (2018).

[25] A. Sen and F. D. Santo. *Superluminal transformations and indeterminism*. New J. of Phys. 28, 074514 (2026).

[26] V. Vilasini and R. Renner. *Fundamental Limits for Realizing Quantum Processes in Spacetime*. Phys. Rev. Lett. 133, 080201 (2024).

[27] T. van der Lugt. *Causal and compositional structure in quantum theory*. Ph.D. thesis, University of Oxford (2025).

[28] J. Butterfield. *Reconsidering Relativistic Causality*. International Studies in the Philosophy of Science 21(3), 295–328 (2007).

[29] C. Brukner. *Quantum causality*. Nat. Phys. 10, 4, pp. 259–263 (2014).

[30] N. Gisin. *Weinberg's non-linear quantum mechanics and supraluminal communications*. Phys. Lett. A 143, 1–2 (1990).

[31] D. S. Abrams, and S. Lloyd. *Nonlinear quantum mechanics implies polynomial-time solution for NP-complete and #P problems*. Phys. Rev. Lett. 81, 3992–3995 (1998).

[32] C. Simon, V. Buzek, and N. Gisin. *The no-signaling condition and quantum dynamics*. Phys. Rev. Lett. 87, 170405 (2001).

[33] A. Bassi and K. Hejazi. *No-faster-than-light-signaling implies linear evolution. A re-derivation*. Eur. J. of Phys. 36 (5), 055027 (2015).

[34] M. Eckstein, T. Miller, R. Horodecki, R. Ramanathan, and P. Horodecki. *The operational no-signalling constraints and their implications*. arXiv:2512.23702 (2025).

[35] O. Oreshkov, F. Costa, and C. Brukner. *Quantum correlations with no causal order*. Nat. Commun. 3, 1092 (2012).

[36] J. Barrett, L. Hardy, and A. Kent. *No signaling and quantum key distribution*. Phys. Rev. Lett. 95, 010503 (2005).

[37] R. D. Sorkin. *Impossible measurements on quantum fields*. Directions in General Relativity: Proceedings of the 1993 International Symposium, Maryland, Vol. 2, pp. 293-305, Cambridge University Press (1993).

[38] R. Colbeck. *Quantum and relativistic protocols for secure multi-party computation*. Ph.D. thesis, University of Cambridge (2007). arXiv:0911.3814.

[39] F. G. S. L. Brandão, R. Ramanathan, A. Grudka, K. Horodecki, M. Horodecki, P. Horodecki, T. Szarek, H. Wojewódka. *Realistic noise-tolerant randomness amplification using finite number of devices*. Nat. Commun. 7, 11345 (2016).

[40] R. Ramanathan, Y. Liu, and Y. Wu. *Accumulation of Device-Independent Quantum Randomness against Time-Ordered No-Signalling Adversaries*. arXiv:2506.17020 (2025).

[41] G. Chiribella, G. M. D'Ariano, P. Perinotti. *Informational derivation of Quantum Theory*. Phys. Rev. A 84, 012311 (2011).

[42] L. P. Hughston, R. Jozsa, W. K. Wootters. *A complete classification of quantum ensembles having a given density matrix*. Phys. Lett. A 183, 14–18 (1993).

[43] N. Gisin. *Stochastic quantum dynamics and relativity*. Helv. Phys. Acta 62, 363–371 (1989).

[44] D. Mayers. *Unconditionally secure quantum bit commitment is impossible*. Phys. Rev. Lett. 78, 3414 (1997).

[45] H.-K. Lo, H. F. Chau. *Is quantum bit commitment really possible?*. Phys. Rev. Lett. 78, 3410–3413 (1997).

[46] M. Ben-Or, S. Goldwasser, J. Kilian, A. Wigderson. *Multi-prover interactive proofs: how to remove intractability assumptions*. Proc. 20th Ann. ACM Symp. on Theory of Computing, 113–131 (1988).

[47] G. Svetlichny. *Distinguishing three-body from two-body nonseparability by a Bell-type inequality*. Phys. Rev. D 35, 3066 (1987).

Appendix A: Unified Operator Framework for Relativistic Causal Correlations

Theorem A.1. *An arbitrary tripartite behavior $\{P(a,b,c|x,y,z)\}$ satisfies the Relativistic Causality $\mathcal{RC}$ constraints for a jamming space-time configuration (where the intersection of Alice and Charlie's future light cones is contained within Bob's future light cone) if and only if there exist:*

- *Local Hilbert spaces $\mathcal{H}_A, \mathcal{H}_B, \mathcal{H}_C$ and sets of local quantum measurements $\{M_a^x\}, \{M_b^y\}, \{M_c^z\}$ (that sum to their respective local identities, e.g., $\sum_b M_b^y = \mathbb{I}_B$),*
- *a Hermitian operator $\Omega_{ABC} \in \mathcal{L}(\mathcal{H}_A \otimes \mathcal{H}_B \otimes \mathcal{H}_C)$ of unit trace,*
- *a family of linear, trace-preserving maps $\Phi_y : \mathcal{L}(\mathcal{H}_A \otimes \mathcal{H}_C) \to \mathcal{L}(\mathcal{H}_A \otimes \mathcal{H}_C)$, parameterized by Bob's setting y,*

such that:

$$P(a,b,c|x,y,z) = tr\Big[\Omega_{ABC}\big(\Phi_y(M_a^x \otimes M_c^z) \otimes M_b^y\big)\Big], \tag{A1}$$

for all a,b,c,x,y,z. And the map Φ_y for any y is locally identity-preserving, i.e., its action on any local observable $X_A \in \mathcal{L}(\mathcal{H}_A)$ and $X_C \in \mathcal{L}(\mathcal{H}_C)$ satisfies:

$$\begin{aligned} \Phi_y(X_A \otimes \mathbb{I}_C) &= X_A \otimes \mathbb{I}_C, \\ \Phi_y(\mathbb{I}_A \otimes X_C) &= \mathbb{I}_A \otimes X_C. \end{aligned} \tag{A2}$$

Proof. Sufficiency (If). We first prove that if the behavior is generated as per (A1), it satisfies all the RC constraints required to prevent causal loops [4].

We can readily check that:

$$\begin{aligned} \sum_a P(a,b,c|x,y,z) &= \sum_a \mathrm{tr}\Big(\Omega_{ABC}\big(\Phi_y(M_a^x \otimes M_c^z) \otimes M_b^y\big)\Big) \\ &= \mathrm{tr}\Big(\Omega_{ABC}\big(\Phi_y(\mathbb{I}_A \otimes M_c^z) \otimes M_b^y\big)\Big) \\ &= \mathrm{tr}\Big(\Omega_{ABC}\big(\mathbb{I}_A \otimes M_c^z \otimes M_b^y\big)\Big) = P(b,c|y,z). \end{aligned} \tag{A3}$$

Here we have used that $\sum_a M_a^x = \mathbb{I}_A$ and that $\Phi_y(\mathbb{I}_A \otimes M_c^z) = \mathbb{I}_A \otimes M_c^z$. The result is manifestly independent of x, precluding Alice from signaling to Bob and Charlie. By symmetry, summing over Charlie's outcomes yields $P(a,b|x,y)$, which is independent of z.

Similarly, we check that

$$\begin{aligned} \sum_{b,c} P(a,b,c|x,y,z) &= \mathrm{tr}\Big(\Omega_{ABC}\big(\Phi_y(M_a^x \otimes \mathbb{I}_C) \otimes \mathbb{I}_B\big)\Big) \\ &= \mathrm{tr}\Big(\Omega_{ABC}\big(M_a^x \otimes \mathbb{I}_C \otimes \mathbb{I}_B\big)\Big) = P(a|x). \end{aligned} \tag{A4}$$

Here we have again used that $\Phi_y(M_a^x \otimes \mathbb{I}_C) = M_a^x \otimes \mathbb{I}_C$. We verify that this marginal is independent of Bob and Charlie's settings y and z. By symmetry, Charlie's marginal distribution is independent of Alice and Bob's settings x and y. Furthermore,

$$\begin{aligned} \sum_{a,c} P(a,b,c|x,y,z) &= \sum_{a,c} \mathrm{tr}\Big(\Omega_{ABC}\big(\Phi_y(M_a^x \otimes M_c^z) \otimes M_b^y\big)\Big) \\ &= \mathrm{tr}\Big(\Omega_{ABC}\big(\mathbb{I}_A \otimes \mathbb{I}_C \otimes M_b^y\big)\Big) = P(b|y). \end{aligned} \tag{A5}$$

Here again we have used that $\Phi_y(M_a^x \otimes \mathbb{I}_C) = M_a^x \otimes \mathbb{I}_C$. We verify that Bob's marginal distribution is independent of Alice and Charlie's settings x and z.

Now, we have that

$$P(a,c|x,y,z) = \mathrm{tr}\Big(\Omega_{ABC}\big(\Phi_y(M_a^x \otimes M_c^z) \otimes \mathbb{I}_B\big)\Big). \tag{A6}$$

Here Φ_y maps $M_a^x \otimes M_c^z$ to a y-dependent operator, so that the joint AC correlation depends on y, as per the jamming condition. We conclude that the generated behavior satisfies all the RC constraints and at the same time allows for a jamming influence from Bob to the joint correlations between Alice and Charlie's systems.

Necessity (Only If). We now proceed to show that any behavior $P(a,b,c|x,y,z)$ satisfying the RC constraints can be written in terms of (A1) with a suitable operator Ω_{ABC} and a linear trace-preserving map Φ_y. We follow the proof of [1] here, we first choose local Hilbert spaces $\mathcal{H}_A, \mathcal{H}_B, \mathcal{H}_C$ and for each party we define a set of local quantum measurements (POVMs) such that for all settings, the operators excluding the final outcome—e.g., $\{M_a^x\}_{a<r_A}$—along with the local identity $\mathbb{I}_A$, form a linearly independent set. The final outcome is determined by $M_{r_A}^x = \mathbb{I}_A - \sum_{a<r_A} M_a^x$. Now, since these primal operators are linearly independent, there exists a uniquely defined dual basis for each party [1]. For Alice, this is $\{\tilde{\Lambda}_A, \tilde{M}_a^x\}$, which satisfies the orthogonality conditions:

$$\mathrm{tr}(\tilde{M}_a^x M_{a'}^{x'}) = \delta_{aa'}\delta_{xx'}, \qquad \mathrm{tr}(\tilde{\Lambda}_A \mathbb{I}_A) = 1,$$

$$\mathrm{tr}(\tilde{M}_a^x \mathbb{I}_A) = 0, \qquad \mathrm{tr}(\tilde{\Lambda}_A M_a^x) = 0.$$

We define similar primal/dual bases for Bob ($\{\mathbb{I}_B, M_b^y\}$, $\{\tilde{\Lambda}_B, \tilde{M}_b^y\}$) and Charlie ($\{\mathbb{I}_C, M_c^z\}$, $\{\tilde{\Lambda}_C, \tilde{M}_c^z\}$).

Now define a fully non-signaling reference behavior $Q(a,b,c|x,y,z)$ that shares all the 1-body and 2-body marginals of $P(a,b,c|x,y,z)$ that involve Bob. Specifically consider the behavior

$$Q(a,b,c|x,y,z) = \frac{P(a,b|x,y)P(b,c|y,z)}{P(b|y)}, \tag{A7}$$

defined for all b, y such that $P(b|y) \neq 0$ and utilized in the proofs in [6, 7]. Since Q is explicitly no-signaling, by the result of [1] we can map it into an operator Ω_{ABC} of unit trace using the expansion:

$$\Omega_{ABC} = \sum_{a,b,c,x,y,z} Q(a,b,c|x,y,z)\, \tilde{M}_a^x \otimes \tilde{M}_b^y \otimes \tilde{M}_c^z + \sum_{a,b,x,y} Q(a,b|x,y)\, \tilde{M}_a^x \otimes \tilde{M}_b^y \otimes \tilde{\Lambda}_C + \cdots + \tilde{\Lambda}_A \otimes \tilde{\Lambda}_B \otimes \tilde{\Lambda}_C. \tag{A8}$$

We check that by the orthogonality of the dual basis, we have that $\mathrm{tr}(\Omega_{ABC}\,[M_a^x \otimes M_b^y \otimes M_c^z)) = Q(a,b,c|x,y,z]$.

To account for the additional jamming influence, we define the linear map Φ_y in terms of a perturbation as:

$$\Phi_y(M_a^x \otimes M_c^z) = M_a^x \otimes M_c^z + \Delta_{ac}^{xz}(y).$$

We require this perturbation operator to obey the constraint:

$$\begin{aligned} \mathrm{tr}\Big[\Omega_{ABC}\big(\Phi_y(M_a^x \otimes M_c^z) \otimes M_b^y\big)\Big] &= \mathrm{tr}\Big[\Omega_{ABC}\big((M_a^x \otimes M_c^z + \Delta_{ac}^{xz}(y)) \otimes M_b^y\big)\Big] \\ &= P(a,b,c|x,y,z). \end{aligned} \tag{A9}$$

Since the unperturbed trace yields Q, we obtain the explicit constraint for the perturbation as:

$$\mathrm{tr}\Big[\Omega_{ABC}\big(\Delta_{ac}^{xz}(y) \otimes M_b^y\big)\Big] = P(a,b,c|x,y,z) - Q(a,b,c|x,y,z) \equiv D(a,b,c|x,y,z). \tag{A10}$$

Finally, to guarantee that $\Phi_y(X_A \otimes \mathbb{I}_C) = X_A \otimes \mathbb{I}_C$, the perturbation $\Delta_{ac}^{xz}(y)$ must vanish when traced against any local identity. We enforce this algebraically by expanding Δ in the non-identity primal matrices:

$$\Delta_{ac}^{xz}(y) = \sum_{x',a'<r_A} \sum_{z',c'<r_C} \Gamma_{a'c'}^{x'z'}(a,c,x,z,y)\, M_{a'}^{x'} \otimes M_{c'}^{z'}. \tag{A11}$$

Since $M_{a'}^{x'}$ and $M_{c'}^{z'}$ are orthogonal to the dual identity elements $\tilde{\Lambda}_A$ and $\tilde{\Lambda}_C$, this guarantees Δ only acts on the bipartite correlation subspace, isolating Bob's jamming from Alice and Charlie's local marginals.

We can now solve for the coefficients Γ, by substituting the primal expansion of Δ in (A11) into the constraint (A10). Since Ω_{ABC} in (A8) is written in the dual basis, tracing it with the primal matrices $M_{a'}^{x'} \otimes M_b^y \otimes M_{c'}^{z'}$ directly extracts the coefficients $Q(a',b,c'|x',y,z')$ and we obtain:

$$\sum_{a',c',x',z'} \Gamma_{a'c'}^{x'z'}(a,c,x,z,y) \cdot Q(a',b,c'|x',y,z') = D(a,b,c|x,y,z). \tag{A12}$$

This forms a system of linear equations. Specifically for a fixed (a,c,x,y,z), Eq.(A12) represents a system of linear equations evaluated across all of Bob's non-deterministic outcomes $b < r_B$. We write this as a matrix equation

$\mathbf{Q}\,\vec{\Gamma} = \vec{D}$, where $\vec{D}$ is a vector of length $(r_B - 1)$, $\vec{\Gamma}$ is a vector of length $N = m_A(r_A - 1) \times m_C(r_C - 1)$, and $\mathbf{Q}$ is a matrix of dimensions $(r_B - 1) \times N$ where each row corresponds to a specific outcome b and each column corresponds to an evaluation of the reference behavior Q at the basis variables (a', c', x', z').

To guarantee that an exact solution for the vector $\vec{\Gamma}$ exists, the matrix Q must possess full row rank (rank $r_B - 1$). If Q is built using only the physical measurement settings x and z, this full rank is not guaranteed. To force full row rank without altering the physical behavior, the local measurement sets for Alice and Charlie are expanded to include $K = r_B - 1$ auxiliary dummy settings: $\{x_1^*, \ldots, x_K^*\}$ and $\{z_1^*, \ldots, z_K^*\}$ with binary outcomes $a', c' \in \{0, 1\}$. Since these are mathematical constructs, their reference distribution Q can be engineered arbitrarily, provided it remains non-signaling and preserves Bob's physical marginal $P(b|y)$. To be specific, we define these as follows. For each dummy input pair corresponding to $k \in \{1, \ldots, r_B - 1\}$, we define the joint reference probabilities as follows:

$$Q(0, b, 0|x_k^*, y, z_k^*) = \frac{1}{4}P(b|y) + \epsilon \cdot P(b|y)\delta_{bk},$$

$$Q(1, b, 0|x_k^*, y, z_k^*) = \frac{1}{4}P(b|y) - \epsilon \cdot P(b|y)\delta_{bk}$$

$$Q(0, b, 1|x_k^*, y, z_k^*) = \frac{1}{4}P(b|y) - \epsilon \cdot P(b|y)\delta_{bk},$$

$$Q(1, b, 1|x_k^*, y, z_k^*) = \frac{1}{4}P(b|y) + \epsilon \cdot P(b|y)\delta_{bk},$$

where $\epsilon > 0$ is a sufficiently small constant (e.g., $\epsilon \leq 1/4$) ensuring all probabilities remain non-negative.

We verify that this block is fully non-signaling by computing:

$$Q(0, b, c'|x_k^*, y, z_k^*) + Q(1, b, c'|x_k^*, y, z_k^*) = \frac{1}{2}P(b|y).$$

This is independent of Alice's setting x_k^*, meaning Alice cannot signal to Bob or Charlie. Similarly we compute

$$Q(a', b, 0|x_k^*, y, z_k^*) + Q(a', b, 1|x_k^*, y, z_k^*) = \frac{1}{2}P(b|y).$$

This is independent of Charlie's setting z_k^*. Finally, Bob's isolated marginal (obtained by summing over a' and c') yields exactly $P(b|y)$.

We now isolate a specific submatrix of $\mathbf{Q}$ by looking at the columns corresponding to the outcomes $a' = 0, c' = 0$ evaluated at the correlated dummy settings (x_k^*, z_k^*). Let column k (for $k = 1, \ldots, r_B - 1$) be the vector $\vec{v}_k$ representing the equation coefficients evaluated across all rows b:

$$\vec{v}_k = \left[Q(0, b, 0|x_k^*, y, z_k^*)\right]_{b=1}^{r_B-1} = \frac{1}{4}\vec{P}(b|y) + \epsilon P(k|y)\hat{e}_k,$$

where $\hat{e}_k$ is the standard basis vector with a 1 at position k and 0 elsewhere. The square $(r_B - 1) \times (r_B - 1)$ submatrix formed by these columns is:

$$\mathbf{Q}_{\text{sub}} = \frac{1}{4}\vec{P}(b|y) \cdot \vec{1}^T + \epsilon \cdot \text{diag}(P(1|y), P(2|y), \ldots, P(r_B - 1|y)). \tag{A13}$$

This $\mathbf{Q}_{\text{sub}}$ is a diagonal matrix of positive entries (since $P(b|y) > 0$) combined with a rank-1 update (the outer product $\frac{1}{4}\vec{P}(b|y)\mathbf{1}^T$). By the Matrix Determinant Lemma, this guarantees that Q_{sub} is invertible. Now since $\mathbf{Q}$ contains a fully invertible $(r_B - 1) \times (r_B - 1)$ submatrix, the entire matrix $\mathbf{Q}$ is guaranteed to possess full row rank. By comparing the rank of the coefficient matrix and the augmented matrix, we see that the surjective linear map $\mathbf{Q}\,\vec{\Gamma} = \vec{D}$ will always yield at least one exact solution vector $\vec{\Gamma}$. The coefficients $\Gamma_{a'c'}^{x'z'}$ are obtained simply by computing the pseudo-inverse (or directly inverting the submatrix and setting all other coefficients to zero):

$$\vec{\Gamma}_{\text{sub}} = \mathbf{Q}_{\text{sub}}^{-1}\vec{D}.$$

Finally, we verify Bob's final outcome ($b = r_B$). Summing the constraint in (A10) over all b yields:

$$\begin{aligned}
&\sum_b \mathrm{tr}\Big[\Omega_{ABC}\big(\Delta^{xz}_{ac}(y)\otimes M^y_b\big)\Big] \\
&= \mathrm{tr}\Big[\Omega_{ABC}\big(\Delta^{xz}_{ac}(y)\otimes \mathbb{I}_B\big)\Big] \\
&= \sum_b D(a,b,c|x,y,z) = P(a,c|x,y,z) - Q(a,c|x,z).
\end{aligned} \tag{A14}$$

This maps Bob's input parameter y onto the joint AC marginal, indicating the jamming.

□

Remark that in the limit when $\Phi_y = \mathbb{I}_{AC}$ (the identity superoperator), the Thm. A.1 recovers the framework of [1] for fully no-signalling correlations.

As an illustrative example of Thm. A.1, we identify the operator Ω_{ABC}, the measurements $\{M^x_a\}$, $\{M^y_b\}$ and $\{M^z_c\}$, and the maps Φ_y that produce the paradigmatic RC box given as Example 4 in [4]. Specifically, this example considers a Bell scenario with three players Alice, Bob and Charlie in the jamming configuration. Each player performs one of two measurements with binary outputs resulting in the box $P(a,b,c|x,y,z)$. The box is characterised by deterministic output $b = 0$ for both inputs $y = 0, 1$ of Bob, with the marginal $P(a,c|xx,y,z)$ being a local box that returns uniformly random outcomes $a = c$ for all x, z when Bob inputs $y = 0$, and the marginal $P(a,c|x,y,z)$ being a Popescu-Rohrlich (PR) box [18] that obeys $a \oplus c = x \cdot z$ when Bob inputs $y = 1$.

To build this behavior within the framework of Thm. A.1, we formally present the following example.

Example A.1. *Let the Hilbert spaces for Alice, Bob, and Charlie correspond to qubits: $\mathcal{H}_A = \mathcal{H}_B = \mathcal{H}_C = \mathbb{C}^2$. As the operator Ω_{ABC}, consider*

$$\Omega_{ABC} = |\Phi^+\rangle\langle\Phi^+|_{AC} \otimes |0\rangle\langle 0|_B. \tag{A15}$$

Specifically, the AC state is

$$\rho_{AC} = \frac{1}{4}(\mathbb{I}\otimes\mathbb{I} + \sigma_x\otimes\sigma_x - \sigma_y\otimes\sigma_y + \sigma_z\otimes\sigma_z).$$

As the local measurement operators, we choose the following. Alice measures in the standard Pauli Z and X bases: $M^0_a = \frac{1}{2}(\mathbb{I}_A + (-1)^a\sigma_z)$, $M^1_a = \frac{1}{2}(\mathbb{I}_A + (-1)^a\sigma_x)$. Charlie measures in a rotated bases given by $M^0_c = \frac{1}{2}(\mathbb{I}_C + (-1)^c\frac{\sigma_z+\sigma_x}{\sqrt{2}})$, $M^1_c = \frac{1}{2}(\mathbb{I}_C + (-1)^c\frac{\sigma_z-\sigma_x}{\sqrt{2}})$. Since Bob's output is deterministically $b = 0$ for all inputs, his measurement operators are trivial for both settings $y = 0$ and $y = 1$, namely $M^y_0 = \mathbb{I}_B$, $M^y_1 = 0$.

The jamming maps $\Phi_y : \mathcal{L}(\mathcal{H}_A\otimes\mathcal{H}_C) \to \mathcal{L}(\mathcal{H}_A\otimes\mathcal{H}_C)$ are required to be linear, trace-preserving, and locally identity-preserving. We define these by their action on the basis of joint Pauli operators $\sigma_i\otimes\sigma_j$. We set when $y = 0$:

$$\begin{aligned}
\Phi_0(\sigma_z\otimes\sigma_z) &= \sqrt{2}\sigma_z\otimes\sigma_z, \\
\Phi_0(\sigma_x\otimes\sigma_z) &= \sqrt{2}\sigma_x\otimes\sigma_x, \\
\Phi_0(\sigma_y\otimes\sigma_y) &= 0.
\end{aligned} \tag{A16}$$

With all other Pauli basis operators mapping to themselves (or zero as the degree of freedom involved in the characterization). When $y = 1$, we set:

$$\begin{aligned}
\Phi_1(\sigma_z\otimes\sigma_z) &= \sqrt{2}\sigma_z\otimes\sigma_z, \\
\Phi_1(\sigma_x\otimes\sigma_x) &= \sqrt{2}\sigma_x\otimes\sigma_x, \\
\Phi_1(\sigma_y\otimes\sigma_y) &= 0.
\end{aligned} \tag{A17}$$

Again, all other Pauli basis operators map to themselves or to zero.

We can verify that the exact RC behavior is achieved by computing

$$P(a,b,c|x,y,z) = tr[\Omega_{ABC}(\Phi_y(M^x_a\otimes M^z_c)\otimes M^y_b)].$$

Since $b = 0$ is guaranteed, Bob's subspace trivially traces to 1. Using the cyclic property of the trace, we can apply the adjoint map $\Phi^\dagger_y$ to the AC state instead:

$$P(a,0,c|x,y,z) = tr[\Phi^\dagger_y(\rho_{AC})(M^x_a\otimes M^z_c)].$$

For $y = 0$, we have that $\Phi_0^\dagger(\rho_{AC}) =: \tilde{\rho}_0$ is given by

$$\tilde{\rho}_0 = \frac{1}{4}(\mathbb{I}\otimes\mathbb{I} + \sqrt{2}\sigma_z\otimes\sigma_z + \sqrt{2}\sigma_x\otimes\sigma_z).$$

Evaluating the expectation values of Alice and Charlie's observables on this state yields exactly 1 for all four combinations of $x, z \in \{0,1\}$, so that $a = c$ with probability 1, reproducing the local box from Example 4. For $y = 1$, we have that $\Phi_1^\dagger(\rho_{AC}) =: \tilde{\rho}_1$ is given by

$$\tilde{\rho}_1 = \frac{1}{4}(\mathbb{I}\otimes\mathbb{I} + \sqrt{2}\sigma_z\otimes\sigma_z + \sqrt{2}\sigma_x\otimes\sigma_x).$$

Evaluating the expectation values on this state yields 1 for $(x, z) \in \{(0,0),(0,1),(1,0)\}$ and -1 for $(x, z) = (1,1)$. This reproduces the $a \oplus c = x \cdot z$ PR box correlations required by Example 4 from [4].

The above theorem applies to the paradigmatic three-party jamming configuration highlighted in [3–5]. The definition of relativistic causal correlations readily generalizes to the n-party configuration.

Definition A.1. *(The n-Party $\mathcal{RC}$ Set). Let $N = \{1, 2, \ldots, n\}$ be a set of n space-like separated parties in a spacetime configuration $\mathcal{M}$. Let p_i and q_i denote the input and output events for party $i \in N$, respectively. For any subset of parties $S \subseteq N$, the Jamming Set $\mathcal{J}_\mathcal{M}(S) \subseteq N \setminus S$ is defined as the set of parties j satisfying: $\bigcap_{i\in S} J^+(q_i) \subseteq J^+(p_j)$. A behavior $P(\vec{a}|\vec{x})$ belongs to $\mathcal{RC}$ if and only if there exist:*

- *Finite-dimensional Hilbert spaces $\{\mathcal{H}_i\}_{i\in N}$, with $\mathcal{H}_N = \bigotimes_{i\in N}\mathcal{H}_i$.*
- *For each $i \in N$ and input x_i, a set of POVM elements $\{M_{a_i}^{x_i}\} \subset \mathcal{L}(\mathcal{H}_i)$ satisfying $M_{a_i}^{x_i} \geq 0$ and $\sum_{a_i} M_{a_i}^{x_i} = I_{\mathcal{H}_i}$.*
- *A Hermitian operator $\Omega \in \mathcal{L}(\mathcal{H}_N)$ of unit trace.*

Further, there exists a hierarchy of linear, trace-preserving maps defined for all subsets $S \subseteq N$: $\Phi_{\vec{u}}^{(S)} : \mathcal{L}(\mathcal{H}_S) \to \mathcal{L}(\mathcal{H}_S)$ where the parameter $\vec{u} = \vec{x}_{S\cup\mathcal{J}_\mathcal{M}(S)}$ depends only on the inputs of the parties in S and their valid jammers. This hierarchy satisfies the following three axioms:

1. *For every proper subset inclusion $S \subsetneq T \subseteq N$ and every operator $X_S \in \mathcal{L}(\mathcal{H}_S)$, the map acting on T must perfectly intertwine with the map acting on S. As an equality of operators on $\mathcal{H}_T$, it must hold that:*

$$\Phi^{(T)}_{\vec{x}_{T\cup\mathcal{J}_\mathcal{M}(T)}}\big(X_S \otimes I_{T\setminus S}\big) = \Phi^{(S)}_{\vec{x}_{S\cup\mathcal{J}_\mathcal{M}(S)}}(X_S)\otimes I_{T\setminus S}. \tag{A18}$$

2. *The map must fix all single-party operators that are not acting as jammers for the rest of the subset. For every $S \subseteq N$, every $i \in S$, and every $X_i \in \mathcal{L}(\mathcal{H}_i)$, if $i \notin \mathcal{J}_\mathcal{M}(S \setminus \{i\})$, then:*

$$\Phi^{(S)}_{\vec{x}_{S\cup\mathcal{J}_\mathcal{M}(S)}}\big(X_i \otimes I_{S\setminus\{i\}}\big) = X_i \otimes I_{S\setminus\{i\}}. \tag{A19}$$

3. *If the spacetime configuration permits no jamming at all ($\mathcal{J}_\mathcal{M}(S) = \emptyset$ for all S), then the global map reduces to the identity:*

$$\Phi^{(N)}_{\vec{x}} = \mathrm{id}_{\mathcal{L}(\mathcal{H}_N)} \quad \forall\vec{x}. \tag{A20}$$

The n-party relativistic causal correlations are written as:

$$P(\vec{a}|\vec{x}) = \mathrm{tr}\Big(\Omega\cdot\Phi^{(N)}_{\vec{x}}\Big(\bigotimes_{i=1}^{n} M_{a_i}^{x_i}\Big)\Big). \tag{A21}$$

In simple terms, in an n-party configuration, the first axiom on the jamming maps ensures that the correlations of a smaller subgroup are consistent and the jamming effects on that subgroup are self-contained. That is, it guarantees that the subgroup's measurement operators only depend upon the inputs within the subgroup and the specific parties explicitly permitted by the spacetime configuration to jam them. The second axiom guarantees that the measurement operators of any subset of parties is untouched by the jamming map, unless the spacetime configuration is such that an outside party is permitted to jam them. Furthermore, even an active jammer's own local measurement operator is unaffected by their jamming map. The third axiom simply states that if the spacetime configuration forbids jamming entirely, then the jamming maps reduce to the identity and the system behaves exactly like standard quantum mechanics.

Appendix B: Jamming extension of Quantum Correlations leads to Hidden Superluminal Signaling

In this section, we examine the physicality of the intermediate set of quantum-like correlations provided in the main text and termed $\mathcal{RCQ}$. We recall the definition here for convenience, noting that the distinction with the general $\mathcal{RC}$ correlations of the previous section arises here from the requirement that the state be a positive semidefinite Hermitian operator ρ_{ABC} of unit trace. Again, we stick to the three-party picture here for convenience, the general n-party case following immediately from Def. A.1 by imposing the requirement of positive semidefiniteness for the state. As in the rest of the paper, in the following definition, $\mathcal{L}(\mathcal{H})$ denotes the space of linear operators on the Hilbert space $\mathcal{H}$, $\mathrm{Herm}(\mathcal{H})$ denotes the space of all Hermitian (self-adjoint) operators acting on the Hilbert space $\mathcal{H}$, and $\mathcal{D}(\mathcal{H})$ denotes the space of density operators on the Hilbert space $\mathcal{H}$.

Definition B.1. *A tripartite behavior $P(a,b,c|x,y,z)$ belongs to the set of relativistically-causal quantum ($\mathcal{RCQ}$) correlations for a jamming space-time configuration if and only if there exist:*

- *local finite-dimensional Hilbert spaces $\mathcal{H}_A, \mathcal{H}_B, \mathcal{H}_C$ and sets of local quantum measurements $\{M_a^x\}, \{M_b^y\}, \{M_c^z\}$ (that sum to their respective local identities, e.g., $\sum_b M_b^y = \mathbb{I}_B$),*
- *a positive semidefinite Hermitian operator $\rho_{ABC} \geq 0$ in $\mathcal{L}(\mathcal{H}_A \otimes \mathcal{H}_B \otimes \mathcal{H}_C)$ of unit trace ($tr(\rho_{ABC}) = 1$),*
- *a family of linear maps $\Phi_y : \mathcal{L}(\mathcal{H}_A \otimes \mathcal{H}_C) \to \mathcal{L}(\mathcal{H}_A \otimes \mathcal{H}_C)$, parameterized by Bob's setting y, that are trace-preserving $\Phi_y^\dagger(\mathbb{I}_A \otimes \mathbb{I}_C) = \mathbb{I}_A \otimes \mathbb{I}_C$, and preserve local identities, i.e., their action on any local observable $X_A \in \mathcal{L}(\mathcal{H}_A)$ and $X_C \in \mathcal{L}(\mathcal{H}_C)$ satisfies:*

$$\Phi_y(X_A \otimes \mathbb{I}_C) = X_A \otimes \mathbb{I}_C,$$

$$\Phi_y(\mathbb{I}_A \otimes X_C) = \mathbb{I}_A \otimes X_C,$$

and furthermore, the maps Φ_y preserve positivity of the initial state, i.e., $\Phi_y^\dagger(\rho_{AC}) \succeq 0$,

such that:

$$P(a,b,c|x,y,z) = tr\Big[\rho_{ABC}\big(\Phi_y(M_a^x \otimes M_c^z) \otimes M_b^y\big)\Big]. \tag{B1}$$

The set $\mathcal{RCQ}$ is a strict superset of the usual quantum set in spacetime configurations which permit jamming. As a paradigmatic example, we present the following.

Example B.1. *Suppose that Alice, Bob and Charlie in the jamming configuration share a 3-qubit state*

$$\rho_{ABC} = |\Phi^+\rangle\langle\Phi^+|_{AC} \otimes |0\rangle\langle 0|_B,$$

where $|\Phi^+\rangle = \frac{|00\rangle+|11\rangle}{\sqrt{2}}$, and perform standard dichotomic projective measurements given as follows.

- *Alice ($x \in \{0,1\}$): $M_a^x = \frac{\mathbb{I}+(-1)^a A_x}{2}$, where $A_0 = \sigma_z$ and $A_1 = \sigma_x$,*
- *Bob ($y \in \{0,1\}$): $M_b^y = \frac{\mathbb{I}+(-1)^b B_y}{2}$, where $B_0 = \sigma_z$ and $B_1 = \sigma_x$,*
- *Charlie ($z \in \{0,1\}$): $M_c^z = \frac{\mathbb{I}+(-1)^c C_z}{2}$, where $C_0 = \frac{\sigma_z+\sigma_x}{\sqrt{2}}$ and $C_1 = \frac{\sigma_z-\sigma_x}{\sqrt{2}}$.*

The jamming maps Φ_y are given as follows. Let $\sigma_0 \equiv \mathbb{I}$, and let $\{\sigma_\mu \otimes \sigma_\nu\}_{\mu,\nu=0}^3$ be the orthogonal Pauli basis. For $y = 0$, Bob chooses not to jam

$$\Phi_0 \equiv \mathbb{I}_{AC}.$$

For $y = 1$, we define Φ_1 uniquely by its linear extension on the basis vectors as

$$\Phi_1(\sigma_\mu \otimes \sigma_\nu) = \sigma_\mu \otimes \sigma_\nu \quad \text{if } \mu = 0 \text{ or } \nu = 0,$$

$$\Phi_1(\sigma_\mu \otimes \sigma_\nu) = 0 \quad \text{if } \mu \neq 0 \text{ and } \nu \neq 0.$$

Since Φ_1 maps Hermitian basis operators to themselves or zero, it is self-adjoint ($\Phi_1^\dagger = \Phi_1$). Furthermore, for any Alice observable $X_A = \sum_\mu c_\mu \sigma_\mu$, we evaluate:

$$\Phi_1(X_A \otimes \mathbb{I}) = \sum_\mu c_\mu \Phi_1(\sigma_\mu \otimes \sigma_0).$$

Because $\nu = 0$, the map preserves the term: $\sum_\mu c_\mu(\sigma_\mu \otimes \sigma_0) = X_A \otimes \mathbb{I}$. Similar considerations hold for $\mathbb{I} \otimes X_C$. We compute the joint probability distributions of outcomes given inputs as

$$\begin{aligned} P(a,b,c|x,y,z) &= tr[\rho_{ABC}(\Phi_y(M_a^x \otimes M_c^z) \otimes M_b^y)] \\ &= tr[\Phi_y^\dagger(\rho_{AC})(M_a^x \otimes M_c^z)] \cdot tr[|0\rangle\langle 0|_B M_b^y]. \end{aligned} \tag{B2}$$

We evaluate the effective state $\tilde{\rho}_{AC|y} = \Phi_y^\dagger(|\Phi^+\rangle\langle\Phi^+|)$. For $y = 0$: $\tilde{\rho}_{AC|0} = \Phi_0(|\Phi^+\rangle\langle\Phi^+|) = |\Phi^+\rangle\langle\Phi^+|$. With this combination of state and measurements, one has $\langle CHSH\rangle_{AC|y=0} = 2\sqrt{2}$ for $\langle CHSH\rangle_{AC|y=0} = Tr\left[(A_0C_0 + A_0C_1 + A_1C_0 - A_1C_1)\,\rho_{AC|y=0}\right]$. For $y = 1$, the map $\Phi_1^\dagger$ perfectly annihilates the correlation terms, leaving: $\tilde{\rho}_{AC|1} = \frac{1}{4}(\mathbb{I} \otimes \mathbb{I}) = \frac{\mathbb{I}_{AC}}{4}$. We see that $\tilde{\rho}_{AC|y}$ is positive semidefinite for all y, and the measurement operators are valid POVMs, so that $P(a,b,c|x,y,z) \geq 0$ for all a,b,c,x,y,z. In this case, since the state is fully mixed, we see that one has $\langle CHSH\rangle_{AC|y=1} = 0$.

This set is similar to the $\mathcal{RC}$ set considered in Appendix A except that we now impose that the state is a positive semidefinite Hermitian operator ρ_{ABC} of unit trace. The set differs from the standard set of quantum correlations $\mathcal{Q}$ in allowing for jamming maps Φ_y that purportedly still respect relativistic causality.

To maintain non-negativity of probabilities, it is necessary to impose that $\Phi_y^\dagger(\rho_{AC}) \geq 0$. In Thm. 2, we have seen that if we instead require that $\Phi_y^\dagger$ is positive on every state (and not only the given initial state ρ_{AC}), then the condition of being locally identity preserving enforces that $\Phi_y = \mathbb{I}_{AC}$, i.e., no nontrivial positive jamming maps exist. We now consider the more restrictive condition that $\Phi_y^\dagger$ acts as a state-dependent positive map, i.e., only maintains positivity of the initial shared state ρ_{AC}. We prove that even this state-dependent version of jamming is unphysical, in that it reintroduces hidden superluminal signaling but this time via a different mechanism akin to that noted by Gisin [30, 32] for nonlinear modifications of quantum theory.

We first recall the following Lemma on no-superluminal-signalling implying linearity of evolution, adapted to our particular evolution under jamming maps $\Phi_y^{\rho_{AC}}$. The proof follows arguments by Gisin in [30] and Simon, Buzek and Gisin in [32] (see also the rederivation in [33]). As usual, here an ensemble for a density operator $\rho \in \mathcal{D}(\mathcal{H})$ is a finite collection $\{(\rho^i, p_i)\}_{i\in I}$ where each $\rho^i \in \mathcal{D}(\mathcal{H})$, $p_i \geq 0$, $\sum_i p_i = 1$, and $\sum_i p_i \rho^i = \rho$. Furthermore, let $|\Psi\rangle \in \mathcal{H} \otimes \mathcal{H}_D$ be a purification of ρ (i.e., $\mathrm{tr}_D\,|\Psi\rangle\langle\Psi| = \rho$). A projective measurement $\{P_i\}_{i\in I}$ on $\mathcal{H}_D$ steers the ensemble $\{(\rho^i, p_i)\}$ if

$$p_i = \langle\Psi|I \otimes P_i|\Psi\rangle \quad \text{and} \quad \rho^i = \frac{1}{p_i}\,\mathrm{tr}_R\big[(I \otimes P_i)|\Psi\rangle\langle\Psi|\big] \tag{B3}$$

for $p_i > 0$.

Assumption B.1. *Let $\mathcal{H}_A$ and $\mathcal{H}_C$ be finite-dimensional Hilbert spaces and write $\mathcal{H}_{AC} = \mathcal{H}_A \otimes \mathcal{H}_C$. Let $\{\Phi_y^{\rho_{AC}}\}_{\rho_{AC}\in\mathcal{D}(\mathcal{H}_{AC}),y}$ be a family of linear, trace preserving maps. For every density operator $\rho_{AC} \in \mathcal{D}(\mathcal{H}_{AC})$ describing Alice-Charlie's system, every finite-dimensional reference system $\mathcal{H}_D$ that may be held by a space-like separated party David, and every purification $|\Psi_{ACD}\rangle \in \mathcal{H}_{AC} \otimes \mathcal{H}_D$ of ρ_{AC}, the following holds.*

Let $\{P_i\}_{i\in I}$ and $\{Q_j\}_{j\in J}$ be any two projective measurements on $\mathcal{H}_D$ that steer ensembles $\{(\rho_{AC}^i, p_i)\}$ and $\{(\sigma_{AC}^j, q_j)\}$ of ρ_{AC} respectively. After the evolution, the states of the Alice-Charlie system must be identical:

$$\sum_{i\in I} p_i (\Phi_y^{\rho_{AC}^i})^\dagger(\rho_{AC}^i) = \sum_{j\in J} q_j (\Phi_y^{\sigma_{AC}^j})^\dagger(\sigma_{AC}^j). \tag{B4}$$

Specifically, if the two mixtures in (B4) were different, a measurement performed on Alice-Charlie's system would reveal which of the two measurements $\{P_i\}_{i\in I}$ and $\{Q_j\}_{j\in J}$ were chosen by David on the spacelike separated system, thereby violating the principle of no-superluminal-signaling.

Lemma B.1. *Let $\mathcal{H}_A$ and $\mathcal{H}_C$ be finite-dimensional Hilbert spaces and write $\mathcal{H}_{AC} = \mathcal{H}_A \otimes \mathcal{H}_C$. Let $\{\Phi_y^{\rho_{AC}}\}_{\rho_{AC}\in\mathcal{D}(\mathcal{H}_{AC}),y}$ be a family of linear, trace preserving maps. Under Assumption B.1, for every pair of density operators $\rho_{AC}^0, \rho_{AC}^1 \in \mathcal{D}(\mathcal{H}_{AC})$, every $p \in [0,1]$ and every input y, it holds that*

$$\left(\Phi_y^{p\rho_{AC}^0 + (1-p)\rho_{AC}^1}\right)^\dagger \left(p\rho_{AC}^0 + (1-p)\rho_{AC}^1\right) = p\,\left(\Phi_y^{\rho_{AC}^0}\right)^\dagger(\rho_{AC}^0) + (1-p)\,\left(\Phi_y^{\rho_{AC}^1}\right)^\dagger(\rho_{AC}^1). \tag{B5}$$

Proof. We are given arbitrary $\rho^0_{AC}, \rho^1_{AC} \in \mathcal{D}(\mathcal{H}_{AC})$ and $p \in [0,1]$. Let us define

$$\rho_{AC} := p\rho^0_{AC} + (1-p)\rho^1_{AC}. \tag{B6}$$

By the GHJW theorem [42, 43], for any such ρ^0_{AC}, ρ^1_{AC} and p, there exists a purification $|\Psi_{ACD}\rangle \in \mathcal{H}_{AC} \otimes \mathcal{H}_D$ of ρ_{AC} and a two-outcome projective measurement $\{P_0, P_1\}$ on $\mathcal{H}_D$ that steers the ensemble $\{(\rho^0_{AC}, p), (\rho^1_{AC}, (1-p)\}$, while the single ensemble $\{\rho_{AC}, I_{AC}\}$ is steered by the trivial measurement I_D. By the no-superluminal-signaling condition in (B4) of Assumption B.1 applied to these two measurements $\{P_0, P_1\}$ and $\{I_D\}$, we have that the states after the evolution must be identical, i.e., it holds that

$$\left(\Phi_y^{\rho_{AC}}\right)^\dagger(\rho_{AC}) = p\,\left(\Phi_y^{\rho^0_{AC}}\right)^\dagger(\rho^0_{AC}) + (1-p)\,\left(\Phi_y^{\rho^1_{AC}}\right)^\dagger(\rho^1_{AC}), \tag{B7}$$

which is equivalent to (B5) by (B6). □

The following theorem shows that under the no-superluminal-signaling Assumption B.1 and the locally identity-preserving constraints on the maps $\Phi_y^{\rho_{AC}}$, the only maps that survive are those for which the effective state transformation is trivial: $\left(\Phi_y^{\rho_{AC}}\right)^\dagger(\rho_{AC}) = \rho_{AC}$ for all ρ_{AC}, y. In other words, no nontrivial jamming is possible even in this state-dependent setting.

Theorem B.1. *Let $\mathcal{H}_A$ and $\mathcal{H}_C$ be finite-dimensional Hilbert spaces and write $\mathcal{H}_{AC} = \mathcal{H}_A \otimes \mathcal{H}_C$. Let $\{\Phi_y^{\rho_{AC}}\}_{\rho_{AC}\in\mathcal{D}(\mathcal{H}_{AC}),y}$ be a family of maps $\Phi_y^{\rho_{AC}} : \mathcal{L}(\mathcal{H}_{AC}) \to \mathcal{L}(\mathcal{H}_{AC})$ that satisfies, for every density operator ρ_{AC} and every input y,*

- *$\Phi_y^{\rho_{AC}}$ is linear,*
- *$\Phi_y^{\rho_{AC}}$ is trace-preserving $tr\left(\Phi_y^{\rho_{AC}}(X)\right) = tr(X)$ for every $X \in \mathcal{L}(\mathcal{H}_{AC})$,*
- *$\Phi_y^{\rho_{AC}}$ is locally identity-preserving, i.e., $\Phi_y^{\rho_{AC}}(X_A \otimes \mathbb{I}_C) = X_A \otimes \mathbb{I}_C$ and $\Phi_y^{\rho_{AC}}(\mathbb{I}_A \otimes X_C) = \mathbb{I}_A \otimes X_C$ for all observables $X_A \in \mathcal{L}(\mathcal{H}_A)$ and $X_C \in \mathcal{L}(\mathcal{H}_C)$,*
- *$\left(\Phi_y^{\rho_{AC}}\right)^\dagger(\rho_{AC}) \geq 0$.*

Then under the no-superluminal-signaling Assumption B.1, it holds that

$$\left(\Phi_y^{\rho_{AC}}\right)^\dagger(\rho_{AC}) = \rho_{AC}, \tag{B8}$$

for every $\rho_{AC} \in \mathcal{D}(\mathcal{H}_{AC})$ and every input y.

Proof. Fix an input y, since y is arbitrary it is sufficient to prove the claim for this fixed value. Define $F_y : \mathcal{D}(\mathcal{H}_{AC}) \to \mathcal{L}(\mathcal{H}_{AC})$ by

$$F_y(\rho_{AC}) := \left(\Phi_y^{\rho_{AC}}\right)^\dagger(\rho_{AC}). \tag{B9}$$

By assumption, $F_y(\rho_{AC}) \geq 0$. From Lemma B.1, for any ensemble $\rho_{AC} = \sum_i p_i \rho^i_{AC}$, the no-superluminal-signaling Assumption B.1 implies that

$$F_y(\rho_{AC}) = \sum_i p_i F_y(\rho^i_{AC}), \tag{B10}$$

i.e., F_y is affine on the convex set $\mathcal{D}(\mathcal{H}_{AC})$.

We now show that F_y preserves both reduced states on $\mathcal{H}_A$ and $\mathcal{H}_C$. To that end, observe that for every $X_A \in \mathcal{L}(\mathcal{H}_A)$, we have

$$\mathrm{tr}\left[F_y(\rho_{AC})(X_A \otimes \mathbb{I}_C)\right] = \mathrm{tr}\left[\left(\Phi_y^{\rho_{AC}}\right)^\dagger(\rho_{AC})(X_A \otimes \mathbb{I}_C)\right] = \mathrm{tr}\left[\rho_{AC}\Phi_y^{\rho_{AC}}(X_A \otimes \mathbb{I}_C)\right] = \mathrm{tr}\left[\rho_{AC}(X_A \otimes \mathbb{I}_C)\right]. \tag{B11}$$

In the above, the first equality is the definition of F_y, the second equality is by the definition of the adjoint, and the third equality is by the local identity preservation of $\Phi_y^{\rho_{AC}}$. We therefore obtain that

$$\begin{aligned} &\mathrm{tr}\left[F_y(\rho_{AC})(X_A \otimes \mathbb{I}_C)\right] = \mathrm{tr}\left[\rho_{AC}(X_A \otimes \mathbb{I}_C)\right] \\ \implies &\mathrm{tr}\left[\mathrm{tr}_C F_y(\rho_{AC})X_A - \mathrm{tr}_C\rho_{AC}X_A\right] = 0 \quad \forall X_A \in \mathcal{L}(\mathcal{H}_A) \\ \implies &\mathrm{tr}_C F_y(\rho_{AC}) = \mathrm{tr}_C\rho_{AC}. \end{aligned} \tag{B12}$$

By an analogous argument using $\Phi_y^{\rho_{AC}}(\mathbb{I}_A \otimes X_C) = \mathbb{I}_A \otimes X_C$ we obtain that

$$\mathrm{tr}_A F_y(\rho_{AC}) = \mathrm{tr}_A \rho_{AC}. \tag{B13}$$

We now proceed to show that F_y fixes every pure product state. To that end, let $\sigma_{AC} = |\psi\rangle\langle\psi|_A \otimes |\phi\rangle\langle\phi|_C$ be an arbitrary pure product state, and set $\tau_{AC} := F_y(\sigma_{AC})$. By assumption, we have that $\tau \geq 0$. From the properties (B12) and (B13) derived above we have that

$$\begin{aligned} \mathrm{tr}_C \tau_{AC} &= |\psi\rangle\langle\psi|_A, \\ \mathrm{tr}_A \tau_{AC} &= |\phi\rangle\langle\phi|_C. \end{aligned} \tag{B14}$$

Since $\tau_{AC} \geq 0$ and $\mathrm{tr}_C \tau_{AC} = |\psi\rangle\langle\psi|_A$ is pure, we obtain that $\tau_{AC} = |\psi\rangle\langle\psi|_A \otimes \sigma_C$ for some state σ_C. By the second marginal condition, we then obtain that $\sigma_C = \mathrm{tr}_A \tau_{AC} = |\phi\rangle\langle\phi|_C$ giving $\tau_{AC} = |\psi\rangle\langle\psi|_A \otimes |\phi\rangle\langle\phi|_C$. We therefore obtain that

$$F_y(\sigma_{AC}) = |\psi\rangle\langle\psi|_A \otimes |\phi\rangle\langle\phi|_C = \sigma_{AC}. \tag{B15}$$

In other words, F_y fixes every pure product state. It remains to show that this implies that F_y is the identity on all density operators.

Now since $\mathcal{H}_A$ and $\mathcal{H}_C$ are finite-dimensional, the real vector spaces $\mathrm{Herm}(\mathcal{H}_A)$ and $\mathrm{Herm}(\mathcal{H}_C)$ are also finite-dimensional and the pure-state projectors span these spaces (every Hermitian operator on $\mathcal{H}_A$ is a real linear combination of projectors $|\psi\rangle\langle\psi|_A$, and similarly for $\mathcal{H}_C$). It follows that the operators $|\psi\rangle\langle\psi|_A \otimes |\phi\rangle\langle\phi|_C$ span the real vector space $\mathrm{Herm}(\mathcal{H}_A \otimes \mathcal{H}_C)$. Now, since $\mathcal{D}(\mathcal{H}_{AC})$ has nonempty relative interior in the real affine space of trace-one Hermitian operators, F_y admits a unique real-linear extension $\tilde{F}_y$ to $\mathrm{Herm}(\mathcal{H}_{AC})$. Furthermore, since $\tilde{F}_y$ fixes every pure product projector, by linearity we have that $\tilde{F}_y = \mathbb{I}_{AC}$ on $\mathrm{Herm}(\mathcal{H}_{AC})$. Therefore, we obtain that

$$F_y(\rho_{AC}) = \left(\Phi_y^{\rho_{AC}}\right)^\dagger (\rho_{AC}) = \rho_{AC}, \tag{B16}$$

for every density operator ρ_{AC}. □

Therefore, we conclude that no nontrivial jamming extension of quantum correlations is possible - any attempt to introduce nontrivial jamming maps in either state-independent or state-dependent fashion leads to hidden superluminal signaling.

Appendix C: Jamming Attacks on Device-Independent Cryptographic Protocols

In this section, we investigate the security of Device-Independent (DI) protocols against adversaries constrained only by the principle of Relativistic Causality. As we have seen, the set $\mathcal{RC}$ is the correct set to consider when considering relativistic attacks on DI protocols rather than the hitherto considered no-signaling polytope $\mathcal{NS}$ [2, 8, 18].

1. Jamming attacks on Device-Indpendent Bit Commitment

In this section, we outline how jamming correlations affect the security of device-independent schemes for bit commitment. Specifically, we show how the two-prover bit commitment scheme of Fehr-Fillinger [15] which was proven secure solely on the impossibility of superluminal signaling (when the devices held by the dishonest provers correspond to an arbitrary no-signaling behavior), fails to be secure in configurations which allow jamming.

Recall that a bit commitment scheme lets a prover (committer) lock in a bit $b \in \{0, 1\}$ such that two properties are obeyed:

- Hiding: after the commit phase, a verifier has (almost) no information about b,
- Binding: the prover cannot later open the commitment to both 0 and 1.

It is well-known that in the classical and quantum single-prover scenarios it is impossible to achieve both properties unconditionally [44, 45]. Ben-Or, Goldwasser, Kilian and Wigderson [46] therefore introduced the multi-prover model, whereby the prover is split into two or more agents that are assumed unable to communicate with each other once the protocol begins. The hope then is that the no-communication assumption alone is sufficient to enforce binding.

Considering two-prover schemes, it is also known that every single-round two-prover commitment scheme that is almost perfectly hiding can be broken by a no-signaling attack. Further, the attack is essentially optimal: the dishonest provers can open the commitment to either bit with success probability equal to the honest opening probability (probability 1 when the scheme is perfectly sound.

On the other hand, as shown by Fehr-Fillinger, there exists a three-prover commitment scheme that is secure against every non-signaling attack [15]. Here, we show that for the same scheme is rendered fully insecure by jamming attacks in the specific spacetime configurations where the third prover P_3 can jam the first two provers P_1 and P_2 (i.e., where the intersection of the future light cones of the output events of P_1 and P_2 lies within the future light cone of the input event of P_3).

We briefly sketch the (simple) argument before presenting a formal proof in Thm. C.1. For a scheme to be considered binding against every no-signaling (NS) strategy, the combined probability of successfully opening to bit 0 and bit 1 must be strictly limited as $p_0^{NS} + p_1^{NS} \leq 1 + \eta$ for some $\eta < 1$ [15]. Now, RC constraints in the jamming configuration permit the joint distribution of messages from P_1, P_2 to depend on an auxiliary classical bit c given to P_3. This allows a strategy whereby both P_1 and P_2 sample and store opening strings for both possible bits using shared randomness, while P_3 uses c to sample and output the corresponding marginals for the chosen string. By executing this RC strategy, the provers can successfully open the commitment to either bit with a combined probability of $p_0^{RC} + p_1^{RC} \geq 2 - 2\epsilon - 2\delta$ where ϵ is the hiding parameter and δ is the soundness error. In an ideal commitment scheme with $\epsilon, \delta \to 0$, this success rate exceeds the binding bound $1 + \eta$ that holds against NS adversaries, showing that RC attacks successfully break the commitment scheme.

To establish this formally, let us define the three-prover bit commitment scheme Π of [15]. The verifier samples a challenge tuple $a = (a_1, a_2, a_3)$ according to a prior distribution $\mu(a)$ and sends a_i to prover P_i. Each prover P_i replies with a commit-phase message $x_i \in X_i$. To open a committed bit $b \in \{0, 1\}$, the provers subsequently return opening messages $y_i \in Y_i$. The verifier evaluates a predicate $V(a, x, y, b) \in \{0, 1\}$ and accepts if and only if $V = 1$. Let $p_b^{\text{hon}}(x, y|a)$ denote the joint probability distribution of all messages when the provers execute the honest strategy to commit to bit b and later open to b. The scheme satisfies three properties:

- $\varepsilon-Hiding$: The commit-phase messages are statistically indistinguishable to the verifier: $d(p_0^{\text{hon}}(x), p_1^{\text{hon}}(x)) \leq \varepsilon$, where $d(P, Q) = \frac{1}{2}\sum_x |P(x) - Q(x)|$ is the total variation distance.
- $\delta - Sound$: The honest strategies succeed with high probability: $\mathbb{E}_{a\sim\mu}\left[p_b^{\text{hon}}(V(a, x, y, b) = 1)\right] \geq 1 - \delta \quad \forall b \in \{0, 1\}$.
- Binding against Non-Signaling: Against any non-signaling (NS) adversary, the sum of success probabilities for opening to 0 and 1 is bounded as $p_0^{\text{NS}} + p_1^{\text{NS}} \leq 1 + \eta$ for some $\eta < 1$.

Theorem C.1. *Suppose that the three provers P_1, P_2 and P_3 executing protocol Π above are embedded in a spacetime configuration $\mathcal{M}$ where P_3 is permitted to jam the pair $\{P_1, P_2\}$. There exists a behavior in the Relativistically Causal (RC) set of $\mathcal{M}$ such that the provers can successfully open to both $b = 0$ and $b = 1$ with a combined probability:*

$$p_0^{\text{RC}} + p_1^{\text{RC}} \geq 2 - 2\varepsilon - 2\delta, \tag{C1}$$

where ϵ is the hiding parameter and δ is the soundness error of the protocol Π. In the ideal scheme with $\epsilon, \delta \to 0$, this breaks the NS binding bound without violating relativistic causality.

Proof. We first recall the following lemma from [15].

Lemma C.1 ([15])**.** *Let $p_0(u, v)$ and $p_1(u, v)$ be two probability distributions defined on a finite product space $\mathcal{U} \times \mathcal{V}$ such that the total variation distance between their marginals on $\mathcal{U}$ is bounded as $d(p_0(u), p_1(u)) \leq \varepsilon$. Then there exists a joint distribution $q(u_0, u_1, v_0, v_1)$ such that:*

$$\begin{aligned} \sum_{u_1, v_1} q(u_0, u_1, v_0, v_1) &= p_0(u_0, v_0) \\ \sum_{u_0, v_0} q(u_0, u_1, v_0, v_1) &= p_1(u_1, v_1) \\ q(u_0 \neq u_1) &\leq \varepsilon. \end{aligned} \tag{C2}$$

Proof. By definition of statistical distance, there is a coupling $r(u_0, u_1)$ of the marginals $p_0(u)$ and $p_1(u)$ (with $\sum_{u_1} r(u_0, u_1) = p_0(u_0)$ and $\sum_{u_0} r(u_0, u_1) = p_1(u_1)$) satisfying $P_r(u_0 \neq u_1) = \sum_{u_0, u_1 : u_0 \neq u_1} r(u_0, u_1) = \varepsilon$. We set:

$$q(u_0, u_1, v_0, v_1) := r(u_0, u_1)\, p_0(v_0|u_0)\, p_1(v_1|u_1). \tag{C3}$$

The three properties hold by direct inspection of the marginals. □

Now, for each fixed challenge a, apply the Lemma C.1 to the two honest distributions $p_0^{\text{hon}}(\cdot|a)$ and $p_1^{\text{hon}}(\cdot|a)$, taking $u = x$ and $v = y$. This produces a family of distributions $q(x^{(0)}, x^{(1)}, y^{(0)}, y^{(1)}|a)$ such that, for every a [15]:

- The marginal on $(x^{(0)}, y^{(0)})$ equals $p_0^{\text{hon}}(\cdot|a)$,
- The marginal on $(x^{(1)}, y^{(1)})$ equals $p_1^{\text{hon}}(\cdot|a)$,
- $q(x^{(0)} \neq x^{(1)}|a)$ satisfies $\mathbb{E}_{a\sim\mu}[q(x^{(0)} \neq x^{(1)})] \leq \varepsilon$.

Define a new distribution $r(x, y^{(0)}, y^{(1)}|a)$ on the set of triples $(x, y^{(0)}, y^{(1)})$ by:

$$r(x, y^{(0)}, y^{(1)}|a) := q(x^{(0)} = x, x^{(1)} = x, y^{(0)}, y^{(1)}|a) + \frac{1}{2}q(x^{(0)} \neq x^{(1)}, y^{(0)}, y^{(1)}|a). \tag{C4}$$

We now compute the acceptance probabilities under r. On the event $x^{(0)} = x^{(1)} = x$, the pair $(x, y^{(0)})$ is distributed exactly as p_0^{hon} and the pair $(x, y^{(1)})$ is distributed exactly as p_1^{hon}. On the complementary event (which has probability at most ε), the acceptance probability is non-negative. Therefore, we obtain

$$\begin{aligned} \mathbb{E}_{a\sim\mu}\big[r(V(a, x, y^{(0)}, 0) = 1)\big] &\geq (1-\varepsilon)(1-\delta) + \varepsilon \cdot 0 \\ &= 1 - \delta - \varepsilon + \varepsilon\delta \geq 1 - \delta - \varepsilon. \end{aligned} \tag{C5}$$

Similarly, for $b = 1$:

$$\mathbb{E}_{a\sim\mu}\big[r(V(a, x, y^{(1)}, 1) = 1)\big] \geq 1 - \delta - \varepsilon. \tag{C6}$$

We now embed this into an RC behavior via the following construction:

- Using shared randomness that realizes the distribution $r(\cdot|a)$, the provers P_1 and P_2 sample the tuple $(x_1, x_2, y_1^{(0)}, y_1^{(1)}, y_2^{(0)}, y_2^{(1)})$ and output the commit-phase messages x_1, x_2.
- Prover P_3 receives their challenge a_3 and the auxiliary bit c. P_3 outputs x_3 together with the single opening string $y_3^{(c)}$.
- In the opening phase, P_1 and P_2 output the strings $y_1^{(c)}$ and $y_2^{(c)}$.

The construction ensures that the resulting joint distribution of all messages, conditioned on c, is exactly the corresponding marginal of r. Furthermore, the single-party marginals are independent of c by construction. Therefore, the behavior lies within the set $\mathcal{RC}$ of the jamming configuration.

By construction, it holds that if the auxiliary bit $c = 0$, the verifier's predicate for bit 0 evaluates to 1 with probability at least $1 - \delta - \varepsilon$. Analogously, if the auxiliary bit $c = 1$, the predicate for bit 1 evaluates to 1 with probability at least $1 - \delta - \varepsilon$. Therefore we obtain,

$$p_0^{\text{RC}} + p_1^{\text{RC}} \geq 2 - 2\varepsilon - 2\delta. \tag{C7}$$

We thus obtain a behavior that surpasses the NS binding bound without violating relativistic causality. □

2. Jamming attacks on Device-Independent Secret Sharing

Device-Independent (DI) Secret Sharing aims to distribute a secret among multiple receivers so that the secret can be reconstructed only when a designated subset of receivers collaborate, while remaining hidden from any proper subset and from external eavesdroppers. The protocol of [14] involves three distant parties - Charlie (the sender) and Alice and Bob (the receivers), who all share a device that takes binary inputs x, y, z and produces binary outputs a, b, c respectively. For a fixed input triple $(x^*, y^*, z^*) = (0, 0, 0)$, the outputs are required to satisfy the relation $c = a \oplus b$ (where $a \oplus b$ denotes the XOR of the two bits a, b). So that the sender Charlie's secret bit c is the parity of the two receivers' bits a and b which appear random and hold no information individually. The protocol is required to be secure against an external eavesdropper but also against an untrusted receiver Alice, who may produce an additional output e in an attempt to guess the secret, i.e., the dishonest Alice aims to maximize her guessing probability of the secret $P_{\text{guess}}(e = c|x^*, y^*, z^*)$ without collaborating with Bob. Security against such attacks is certified by the parties

testing for sufficiently strong genuine multipartite nonlocality in the form of the violation of Svetlichny inequality $S \leq 4$ [47] by the observed behavior $\{P(a,b,c|x,y,z)\}$. Formally, the Svetlichny inequality is given as

$$\begin{aligned} S \;=\; \sum_{a,b,c=0}^{1} (-1)^{a+b+c} \Big[& P(a,b,c|0,0,0) + P(a,b,c|0,0,1) + P(a,b,c|0,1,0) - P(a,b,c|0,1,1) \\ & + P(a,b,c|1,0,0) - P(a,b,c|1,0,1) - P(a,b,c|1,1,0) - P(a,b,c|1,1,1) \Big] \leq 4. \end{aligned} \tag{C8}$$

Under the full set of no-signaling constraints (i.e., that the observed behavior lies within $\mathcal{NS}$) the maximal violation of the Svetlichny inequality enforces that $P_{\text{guess}}(e = c|x^*, y^*, z^*) = 1/2$.

Theorem C.2. *Let* Π *denote the tripartite DI secret sharing protocol of [14] explained above. Suppose that the protocol is executed in a jamming configuration where Bob's measurement event lies in the causal past of the intersection of the future light cones of Alice and Charlie, so that Bob is permitted to jam the Alice-Charlie marginal. Then there exists a behavior* $\{P(a,b,c,e|x,y,z)\}$ *in the relativistically causal set* $\mathcal{RC}$ *such that*

- *the observable behavior with marginals* $P(a,b,c|x,y,z) = \sum_e P(a,b,c,e|x,y,z)$ *attains the algebraic maximum of the Svetlichny expression* (C8) $S = 8$,
- *a dishonest Alice is able to perfectly guess the secret with an auxiliary output* e, *i.e., Alice achieves* $P_{guess}(e = c|x^*, y^*, z^*) = 1$ *without collaborating with Bob.*

Proof. The explicit behavior $\{P(a,b,c,e|x,y,z)\} \in \mathcal{RC}$ achieving the two conditions stated in the theorem is computed by optimizing the attack strategies over the convex polytope $\mathcal{RC}$ and is given by

$$P(a,b,c,e|x,y,z) = \frac{1}{4}\, \delta_{a\oplus b\oplus c, xy\oplus xz\oplus yz} \cdot \delta_{e,a\oplus b}, \tag{C9}$$

where $\delta_{i,j}$ denotes the Kronecker delta ($\delta_{i,j} = 11$ if $i = j$ and 0 otherwise), and $\oplus$ denotes addition modulo 2.

Firstly, we sum over the auxiliary output e to obtain the observed behavior as

$$P(a,b,c|x,y,z) = \sum_e P(a,b,c,e|x,y,z) = \frac{1}{4}\, \delta_{a\oplus b\oplus c, xy\oplus xz\oplus yz}. \tag{C10}$$

We see that for $(x,y,z) \in \{(0,0,0),(0,0,1),(0,1,0),(1,0,0)\}$ we have $xy \oplus xz \oplus yz = 0$ giving $P(a,b,c|x,y,z) = \frac{1}{4}$ for $a \oplus b \oplus c = 0$ meaning $(-1)^{a+b+c} = 1$. Similarly for $(x,y,z) \in \{(0,1,1),(1,0,1),(1,1,0),(1,1,1)\}$ we have $xy \oplus xz \oplus yz = 1$ giving $P(a,b,c|x,y,z) = \frac{1}{4}$ for $a \oplus b \oplus c = 1$ meaning $(-1)^{a+b+c} = -1$. Substituting in (C8) gives $S = 8$.

Secondly, we verify that the behavior $\{P(a,b,c,e|x,y,z)\}$ satisfies all the relativistic causality constraints for the jamming configuration. Recall that [4] here all marginal behaviors of any subset of parties is required to be independent of the inputs of the complementary set, except the joint behavior of Alice-Charlie which may depend upon the input of Bob. We verify the relativistic causality constraints by explicitly computing the marginal probabilities as

$$\begin{aligned} P(a,e|x,y,z) &= \sum_{b,c} P(a,b,c,e|x,y,z) = \frac{1}{4} && \text{independent of } y,z, \\ P(b|x,y,z) &= \sum_{a,c,e} P(a,b,c,e|x,y,z) = \frac{1}{2} && \text{independent of } x,z, \\ P(c|x,y,z) &= \sum_{a,b,e} P(a,b,c,e|x,y,z) = \frac{1}{2} && \text{independent of } x,y, \\ P(a,e,b|x,y,z) &= \sum_{c} P(a,b,c,e|x,y,z) = \frac{1}{4}\,\delta_{e,a\oplus b} && \text{independent of } z, \\ P(b,c|x,y,z) &= \sum_{a,e} P(a,b,c,e|x,y,z) = \frac{1}{4} && \text{independent of } x. \end{aligned} \tag{C11}$$

On the other hand, we also compute that

$$P(a,c,e|x,y,z) \;=\; \sum_{b} P(a,b,c,e|x,y,z) = \frac{1}{4}\,\delta_{e\oplus c,\, xy\oplus xz\oplus yz} \quad \text{dependent on } y, \tag{C12}$$

showing that the behavior $\{P(a,b,c,e|x,y,z)\} \in \mathcal{RC}$ but $\{P(a,b,c,e|x,y,z)\} \notin \mathcal{NS}$.

Finally, we have by construction that $e = a \oplus b$ in the behavior. For the input $(x^*, y^*, z^*) = (0,0,0)$ one has $x^*y^* \oplus x^*z^* \oplus y^*z^* = 0$ giving $a \oplus b \oplus c = 0$. Under this condition $e = a \oplus b$ ensures $e = c$ for this input, and we conclude that $P_{\text{guess}}(e = c|x^*, y^*, z^*) = 1$.

□