Continuous-variable supraquantum nonlocality

Andreas Ketterer,[1,2,*] Adrien Laversanne-Finot,[2,†] and Leandro Aolita[3,4,‡]
[1]*Naturwissenschaftlich-Technische Fakultät, Universität Siegen, Walter-Flex-Str. 3, 57068 Siegen, Germany*
[2]*Laboratoire Matériaux et Phénomènes Quantiques, Sorbonne Paris Cité, Université Paris Diderot, CNRS UMR 7162, 75013 Paris, France*
[3]*Instituto de Física, Universidade Federal do Rio de Janeiro, Caixa Postal 68528, 21941-972 Rio de Janeiro, RJ, Brazil*
[4]*ICTP South American Institute for Fundamental Research Instituto de Física Teórica, UNESP-Universidade Estadual Paulista R. Dr. Bento T. Ferraz 271, Bl. II, São Paulo 01140-070, SP, Brazil*

(Received 12 September 2017; revised manuscript received 7 December 2017; published 31 January 2018)

Supraquantum nonlocality refers to correlations that are more nonlocal than allowed by quantum theory but still physically conceivable in postquantum theories, in the sense of respecting the basic no-faster-than-light communication principle. While supraquantum correlations are relatively well understood for finite-dimensional systems, little is known in the infinite-dimensional case. Here, we study supraquantum nonlocality for bipartite systems with two measurement settings and infinitely many outcomes per subsystem. We develop a formalism for generic no-signaling black-box measurement devices with continuous outputs in terms of probability measures, instead of probability distributions, which involves a few technical subtleties. We show the existence of a class of supraquantum Gaussian correlations, which violate the Tsirelson bound of an adequate continuous-variable Bell inequality. We then introduce the continuous-variable version of the celebrated Popescu–Rohrlich (PR) boxes, as a limiting case of the above-mentioned Gaussian ones. Finally, we characterize the geometry of the set of continuous-variable no-signaling correlations. Namely, we show that that the convex hull of the continuous-variable PR boxes is dense in the no-signaling set. We also show that these boxes are extreme in the set of no-signaling behaviors and provide evidence suggesting that they are indeed the only extreme points of the no-signaling set. Our results lay the grounds for studying generalized-probability theories in continuous-variable systems.

DOI: 10.1103/PhysRevA.97.012133

I. INTRODUCTION

Bell nonlocality refers to correlations incompatible with local hidden-variable theories [1], which explain correlations between spacelike separated measurement outcomes as due exclusively to past common causes. Since the pioneering works of Bell [1] and of Clauser, Horn, Shimony, and Holt [2], it is known that quantum mechanics admits Bell nonlocality, i.e., that local measurements on quantum entangled states produce Bell nonlocal correlations. However, nonlocality is not a phenomenon exclusive of quantum theory. Hypothetical supraquantum theories satisfying the basic *no-signaling principle* of no-faster-than-light communication, in consistency with special relativity, can produce Bell correlations that are even more nonlocal than those compatible with quantum theory. This is generally referred to as *supraquantum Bell nonlocality*. The first known example thereof was the so-called Popescu–Rohrlich (PR) boxes [3]. These are hypothetic black-box measurement devices that can violate the Clauser–Horn–Shimony–Holt inequality up to its algebraic maximum of four, which is above the maximum value of $2\sqrt{2}$ attained by quantum correlations, known as Tsirelson's bound [4].

Importantly, the aim of studying supraquantum nonlocality is by no means to question the validity of quantum mechanics but, rather on the contrary, actually to gain a better understanding of quantum nonlocality itself. For instance, even though unphysical, PR boxes make excellent *units of Bell nonlocality*, serving, in fact, as references to quantify the nonlocal weight of quantum correlations [5,6]. Furthermore, understanding why quantum mechanics is not as nonlocal as allowed by the no-signaling principle gives us valuable insights with foundational implications on the very axiomatic structure of quantum theory. For instance, a seminal result in this direction was the realization that the physical existence of PR boxes would make communication complexity problems trivial [7–9], which is a highly implausible possibility. Hence, if one accepts that communication complexity is not trivial as a postulate, the nonexistence of PR boxes is implied. In fact, in a similar spirit, a large effort has been devoted to proposing physically reasonable postulates from which Tsirelson's bound can be derived from first principles (see, e.g., Refs. [10–14]).

PR boxes have been generalized to arbitrary finite numbers of measurement outcomes [15] and to multipartite systems as well [16]. What is more, in the multipartite scenario, nontrivial tight Bell inequalities are known without a quantum violation, i.e., for which the quantum maximum coincides with the local one and is below the no-signaling one [17]. In addition, supraquantum nonlocality has been explored even in the bipartite scenario where only one part makes measurements [18]. From a broader perspective, Bell nonlocality in generalized probabilistic theories has been extensively studied in the finite-dimensional case (see Ref. [19] and references therein). Nevertheless, in striking contrast, essentially nothing is known

*andreas.ketterer@uni-siegen.de
†adrien.laversanne-finot@univ-paris-diderot.fr
‡aolita@if.ufrj.br

about supraquantum nonlocality in continuous-variable (CV) systems. On the one hand, this is surprising in view of the huge amount of work on CV quantum nonlocality (see, e.g., Refs. [20–30]) and the importance of CV systems for quantum information processing [31–33]. On the other hand, this is at the same time understandable because, for CV systems, the set of local correlations (as well as that of no-signaling ones) is a generic convex set, instead of a (computationally much tamer) convex polytope as in finite-dimensional systems [34–36].

In this article we explore CV supraquantum nonlocality. To begin with, we develop a formalism to deal with generic no-signaling black-box measurement devices with discrete measurement settings (inputs) and CV measurement outcomes (outputs). The correlations produced by such devices are described by probability measures instead of probability distributions. We then show the existence of a class of supraquantum Gaussian PR boxes, for bipartite systems with dichotomic inputs and real, continuous outputs. This is done by showing that these behaviors violate the Cavalcanti–Foster–Reid–Drummond (CFRD) inequality [25], which admits no quantum violation in the bipartite case [30]. Next, we introduce a limiting case of the supraquantum Gaussian behavior, a hierarchy of CV PR boxes, whose ground level consists of local, deterministic points and the upper levels of nonlocal, nondeterministic ones. The CV PR boxes obtained are very similar in structure to the finite-dimensional ones. To end up with, we characterize the set of CV no-signaling behaviors and show that all CV PR boxes are extreme points of the CV no-signaling set, and that their convex hull (i.e., the set of all finite convex sums) is dense therein. In particular, we discuss whether the CV PR boxes are the only extreme no-signaling behaviors and, along with some evidence, conjecture that this is indeed the case.

The paper is structured as follows: In Sec. II, we set up the mathematical framework for CV no-signaling behavior based on probability measures. In Sec. III, we introduce the supraquantum Bell nonlocal Gaussian behavior and the CV PR boxes. Section IV is devoted to the geometrical characterization of the set of CV no-signaling set. Finally, we conclude in Sec. V with some final remarks and perspectives of our work.

II. PRELIMINARIES: MATHEMATICAL REPRESENTATION OF CONTINUOUS VARIABLE BELL CORRELATIONS

We consider a bipartite Bell experiment where two spacelike separated observers, conventionally referred to as Alice (A) and Bob (B), make measurements. We work in the generic device-independent scenario where the measurement apparatuses are treated as unknown black-box measurement devices [see Fig. 1(a)]. Alice's (Bob's) device has a dichotomic input x (y) $\in \{0,1\}$ and a continuous output a (b) $\in \mathbb{R}$. That is, we are considering infinite resolution: we want to investigate the ideal situation where the outputs can take any arbitrary real value. The statistics produced by such devices is most conveniently described in terms of probability measures, which we briefly recap in what follows. We consider probability spaces defined by a triple $\{\Omega,\mathcal{B}(\Omega),\mu\}$, where Ω denotes a sample space, $\mathcal{B}(\Omega)$ is the Borel σ algebra of events on Ω (i.e., the smallest σ algebra that contains all open subsets of Ω) and $\mu : \mathcal{B}(\Omega) \to [0,1]$ is

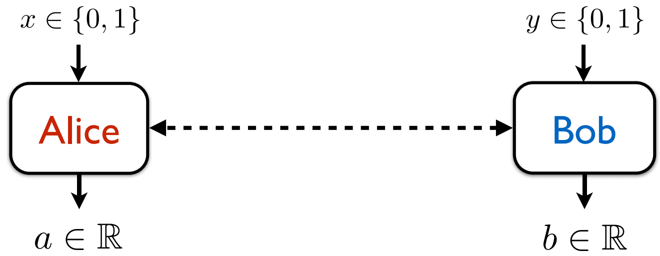

FIG. 1. Schematic representation of a bipartite Bell experiment with continuous measurement outcomes in the so-called device-independent scenario of black-box measurement instruments. Two spacelike separated observers, Alice (A) and (Bob), perform local measurements on their subsystems with dichotomic measurements choices (inputs) x and y, respectively, and obtain continuous-variable measurement outcomes (outputs) a and b.

a Borel probability measure. In our case, the sample space is given by a product space $\Omega = \Omega_{\rm A} \times \Omega_{\rm B}$, with $\Omega_{\rm A} = \Omega_{\rm B} = \mathbb{R}$, where the first and second factors, $\Omega_{\rm A}$ and $\Omega_{\rm B}$, correspond to the outputs of A and B, respectively. The probability measure μ is required to be normalized, $\mu(\mathbb{R}\times\mathbb{R}) = 1$, and to satisfy the additivity property $\mu(\bigcup_{i=1} E_i) = \sum_{i=1}\mu(E_i)$ for every countable sequence $\{E_i\}_i$ of disjoint events $E_i \in \mathcal{B}(\mathbb{R}\times\mathbb{R})$, where $\cup$ stands for the set union. The probability of an event $E \in \mathcal{B}(\mathbb{R}\times\mathbb{R})$ is then given by $P(E) := \mu(E)$. We denote the set of all probability measures on $\mathcal{B}(\mathbb{R}\times\mathbb{R})$ as $\mathcal{M}_{\mathbb{R}\times\mathbb{R}}$.

The connection between a probability measure μ and a probability density p (with respect to the Lebesgue measure) can be made explicit in the integral representation

$$\mu(A\times B) := \int_{A\times B} d\mu(a',b') = \int_A\int_B p(a',b')da'db', \quad (1)$$

where $A\times B \in \mathcal{B}(\mathbb{R}\times\mathbb{R})$, $A,B \in \mathcal{B}(\mathbb{R})$, $p(a',b')$ denotes the corresponding probability density to μ, and $d\mu(a,b)$ and $da'db'$ refer to integrations with respect to μ and the Lebesgue measure on $\mathbb{R}\times\mathbb{R}$, respectively. Note that not every probability measure can be expressed in terms of a probability density as in Eq. (1). The question of the existence of a probability density is answered by the Radon–Nikodym (RN) theorem, whose statement is briefly reviewed in Appendix A. While most assumptions of the RN theorem are fulfilled by any probability measure on $\mathbb{R}\times\mathbb{R}$, for us the crucial prerequisite is that μ has to be absolutely continuous with respect to the Lebesgue measure. However, as we see later on, absolute continuity cannot be guaranteed for all types of probability measures which will become important when dealing with so-called boxes describing idealized unphysical outcome scenarios. Hence, all in all, it is both more general and more convenient to work with measures, because one needs not worry about the existence of a density.

We thus arrive at the following definition:

Definition 1. (CV Bell behavior) A behavior is a joint conditional probability measure represented by a 2×2 matrix $\boldsymbol{\mu} = \{\mu_{x,y}\}_{x,y\in\{0,1\}}$ with arbitrary probability measures $[\boldsymbol{\mu}]_{x,y} := \mu_{x,y} \in \mathcal{M}_{\mathbb{R}\times\mathbb{R}}$ as entries. The set of all behaviors is denoted as $\mathcal{M}^4_{\mathbb{R}\times\mathbb{R}}$.

Note that, for finite-dimensional systems, the sample space has a finite number of events, so that joint conditional probability measures reduce to the more usual notion of joint conditional probability distributions [19]. Also as in the discrete case, since the observers are spacelike separated, $\boldsymbol{\mu}$ must fulfill the no-signaling principle, given, in this language, by the constraints

$$\mu_{x,y}(A \times \mathbb{R}) = \mu_{x,\overline{y}}(A \times \mathbb{R}) \; \forall \; x \in \{0,1\}, \tag{2a}$$

$$\mu_{x,y}(\mathbb{R} \times B) = \mu_{\overline{x},y}(\mathbb{R} \times B) \; \forall \; y \in \{0,1\}, \tag{2b}$$

for all $A,B \in \mathcal{B}(\mathbb{R})$, where $\overline{y} = y \oplus 1$ and $\overline{x} = x \oplus 1$, with $\oplus$ being the sum modulo two.

Conditions (2a) and (2b) imply respectively that Alice's and Bob's marginal measures $\mu_x(A) := \mu_{x,y}(A \times \mathbb{R})$ and $\mu_y(B) := \mu_{x,y}(\mathbb{R} \times B)$ are independent of each others' input, which prevents signaling. We call any $\boldsymbol{\mu}$ satisfying these conditions a *no-signaling behavior*, and denote the set of all no-signaling behaviors by $\mathcal{M}_{\text{NS}} \subset \mathcal{M}^4_{\mathbb{R}\times\mathbb{R}}$.

Quantum correlations, in turn, are those described by the behaviors that can be expressed as

$$\mu_{x,y}(A \times B) = \text{Tr}\,[M_x(A) \otimes M_y(B)\varrho_{AB}] \tag{3}$$

for all $A,B \in \mathcal{B}(\mathbb{R})$, where ρ_{AB} is an arbitrary bipartite quantum state on a Hilbert space $\mathcal{H} := \mathcal{H}_\text{A} \otimes \mathcal{H}_\text{B}$, with $\mathcal{H}_\text{A}$ and $\mathcal{H}_\text{B}$ being the local Hilbert spaces of Alice's and Bob's systems, respectively, and where M_x and M_y are, for all x $(y) \in \{0,1\}$, semispectral measures, also known as positive-operator valued measures (POVMs) [37]. The latter means that $M_x, M_y : \mathcal{B}(\mathbb{R}) \to \mathcal{L}_{\geqslant 0}(\mathcal{H})$ are maps such that, for all A $(B) \in \mathcal{B}(\mathbb{R}), M_x(A) \in \mathcal{L}_{\geqslant 0}(\mathcal{H}_\text{A})$ $[M_y(B) \in \mathcal{L}_{\geqslant 0}(\mathcal{H}_\text{B})]$, with $\mathcal{L}_{\geqslant 0}(\mathcal{H}_\text{A})$ $[\mathcal{L}_{\geqslant 0}(\mathcal{H}_\text{B})]$ being the space of positive semidefinite operators on $\mathcal{H}_\text{A}$ $(\mathcal{H}_\text{B})$; and that $M_x(\mathbb{R}) = \mathbb{1}_\text{A}$ $[M_y(\mathbb{R}) = \mathbb{1}_\text{B}]$, with $\mathbb{1}_\text{A}$ $(\mathbb{1}_\text{B})$ being the identity operator on $\mathcal{H}_\text{A}$ $(\mathcal{H}_\text{B})$. We call any $\boldsymbol{\mu}$ satisfying Eq. (3) a *quantum behavior* and denote the set of all quantum behaviors by $\mathcal{M}_\text{Q}$. For generic Bell scenarios, the relationship $\mathcal{M}_\text{Q} \subseteq \mathcal{M}_\text{NS}$ holds. For the scenario under consideration here, we show below that $\mathcal{M}_\text{Q} \subset \mathcal{M}_\text{NS}$. We call any $\boldsymbol{\mu} \in \mathcal{M}_\text{NS} \setminus \mathcal{M}_\text{Q}$ a *supraquantum behavior*.

The last important class for our purposes is the one of classical correlations, described by the behaviors produced by local hidden-variable models:

$$\mu_{x,y} = \int_\Lambda \delta_{a(x,\lambda),b(y,\lambda)} d\eta(\lambda), \tag{4}$$

where λ is the hidden variable, taking values in a parameter space Λ according to a probability measure $\eta : \mathcal{B}(\Lambda) \to \mathbb{R}_{\geqslant 0}$, and $\delta_{a(x,\lambda),b(y,\lambda)}$ is the CV version of the λth local deterministic response function. More precisely, $\delta_{a,b}$ denotes the Dirac measure at the point $(a,b) \in \mathbb{R}^2$, i.e., the deterministic measure such that

$$\delta_{a,b}(A \times B) := \begin{cases} 1 & \text{if } a \in A \text{ and } b \in B \\ 0 & \text{otherwise,} \end{cases} \tag{5}$$

for all $A,B \in \mathcal{B}(\mathbb{R})$. In turn, for each $\lambda \in \Lambda$, $a(x,\lambda)$ and $b(y,\lambda)$ are respectively deterministic functions of x and y, in a similar spirit to the local deterministic response functions in finite-dimensional scenarios [19]. Since the outputs are locally generated from each input and the pre-established classical correlations encoded in λ, one typically calls any $\boldsymbol{\mu}$ given by Eq. (4) a *local behavior*. We denote the set of all local behaviors by $\mathcal{M}_\text{L} \subseteq \mathcal{M}_\text{Q}$. In turn, any $\boldsymbol{\mu} \in \mathcal{M}_\text{NS} \setminus \mathcal{M}_\text{L}$ is a *nonlocal behavior*.

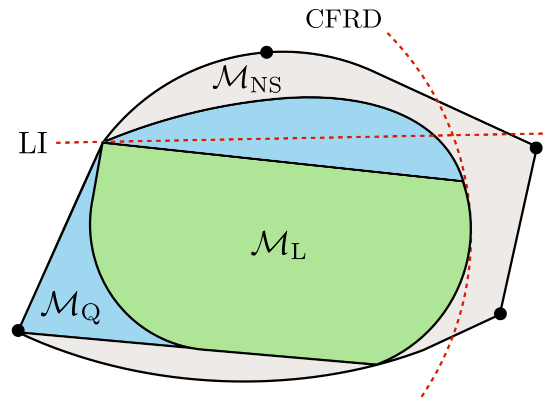

FIG. 2. Pictorial (not rigorous) geometrical representation of the (possible) inner structure of the set $\mathcal{M}_\text{NS}$ of CV no-signaling behaviors in the Bell scenario of Fig. 1. $\mathcal{M}_\text{NS}$ contains the set $\mathcal{M}_\mathcal{Q}$ of quantum behaviors, which contains, in turn, the set $\mathcal{M}_\mathcal{L}$ of local behaviors. All three sets are generic convex sets with infinitely many extreme points, delimited by facets as well as curved hypersurfaces. This is in contrast with the finite-dimensional case, where both $\mathcal{M}_\text{NS}$ and $\mathcal{M}_\mathcal{L}$ are convex polytopes, delimited exclusively by facets that can be characterized by a finite number of linear Bell inequalities. In the plot, an example of a linear Bell inequality is represented as the straight line LI. Such linear inequality can, e.g., correspond to a Bell inequality for finite-dimensional systems, which can be violated by CV quantum correlations using so-called binning procedures [20–24,27,28] (see also references in Ref. [19]). Besides this, a hypothetical quantum extreme point is shown in the figure (light-blue corner). While such points are in principle possible, no explicit example thereof is known. In this paper we consider a nonlinear Bell inequality, the CFRD inequality [25], represented as a curve in the plot. This inequality applies in the genuinely CV scenario of our interest and has, additionally, the appealing feature of admitting violations only by supraquantum behaviors (see Sec. III). Finally, four exemplary CV PR boxes are represented as extreme points of $\mathcal{M}_\text{NS}$ (black dots).

Finally, we emphasize that, in contrast with the finite-dimensional case, $\mathcal{M}_\text{L}$ does not define a polytope (i.e., a convex set with finitely many extreme points); see Fig. 2. This is due to the fact that Dirac measures are extreme in $\mathcal{M}_\text{NS}$ and $\mathcal{M}_\text{L}$ is generated by a continuously infinite number of them. It follows, then, that $\mathcal{M}_\mathcal{L}$ cannot be characterized by a finite set of linear Bell inequalities [34–36]. In the next section, we use a nonlinear Bell inequality to identify not only nonlocal behaviors but supraquantum ones.

III. CONTINUOUS-VARIABLE SUPRAQUANTUM NONLOCALITY

In Ref. [25], Cavalcanti, Foster, Reid, and Drummond derived the nonlinear Bell inequality

$$[\langle A_0 B_0\rangle - \langle A_1 B_1\rangle]^2 + [\langle A_0 B_1\rangle + \langle A_1 B_0\rangle]^2 \leqslant \langle A_0^2 B_0^2\rangle + \langle A_0^2 B_1^2\rangle + \langle A_1^2 B_0^2\rangle + \langle A_1^2 B_1^2\rangle, \tag{6}$$

where A_0 and A_1 (B_0 and B_1) are the real, continuous outputs of Alice's (Bob's) box for the inputs 0 and 1, respectively. By using the integral representation of Eq. (1), the expectation values of such observables appearing in the inequality can be recast as cross-moments of the behavior elements $\mu_{x,y}$:

$$\left\langle A_x^{n_a} B_y^{n_b} \right\rangle = \int_{\mathbb{R}^2} a^{n_a} b^{n_b} d\mu_{x,y}(a,b). \tag{7}$$

Equation (6) can be generalized to a higher number of parties [25] as well as observables per party [26]. We refer to the bipartite dichotomic-input version of inequality, given by Eqs. (6) and (7), as the *CFRD inequality*. The inequality has a number of interesting properties [25,26]. Specially relevant for our purposes is the fact that it cannot be violated by any quantum behavior. This was first shown in Ref. [29] for the restricted case of measurements of (quantum) phase-space quadrature operators, and then extended to the general case of arbitrary quantum measurements in Ref. [30]. Hence, the CFRD constitutes a nontrivial Bell inequality with no quantum violation. Any no-signaling behavior that violates it is thus automatically certified as supraquantum, as we do next.

The first case that we study is a subclass of behaviors that we term *Gaussian PR boxes*. To this end, we first introduce two real vectors, $\boldsymbol{a} := (a_1,\ldots,a_k)$ and $\boldsymbol{b} := (b_1,\ldots,b_k)$, with different components, i.e., such that $a_1 \neq a_2 \neq \cdots a_k$ and $b_1 \neq b_2 \neq \cdots b_k$, and one positive-real vector $\boldsymbol{\sigma} := (\sigma_1,\ldots,\sigma_k)$, all of length $k \in \mathbb{N}$. The vectors $\boldsymbol{a}$ and $\boldsymbol{b}$ determine k points (a_j,b_j) where Gaussian-measure components are centered; while the vector $\boldsymbol{\sigma} := (\sigma_1,\ldots,\sigma_k)$ determines their widths. More precisely, then, we say that $\boldsymbol{\mu} \in \mathcal{M}_{\text{NS}}$ is a Gaussian PR box of order k, with center vector $(\boldsymbol{a},\boldsymbol{b})$ and width vector $\boldsymbol{\sigma}$, if it is of the form

$$\mu_{x,y}^{(k,\boldsymbol{a},\boldsymbol{b},\boldsymbol{\sigma})} := \frac{1}{k}\sum_{j=1}^{k} \mathcal{N}_{(a_j,b_{[j+xy]_k}),\sigma_j}, \tag{8}$$

where $[\ ]_k$ denotes modulo k and $\mathcal{N}_{(a,b),\sigma}$ is the normal (Gaussian) measure centered at (a,b) and with width σ, defined through Eq. (1) with the probability density

$$p_{(a,b),\sigma}(a',b') = \frac{1}{2\pi\sigma^2} e^{-\frac{(a-a')^2+(b-b')^2}{2\sigma^2}}. \tag{9}$$

Whether a Gaussian PR box is supraquantum depends on $(\boldsymbol{a},\boldsymbol{b})$ and $\boldsymbol{\sigma}$. As an example, consider next the simple case with $k=2$, $\boldsymbol{a} = (\ell,-\ell) = \boldsymbol{b}$, for some arbitrary $\ell \in \mathbb{R}_{\neq 0}$, and $\boldsymbol{\sigma} = (\sigma,\sigma)$, graphically represented in Fig. 3(a). It is immediate to see that the resulting behavior violates the CFRD inequality by the amount

$$\max\{8\ell^4 - 4(\sigma^2+\ell^2)^2, 0\}. \tag{10}$$

This violation is plotted in Fig. 3(b) as a function of σ/ℓ. Note that it grows unboundedly with ℓ. The condition for this Gaussian PR box to violate the CFRD inequality is $\ell/\sigma \geqslant (1+\sqrt{2})^{1/2} \approx 1.55$, as can be graphically appreciated in the figure. In turn, taking, for the Gaussian PR box above, the limit $\sigma \to 0$, one obtains the behavior with components

$$\mu_{x,y} = \tfrac{1}{2}[\delta_{\ell,(-1)^{xy}\ell} + \delta_{-\ell,-(-1)^{xy}\ell}], \tag{11}$$

with δ being the measure defined in Eq. (5). This limiting box violates the CFRD inequality by $4\ell^4$. In fact, it is the CV version of the original dichotomic-input dichotomic output PR box [3].

Similarly, to define generic CV PR boxes, we take the $\boldsymbol{\sigma} \to \mathbf{0}$ limit of the Gaussian PR boxes of Eq. (12). That is, we say

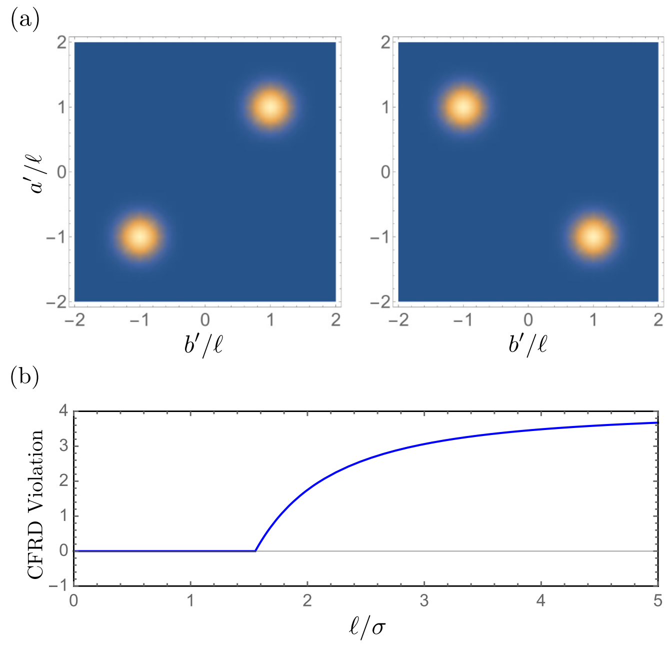

FIG. 3. (a) Density plots of a Gaussian PR box of order two, with center vector characterized by $\boldsymbol{a} = (\ell,-\ell) = \boldsymbol{b}$ and width vector $\boldsymbol{\sigma} = (\ell/5,\ell/5)$, for the inputs $(x,y) = (0,0)$, $(0,1)$, or $(1,0)$ (left) and $(x,y) = (1,1)$ (right). Note that, for both plots, the projections onto the horizontal as well as vertical axes coincide, reflecting the fact that the behavior is no-signaling. Each center point may also have a different width (or squeezing), but we do not consider that here for simplicity. (b) Violation of the CFRD inequality, normalized by the factor ℓ^4, by the Gaussian behavior in question as a function of the parameter ℓ/σ. The CFRD inequality certifies that the Gaussian PR box is supraquantum for the parameter region with $\ell/\sigma \geqslant (1+\sqrt{2})^{1/2} \approx 1.55$.

that $\boldsymbol{\mu}^{(k,\boldsymbol{a},\boldsymbol{b})} \in \mathcal{M}_{\text{NS}}$ is a CV PR box of order k and center vector $(\boldsymbol{a},\boldsymbol{b})$, with different real components such that $a_1 \neq a_2 \neq \cdots a_k$ and $b_1 \neq b_2 \neq \cdots b_k$, if it is of the form

$$\mu_{x,y}^{(k,\boldsymbol{a},\boldsymbol{b})} := \mu_{x,y}^{(k,\boldsymbol{a},\boldsymbol{b},\mathbf{0})} = \frac{1}{k}\sum_{j=1}^{k} \delta_{a_j,b_{[j+xy]_k}}. \tag{12}$$

One can immediately verify that these behaviors fulfill the no-signaling constraints (2). These boxes are the CV version of the finite-dimensional PR boxes generalized to arbitrarily many outputs and dichotomic inputs given in Ref. [15]. Still, Eq. (12) does not yet describe the most general CV PR box, because input and output relabelling symmetries must be taken into account. For dichotomic inputs, the possible local, reversible relabelings are given by $x \to [x+1]_2$ and $y \to [y+1]_2$ [15]. The situation is notably different, however, for the outputs, because they are continuous. For CV outputs, the most general local, reversible relabelings are given by $\boldsymbol{a} \to \boldsymbol{\alpha}_x(\boldsymbol{a})$ and $\boldsymbol{b} \to \boldsymbol{\beta}_y(\boldsymbol{b})$, where $\boldsymbol{\alpha}_x : \mathbb{R}^k \to \mathbb{R}^k$ and $\boldsymbol{\beta}_y : \mathbb{R}^k \to \mathbb{R}^k$ are, for every $x,y \in \{0,1\}$, bijective maps from $\mathbb{R}^k$ to itself. This amounts to reshuffling the components of the center vectors in a reversible, input-dependent fashion, so that the condition $[\boldsymbol{\alpha}_x(\boldsymbol{a})]_1 \neq [\boldsymbol{\alpha}_x(\boldsymbol{a})]_2 \neq \cdots [\boldsymbol{\alpha}_x(\boldsymbol{a})]_k$ and $[\boldsymbol{\beta}_y(\boldsymbol{b})]_1 \neq [\boldsymbol{\beta}_y(\boldsymbol{b})]_2 \neq \cdots [\boldsymbol{\beta}_y(\boldsymbol{b})]_k$ is always maintained.

Since the relabelings are local and reversible, all boxes equivalent under them have the same nonlocality properties. Indeed, all the boxes given by Eq. (12), i.e., for all different center vectors, are equivalent under input-independent relabelings. So, any of them, i.e., for any fixed center vector, can be taken as representative to define (modulo local, reversible, and input-dependent relabelings) the entire class of all CV PR boxes. This is, in turn, equivalent to allowing for input-dependent center vectors $(\boldsymbol{a}_x,\boldsymbol{b}_y)$ directly in the definition:

Definition 2. (Set of CV PR boxes) We define the class $\mathcal{M}_{\mathrm{PR}}$ as the set

$$\mathcal{M}_{\mathrm{PR}} := \{\boldsymbol{\mu}^{(k,\boldsymbol{a}_0,\boldsymbol{a}_1,\boldsymbol{b}_0,\boldsymbol{b}_1)} \in \mathcal{M}_{\mathrm{NS}}\}_{k\in\mathbb{N},\, \boldsymbol{a}_0,\boldsymbol{a}_1,\boldsymbol{b}_0,\boldsymbol{b}_1\in\mathbb{R}^k}, \tag{13}$$

where each behavior component $[\boldsymbol{\mu}^{(k,\boldsymbol{a}_0,\boldsymbol{a}_1,\boldsymbol{b}_0,\boldsymbol{b}_1)}]_{x,y}$ is given by a measure $\mu_{x,y}^{(k,\boldsymbol{a}_x,\boldsymbol{b}_y)}$ as in Eq. (12), with a possibly different vector $(\boldsymbol{a}_x,\boldsymbol{b}_y)$ for each $(x,y)\in\{0,1\}^2$.

Note that for $k=1$, CV PR boxes reduce to local, deterministic behaviors, whose components are given by Dirac delta measures. In contrast, for all $k \geqslant 2$, Def. 2 yields nonlocal, nondeterministic behaviors. Here, for simplicity, we use the term "CV PR box" for all $k\in\mathbb{N}$ indistinctly, the distinction between local, deterministic, and nonlocal, nondeterministic ones being given by the order k. In the next section, we show that every element of $\mathcal{M}_{\mathrm{PR}}$ is an extreme behavior of $\mathcal{M}_{\mathrm{NS}}$ and that the convex hull of $\mathcal{M}_{\mathrm{PR}}$ is dense in $\mathcal{M}_{\mathrm{NS}}$.

IV. CHARACTERIZATION OF SET OF NO-SIGNALLING BEHAVIORS

We start by recapping basic definitions of convex combinations and extremality. The convex hull $\mathrm{Conv}(\mathcal{M})$ of an arbitrary (finite or infinite) set $\mathcal{M}$ of behaviors is the set of all finite convex sums of elements of $\mathcal{M}$:

$$\mathrm{Conv}(\mathcal{M}) = \left\{\sum_{i=1}^{n} q_i\boldsymbol{\mu}_i : \boldsymbol{\mu}_i\in\mathcal{M}\right\}_{q_i\geqslant 0,\ \sum_{i=1}^n q_i=1,\ n\in\mathbb{N}}. \tag{14}$$

In turn, if $\mathcal{M}$ contains an uncountably infinite number of elements, continuous convex combinations (i.e., convex integrals) of infinitely many elements can be considered, too, but are not necessarily contained in $\mathrm{Conv}(\mathcal{M})$.

Clearly, any behavior that admits a decomposition in terms of a convex integral of uncountably infinitely many behaviors, admits also a decomposition in terms of a convex sum of finitely many behaviors. Similarly, any behavior that admits a decomposition in terms of a convex sum of an arbitrary finite number of behaviors admits also a decomposition in terms of a convex sum of two behaviors. This leads us to the same definition of extreme no-signaling behaviors as in discrete variables.

Definition 3 (Extreme no-signaling behaviors). We call $\boldsymbol{\mu}$ an *extreme point* of $\mathcal{M}_{\mathrm{NS}}$ if, for any $\boldsymbol{\mu}^*,\boldsymbol{\mu}'\in\mathcal{M}_{\mathrm{NS}}$ and $0\leqslant q\leqslant 1$ such that

$$\boldsymbol{\mu} = q\boldsymbol{\mu}^* + (1-q)\boldsymbol{\mu}', \tag{15}$$

it holds that either $q=1$ and $\boldsymbol{\mu}^*=\boldsymbol{\mu}$, or $q=0$ and $\boldsymbol{\mu}'=\boldsymbol{\mu}$.

Now we know that every $\boldsymbol{\mu}\in\mathcal{M}_{\mathrm{PR}}$ has a finite number of outcomes with nonzero probability and belongs to $\mathcal{M}_{\mathrm{NS}}$. That is, $\boldsymbol{\mu}$ is either an extreme point of $\mathcal{M}_{\mathrm{NS}}$ or it can be decomposed as the convex sum of at most finitely many points in $\mathcal{M}_{\mathrm{NS}}$. However, the fact that finite-dimensional PR boxes are no-signaling extreme implies that the former is the case. This follows from the fact that finite-dimensional PR boxes are given by an equivalent expression to that in Eq. (12) where Kronecker deltas are in the place of the Dirac ones [15]. This proves, then, that all CV PR boxes are no-signaling extreme:

Observation 1. (Extremality of $\mathcal{M}_{\mathrm{PR}}$) All elements of $\mathcal{M}_{\mathrm{PR}}$ are extreme points of $\mathcal{M}_{\mathrm{NS}}$.

Observation IV constitutes, in turn, a generalization to the CV realm of the result of Ref. [38], where it is shown that any extreme point of the no-signaling set with a given finite number of inputs and outputs is also extreme in the no-signaling set with any higher (but still finite) number of inputs and outputs. In addition, since $\mathcal{M}_{\mathrm{PR}}$ is not finite, the observation also directly implies that $\mathcal{M}_{\mathrm{NS}}$ is not a polytope. On the other hand, the fact that $\mathcal{M}_{\mathrm{NS}}$ contains behaviors with infinitely many outcomes with nonzero probability (e.g., the Gaussian PR boxes of the previous section) automatically implies that $\mathcal{M}_{\mathrm{NS}} \not\subseteq \mathrm{Conv}(\mathcal{M}_{\mathrm{PR}})$, in striking contrast with the finite-dimensional case. This is due to the fact that every behavior in $\mathrm{Conv}(\mathcal{M}_{\mathrm{PR}})$ necessarily has only finitely many outcomes with nonzero probability. Nevertheless, we show in Appendix B that $\mathcal{M}_{\mathrm{NS}}$ is approximated arbitrarily well by $\mathrm{Conv}(\mathcal{M}_{\mathrm{PR}})$, in the formal sense of there existing, for all $\boldsymbol{\mu}\in\mathcal{M}_{\mathrm{NS}}$, a sequence of elements in $\mathrm{Conv}(\mathcal{M}_{\mathrm{PR}})$ that converges to $\boldsymbol{\mu}$. This proves the following:

Theorem 1. [$\mathrm{Conv}(\mathcal{M}_{\mathrm{PR}})$ dense in $\mathcal{M}_{\mathrm{NS}}$] The closure $\overline{\mathrm{Conv}(\mathcal{M}_{\mathrm{PR}})}$ of $\mathrm{Conv}(\mathcal{M}_{\mathrm{PR}})$ equals $\mathcal{M}_{\mathrm{NS}}$. In other words, $\mathrm{Conv}(\mathcal{M}_{\mathrm{PR}})$ is a *dense subset* of $\mathcal{M}_{\mathrm{NS}}$.

The theorem is proven in detail in Appendix B. Let us sketch the proof idea here. We consider first the case of behaviors defined on a compact domain $[-\mathcal{K},\mathcal{K}]^2$. There, we can use standard techniques from measure theory to show that, for any no-signaling behavior $\boldsymbol{\mu}$, one can find a sequence of convex sums of CV PR boxes $\boldsymbol{\mu}_n$ that converges to it. The main idea is then to define the considered sequence in such a way that its components become good approximations of the components of $\boldsymbol{\mu}$, in the limit of large n. This procedure can be seen as a generalization to the approximation of a function by piecewise constant functions as it is used in integration theory. Next, one generalizes this further to an infinite sequence of compact intervals which, in the infinite-length limit, covers the whole space $\mathbb{R}\times\mathbb{R}$.

Even though $\mathcal{M}_{\mathrm{PR}}$ consists exclusively of extreme points of $\mathcal{M}_{\mathrm{NS}}$, the fact that $\mathrm{Conv}(\mathcal{M}_{\mathrm{PR}})$ is a strict subset of $\mathcal{M}_{\mathrm{NS}}$ in principle leaves room for other extreme points in $\mathcal{M}_{\mathrm{NS}}$ that are not contained in $\mathcal{M}_{\mathrm{PR}}$. In the following, we approach this problem systematically by focusing first on behaviors with compact support. In this case, a related problem was addressed by Milman, who proved that, given a compact convex subset $\mathcal{C}$ of a locally convex space $\mathcal{E}$ (see Ref. [39] for a definition of locally convex) and another set $\mathcal{T}\subset\mathcal{C}$ such that $\overline{\mathrm{Conv}(\mathcal{T})}=\mathcal{C}$, it follows that all extreme points of $\mathcal{C}$ are in the closure of $\mathcal{T}$ [39]. The space $\mathcal{M}_{[-\mathcal{K},\mathcal{K}]^2}$ of probability measures with bounded domain $[-\mathcal{K},\mathcal{K}]^2\subset\mathbb{R}^2$, is a compact subset of the locally convex space of all measures on the same domain. The same holds also for the set of behaviors $\mathcal{M}^4_{[-\mathcal{K},\mathcal{K}]^2}$. Moreover, the set of no-signaling behaviors on $[-\mathcal{K},\mathcal{K}]^2$ is a closed subset of $\mathcal{M}^4_{[-\mathcal{K},\mathcal{K}]^2}$ and thus also compact, which enables us to use Milman's theorem to characterize its extreme points. In what follows, we deal with no-signaling and PR box behaviors on a compact domain. To emphasize this, we equip the corresponding no-signaling set and the set of CV PR boxes with a superscript $\mathcal{K}$, i.e., $\mathcal{M}_{\rm NS}^{(\mathcal{K})}$ and $\mathcal{M}_{\rm PR}^{(\mathcal{K})}$. Consequently, we arrive at the following corollary:

Corollary 1. (Characterization of $\mathcal{M}_{\rm NS}^{(\mathcal{K})}$) Every extreme point of $\mathcal{M}_{\rm NS}^{(\mathcal{K})}$ belongs to the closure of $\mathcal{M}_{\rm PR}^{(\mathcal{K})}$.

Further on, it is interesting to investigate if the closure of $\mathcal{M}_{\rm PR}^{(\mathcal{K})}$ contains behaviors that are extreme as well. If this were not the case, it would prove that all extreme points of $\mathcal{M}_{\rm NS}^{(\mathcal{K})}$ are in $\mathcal{M}_{\rm PR}^{(\mathcal{K})}$. We thus have to answer the question if PR boxes of infinite order, i.e., in the limit $k \to \infty$ [see Eq. (12)], are also extreme. In Appendix C we provide evidence suggesting that this is not the case. More precisely, we provide an exemplary sequence of PR boxes whose limiting behavior is not extreme, thus implying that $\mathcal{M}_{\rm PR}^{(\mathcal{K})}$ is not a closed set. This evidence leads us to the following conjecture:

Conjecture 1. (Characterization of $\mathcal{M}_{\rm NS}^{(\mathcal{K})}$) Every extreme point of $\mathcal{M}_{\rm NS}^{(\mathcal{K})}$ belongs to $\mathcal{M}_{\rm PR}^{(\mathcal{K})}$.

Even though the preceding discussion was restricted to behaviors with outcomes on a compact set, we have reasons to believe that the conjecture holds also in the general case of unbounded support. Namely, in probability theory it is a rather standard result that all extreme points of the set of probability measures are given by Dirac measures [see Eq. (5)]. In particular, this is the case for probability measures defined on $\mathbb{R}$. Similarly, the extreme no-signaling behaviors may have also only finite support, which would suggest our Conjecture 1 also in the general case of behaviors defined on $\mathbb{R}$. A proof of Conjecture 1 would, however, require more involved arguments which go beyond the scope of the present article.

Let us finish with some final clarifications on the boundary and the boundedness of $\mathcal{M}_{\rm NS}$. In the finite-dimensional case, the boundary between the no-signaling behaviors and behavior-like objects that still satisfy the no-signaling constraints but involve nonpositive probability distributions is given by the subset of all convex combinations of no-signaling extreme points resulting in nonstrictly positive behaviors [i.e., whose (x,y)th components are probability measures assigning zero probability to some event]. Consequently, the set of no-signaling behaviors has a nonempty interior. In contrast, for infinite dimensional behaviors, the boundary of $\mathcal{M}_{\rm NS}$ is actually $\mathcal{M}_{\rm NS}$ itself, showing that its interior is empty. The latter can be proven by using convergence arguments similar to those used in the proof of Theorem 1, i.e. every no-signaling behavior is arbitrarily close (in the weak-convergence sense) to a nonpositive no-signaling behavior. This may at first blush seem bizarre, but it is actually a typical property of compact convex sets in infinite-dimensional spaces. Indeed, the sets of probability distributions or quantum states for infinite-dimensional systems display exactly the same property (see, e.g., Ref. [40]).

Lastly, we stress that, in the present work, we did not touch the question of whether the set $\mathcal{M}_{\rm NS}$ is bounded. Doing so would require to introduce an appropriate metric and, because we are dealing with infinite-dimensional spaces, the boundedness of the set $\mathcal{M}_{\rm NS}$ might depend on its particular choice. For instance, with respect to the Lévy–Prokhorov metric, which is a metric on the set of probability measures associated with the weak topology, the set of all probability measures is bounded. Hence, for this metric also the no-signaling set is bounded, since the components of behaviors are by definition always probability measures.

V. FINAL DISCUSSION

We have studied supraquantum Bell correlations in a genuinely CV regime, i.e., without discretization procedures such as binning [20–24,27,28]. Here, genuine CV supraquantumness was witnessed by the violation of the CFRD inequality [25], which, for the bipartite case, is known not to admit any quantum violation [29,30]. We found a class of supraquantum Gaussian PR boxes, whose zero-width limit gives the CV PR boxes. Here, we have explicitly checked the supraquantumness of both Gaussian and CV PR boxes of order $k = 2$. Interestingly, due to symmetries in the CFRD inequality, no violation can be found for $k = 3$, but supraquantumness of CV PR boxes of higher orders is guaranteed by the supraquantumness of the equivalent boxes in finite dimensions. In turn, the supraquantumness of finite-width Gaussian PR boxes of higher order can be verified violating—via some appropriate binning—finite-dimensional Bell inequalities above their quantum limit; but this is outside the scope of this paper.

In addition, we have characterized the set of CV no-signaling correlations from a geometrical point of view. To this end, we devised a mathematical framework to deal with arbitrary CV no-signaling behaviors based on conditional probability measures instead of on conditional probability distributions. With this, we have shown that, for CV systems, the convex hull (i.e., the set of all finite convex sums) of all CV PR boxes is dense in the no-signaling set, instead of equal to it as in finite-dimensional systems. In particular, this result tells us that every no-signaling behavior can be approximated arbitrarily well by a sequence of behaviors with a finite number of nonzero probability outcomes. Consequently, the nonlocality of every CV no-signaling behavior can always be detected with discrete Bell inequalities in combination with a binning procedure, for a sufficiently large number of bins.

Since every CV PR box assigns a nonzero probability to a finite number of outcomes, being thus in one-to-one correspondence with a discrete PR box in the usual finite-dimensional scenario, it is not surprising that every CV PR box is extreme in the no-signaling set. In contrast, the possibility that all extreme points of the no-signaling set are given by CV PR boxes, as suggested by Conjecture 1, appears as more surprising. Indeed, it would evidence a qualitative difference between the structure of quantum theory and that of generic probability theories compatible with the no-signaling principle, a question that has been previously considered in other scenarios, too [41]. Namely, in quantum theory we know about the existence of behaviors with an uncountably infinite number of nonzero probability outcomes which are extreme in the set of CV quantum correlations. The latter quantum behaviors can be built, e.g., with extreme quantum POVMs with a continuous spectrum [42–44] acting on pure CV entangled states. We leave the proof (or disproof) of this conjecture as an open question for future investigations.

Another interesting question for future investigations is how to formalize the notion of tightness [36] for CV Bell inequalities, and, in particular, whether the CFRD inequality is tight. In finite dimensions, nontrivial tight Bell inequalities without a quantum violation exist in the multipartite scenario [17], but no equivalent example is known for bipartite systems. If the CFRD inequality were tight, our results would give it the status of the first known example of a nontrivial tight Bell inequality with no quantum violation in the bipartite setting.

To end up with, far from being just a mere abstract exercise, studying supraquantum nonlocality helps us understand quantum nonlocality itself. Efficient tools to study nonlocality for discrete systems—such as semidefinite or linear programming—no longer apply for CV systems; so that the characterization of nonlocal correlations is a much harder task. We thus hope that our findings can be useful for future research, such as, e.g., searching for novel CV Bell inequalities or, more generally, studying generalized-probability theories in CV systems.

ACKNOWLEDGMENTS

L.A. thanks D. Cavalcanti for useful comments and the International Institute of Physics (IIP) Natal, Brazil, where parts of this work were carried out, for their hospitality. L.A.'s work is financed by the Brazilian ministries MEC and MCTIC and Brazilian agencies CNPq, CAPES, FAPESP, FAPERJ, and INCT-IQ. A.L.F. and A.K. acknowledge CAPES-COFECUB project Ph-855/15 and CNPq for financial support. A.K. acknowledges financial support from the ERC (Consolidator Grant 683107/TempoQ), and the DFG. A special thanks goes to T. Cabana, S. Kaakai, and C. Ketterer for the clarification of mathematical details, and to J. P. Pellonpää for valuable discussions about extremality problems in quantum mechanics.

A.K. and A.L.-F. contributed equally to this work.

APPENDIX A: RADON-NIKODYM THEOREM

A measurable space is given by a pair (X,Σ), where X is some nonempty set and Σ denotes a σ algebra on X. We define a measure ν on X to be σ finite if X is a countable union of measurable sets X_i with finite measure $\nu(X_i) < \infty$. Note that every probability measure μ on $\mathbb{R}$ is also σ finite since, on the one hand, $\mathbb{R}$ can be expressed as countable union of measurable set and, on the other hand, we have by definition that $\mu(A) < 1$, for all $A \subset \mathbb{R}$. Furthermore, a measure ν is called absolutely continuous with respect to μ, if from $\nu(A) = 0$ it follows that $\mu(A) = 0$, for every measurable set $A \subset X$. Now, we are in the position to state the Radon–Nikodym Theorem.

Theorem 2. (Radon–Nikodym) Given a σ-finite measure ν on (X,Σ) that is absolutely continuous with respect to a σ-finite measure μ on (X,Σ), then there exists a measurable function $f : X \to [0,\infty)$, referred to as the Radon–Nikodym derivative, such that

$$\nu(A) = \int_A f d\mu, \tag{A1}$$

for any measurable subset $A \subset X$.

APPENDIX B: PROOF OF THEOREM 1

Before turning to the proof of Theorem 1 we provide some preliminary notions of the type of convergence that we use in the following, i.e., the weak convergence. We say that a sequence of measures $(\mu_n)_{n\in\mathbb{N}} \in \mathcal{M}_\Omega$ converges weakly towards the same $\mu \in \mathcal{M}_\Omega$, with $n \to \infty$, if

$$\int_\Omega f d\mu_n \to \int_\Omega f d\mu, \tag{B1}$$

for all $f \in C_b(\Omega)$, where $C_b(\Omega)$ denotes the set of bounded and continuous functions $f : \Omega \to \mathbb{R}$. In what follows, if not stated differently, we always implicitly assume the use of weak convergence for sequences of measures. Moreover, since we often consider behaviors (i.e., matrices with entries given by probability measures), we say that a sequence of behaviors $\boldsymbol{\mu}_n$ weakly converges to $\boldsymbol{\mu}$ if $[\boldsymbol{\mu}_n]_{x,y} \to [\boldsymbol{\mu}]_{x,y}, \forall x,y$.

Weak convergence is a natural choice in the present context because it is directly applicable to sequences of measures without resorting to a specific distributions in terms of some random variables. Other, possibly stronger, notions of convergence do exist but are not required here. Furthermore, as we will see shortly, weak convergence is also meaningful with respect to physical considerations since, from experiments, one usually extracts some statistical moments of a probability measure instead of the measure itself.

Further on, as stated also in the main text, to prove that $\mathrm{Conv}(\mathcal{M}_{\mathrm{PR}})$ is dense in $\mathcal{M}_{\mathrm{NS}}$ we need to show that, for every behavior $\boldsymbol{\mu} \in \mathcal{M}_{\mathrm{NS}}$, one can find a sequence in $\mathrm{Conv}(\mathcal{M}_{\mathrm{PR}})$ that converges weakly to $\boldsymbol{\mu}$. To keep the proof of Theorem 1 as instructive as possible, we first provide a proof for the case of behaviors with compact support meaning that their components are probability measures on $\Omega = [-\mathcal{K},\mathcal{K}]^2$. A generalization of the proof to the most general case $\Omega = \mathbb{R} \times \mathbb{R}$ will then be ensued afterwards.

Then, the following lemma holds:

Lemma 1. [$\mathrm{Conv}(\mathcal{M}_{\mathrm{PR}}^{(\mathcal{K})})$ dense in $\mathcal{M}_{\mathrm{NS}}^{(\mathcal{K})}$] The closure $\overline{\mathrm{Conv}(\mathcal{M}_{\mathrm{PR}}^{(\mathcal{K})})}$ of $\mathrm{Conv}(\mathcal{M}_{\mathrm{PR}}^{(\mathcal{K})})$ equals $\mathcal{M}_{\mathrm{NS}}^{(\mathcal{K})}$. In other words, $\mathrm{Conv}(\mathcal{M}_{\mathrm{PR}}^{(\mathcal{K})})$ is a *dense subset* of $\mathcal{M}_{\mathrm{NS}}^{(\mathcal{K})}$.

Note that, according to the introduced nomenclature in the main text we equipped the corresponding no-signaling set and the set of CV PR boxes with a superscript $\mathcal{K}$, i.e., $\mathcal{M}_{\mathrm{NS}}^{(\mathcal{K})}$ and $\mathcal{M}_{\mathrm{PR}}^{(\mathcal{K})}$.

Proof of Lemma 1. Without loss of generality we can restrict the following proof to the case $\mathcal{K} = 1$, i.e., $\Omega = [-1,1]^2$. The strategy consists of explicitly constructing, for every arbitrary $\boldsymbol{\mu} \in \mathcal{M}_{\mathrm{NS}}^{(1)}$, a sequence of behaviors $\boldsymbol{\mu}_n \in \mathrm{Conv}(\mathcal{M}_{\mathrm{PR}}^{(1)})$ that weakly converges to $\boldsymbol{\mu}$. The proof is divided in three steps: First, for every $\boldsymbol{\mu} \in \mathcal{M}_{\mathrm{NS}}$, we define a sequence of behaviors that weakly converges to $\boldsymbol{\mu}$. Second, we show that each element of this sequence is indeed a no-signaling behavior. Third, we show that all such elements can be expressed as a convex sum of CV PR boxes.

For the first step, we divide the interval $[-1,1]$ in $n \geqslant 1$ segments of the same length, denoting each one by I_n (note that a generalization of the following proof to arbitrary $\mathcal{K}$s would simply involve a rescaling of the defined intervals I_n). Next, we define $\boldsymbol{\mu}_n$ as follows:

$$[\boldsymbol{\mu}_n]_{x,y} = \sum_{k,l=1}^{n} [\boldsymbol{\mu}]_{x,y}(I_k \times I_l)\delta_{a_k,b_l} \;\forall\, x,y, \tag{B2}$$

where (a_k,b_l) is a point located in the interval $I_k \times I_l$ and the Dirac measure is defined according to Eq. (5). The behaviors $\boldsymbol{\mu}_n$ have the same weight as $\boldsymbol{\mu}$ on each of the squares $I_k \times I_l$, but concentrated on a single point (a_k,b_l). In this way, $\boldsymbol{\mu}_n$ becomes a better and better approximation of $\boldsymbol{\mu}$ with increasing n.

To prove that $\boldsymbol{\mu}_n$ is indeed weakly converging to $\boldsymbol{\mu}$, it suffices to prove that each of its components is weakly converging to the components of $\boldsymbol{\mu}$. Let f be a bounded and continuous function defined on the domain $[-1,1] \times [-1,1]$. Integrating f with respect to $\mu_n^{x,y}$ yields

$$\int_{[-1,1]^2} f d[\boldsymbol{\mu}_n]_{x,y} = \sum_{k,l} f(a_k,b_l)[\boldsymbol{\mu}]_{x,y}(I_k \times I_l). \tag{B3}$$

The sum on the right-hand side of Eq. (B3) can be bounded from below and above in the following way:

$$\sum_{k,l} [\boldsymbol{\mu}]_{x,y}(I_k \times I_l)\min_{I_k\times I_l} f \leqslant \int_{[-1,1]\times[-1,1]} f d[\boldsymbol{\mu}_n]_{x,y} \tag{B4}$$

$$\leqslant \sum_{k,l} [\boldsymbol{\mu}]_{x,y}(I_k \times I_l)\max_{I_k\times I_l} f, \tag{B5}$$

where $\min_{I_k\times I_l}$ ($\max_{I_k\times I_l}$) denotes the minimum (maximum) of the function f over the cell $I_k \times I_l$. The same inequality holds if we integrate f with respect to $[\boldsymbol{\mu}]_{x,y}$, and, since f is continuous, this proves that $\int f d[\boldsymbol{\mu}_n]_{x,y} \to \int f d[\boldsymbol{\mu}]_{x,y}$ and that $[\mu_n]_{x,y} \to \mu, \forall x,y$. It follows that $\boldsymbol{\mu}_n \to \boldsymbol{\mu}$.

As for the second step, we now prove that $\boldsymbol{\mu}_n$ is no-signaling for all n. For a given $n > 0$ and $x,y \in \{0,1\}^n$, the marginal of $[\mu_n]_{x,y}$ on Bob's side is given by

$$\begin{aligned}[\boldsymbol{\mu}_n]_{x,y}([-1,1] \times B) &= \sum_{k,l=1}^{n} [\boldsymbol{\mu}]_{x,y}(I_k \times I_l)\delta_{a_k,b_l}([-1,1] \times B)\\ &= \sum_{l} \delta_{b_l}(B) \sum_{k} [\boldsymbol{\mu}]_{x,y}(I_k \times I_l)\\ &= \sum_{l} \delta_{b_l}(B)[\boldsymbol{\mu}]_{x,y}([-1,1] \times I_l),\end{aligned} \tag{B6}$$

where δ_{b_l} is the Dirac measure located at b_l in the lth interval. Since we know that $\boldsymbol{\mu}$ is a no-signaling behavior it follows that $[\boldsymbol{\mu}]_{x,y}([-1,1] \times I_l)$ does not depend on x [compare with Eq. (2a)]. The same argument holds for the Alice's marginal and proves that the $\boldsymbol{\mu}_n$ are no-signaling behaviors.

The third and last step to complete the proof is to show that $\boldsymbol{\mu}_n$ can be written as a convex sum of finitely many CV PR boxes. For this we note that the $\boldsymbol{\mu}_n$ are no-signaling behaviors with a finite number of outcomes (the centers of the intervals $I_{k,l}$) and support $[-1,1]^2$. However, we know from the finite-dimensional case that all behaviors with only finitely many outcomes with nonzero probability can be expressed as a convex combination of finitely many PR boxes. Taking instead their continuous-variable generalizations (12), yields the desired decomposition. ■

With Lemma 1, we next prove Theorem 1.

Proof of Theorem 1. Now we consider the case $\Omega = \mathbb{R} \times \mathbb{R}$. Again, we consider a $\boldsymbol{\mu} \in \mathcal{M}_{\text{PR}}$ and want to prove that there exists a sequence $\boldsymbol{\mu}_n \in \text{Conv}(\mathcal{M}_{\text{PR}})$ for which each component converges weakly to the components of $\boldsymbol{\mu}$. To do so, we divide $[-n,n]$, with $n \geqslant 1$, in $2n^2$ subintervals of length $1/n$ and denote them by I_n as before. Furthermore, we define the components of $\boldsymbol{\mu}_n$ as follows:

$$[\boldsymbol{\mu}_n]_{x,y} = \sum_{k,l=1}^{2n^2} [\boldsymbol{\mu}]_{x,y}(I_k \times I_l)\delta_{a_k,b_l} + [\boldsymbol{\nu}_n]_{x,y}, \tag{B7}$$

where (a_k,b_l) is a point located in the square $I_k \times I_l$, and δ_{a_k,b_l} is the Dirac measure. The first term of Eq. (B7) corresponds to the same construction as in the compact case treated in Lemma B, whereas the second term $\boldsymbol{\nu}_n$ is merely necessary to ensure the no-signaling conditions (2a) and (2b) on $\mathbb{R} \times \mathbb{R}$. It reads as follows:

$$\begin{aligned}[\boldsymbol{\nu}_n]_{x,y} = &\sum_{l=-n}^{n} \big\{[\boldsymbol{\mu}]_{x,y}(]n,\infty[\times I_l)\delta_{n+1,b_l}\\ &+[\boldsymbol{\mu}]_{x,y}(]-\infty,-n[\times I_l)\delta_{-(n+1),b_l}\big\}\\ &+\sum_{k=-n}^{n} \big\{[\boldsymbol{\mu}]_{x,y}(I_k\times]n,\infty[)\delta_{a_k,(n+1)}\\ &+[\boldsymbol{\mu}]_{x,y}(I_k\times]-\infty,-n[)\delta_{a_k,-(n+1)}\big\}\\ &+[\boldsymbol{\mu}]_{x,y}(]n,\infty[\times]n,\infty[)\delta_{n+1,n+1}\\ &+[\boldsymbol{\mu}]_{x,y}(]-\infty,-n[\times]n,\infty[)\delta_{-(n+1),n+1}\\ &+[\boldsymbol{\mu}]_{x,y}(]n,\infty[\times]-\infty,-n[)\delta_{n+1,-(n+1)}\\ &+[\boldsymbol{\mu}]_{x,y}(]-\infty,-n[\times]-\infty,-n[)\delta_{-(n+1),-(n+1)},\end{aligned} \tag{B8}$$

where $]a,b[$ refers to an open interval bounded by a and b, respectively. Note that, in contrast to the compact case treated in Lemma 1, the measures $[\boldsymbol{\mu}_n]_{x,y}$ are defined on different intervals for different n. We now complete the proof of Theorem 1 by showing the weak convergence of this sequence in the general case. The other parts of the proof remain unchanged.

Let $f \in \mathcal{C}_b(\mathbb{R}^2)$ and $\epsilon \in [0,1]$, we want to prove that there exists an $n_0 \in \mathbb{N}$ such that $|\int_{\mathbb{R}^2} f d\boldsymbol{\mu}_n - \int_{\mathbb{R}^2} f d\boldsymbol{\mu}| < \epsilon$ for all $n > n_0$, where this inequality should be understood as component wise inequality. Since $\boldsymbol{\mu}$ is a set of probability measures and f is a bounded function, there exists an $n_1 \in \mathbb{N}$ such that:

$$[\boldsymbol{\mu}]_{x,y}(\mathbb{R}^2 \setminus [-n_1,n_1]) < \min\left(\epsilon, \frac{\epsilon}{\max_{\mathbb{R}^2}|f|}\right) \tag{B9}$$

for all (x,y) and all $n > n_1$. It follows that

$$\begin{aligned}&\left|\int_{\mathbb{R}^2} f d[\boldsymbol{\mu}_n]_{x,y} - \int_{\mathbb{R}^2} f d[\boldsymbol{\mu}]_{x,y}\right|\\ &< \left|\int_{\mathbb{R}^2\setminus[-n_1,n_1]^2} f d[\boldsymbol{\mu}_n]_{x,y} - \int_{\mathbb{R}^2\setminus[-n_1,n_1]^2} f d[\boldsymbol{\mu}]_{x,y}\right|\\ &+ \left|\int_{[-n_1,n_1]^2} f d[\boldsymbol{\mu}_n]_{x,y} - \int_{[-n_1,n_1]^2} f d[\boldsymbol{\mu}]_{x,y}\right|.\end{aligned} \tag{B10}$$

While the first term on the right-hand side of inequality (B10) becomes

$$\begin{aligned}&\left|\int_{\mathbb{R}^2\setminus[-n_1,n_1]^2} f d[\boldsymbol{\mu}_n]_{x,y} - \int_{\mathbb{R}^2\setminus[-n_1,n_1]^2} f d[\boldsymbol{\mu}]_{x,y}\right|\\ &\leqslant \left|\int_{\mathbb{R}^2\setminus[-n_1,n_1]^2} f d[\boldsymbol{\mu}_n]_{x,y}\right| + \left|\int_{\mathbb{R}^2\setminus[-n_1,n_1]^2} f d[\boldsymbol{\mu}]_{x,y}\right|\\ &\leqslant \max_{\mathbb{R}^2}|f|[\boldsymbol{\mu}_n]_{x,y}(\mathbb{R}^2\setminus[-n_1,n_1]^2)+\epsilon\\ &= \max_{\mathbb{R}^2}|f|[\boldsymbol{\mu}]_{x,y}(\mathbb{R}^2\setminus[-n_1,n_1]^2)+\epsilon\\ &\leqslant 2\epsilon, \end{aligned} \tag{B11}$$

the second term contains an integration over a compact area, which allows us to use the statement of Lemma 1. Hence, we can conclude that this term is smaller than ϵ for sufficiently large n. Note that Lemma 1 does not apply directly here since the considered sequence of behaviors is not no-signaling on the compact domain $[-n_1,n_1]^2$, but rather on $\mathbb{R}^2$. However, dropping the no-signaling condition does not contradict with the convergence of this sequence. By combining inequalities (B10) and (B11) we finally arrive at

$$\left|\int_{\mathbb{R}^2} f d[\boldsymbol{\mu}_n]_{x,y} - \int_{\mathbb{R}^2} f d[\boldsymbol{\mu}]_{x,y}\right| < 3\epsilon, \tag{B12}$$

for n sufficiently large. This quantity goes to zero as ϵ goes to zero and thus $\boldsymbol{\mu}_n$ weakly converges to $\boldsymbol{\mu}$.

APPENDIX C: CONCERNING CONJECTURE 1

Here we construct a specific example of a sequence of CV PR boxes, with increasing order k, whose limit is not an extreme no-signaling behavior anymore. This suggests that one cannot obtain extreme no-signaling behaviors as limits of a sequences of CV PR boxes when the order k goes to infinity. We restrict ourselves to measures on $[0,1]^2$ but it can be straightforwardly extended to $\mathbb{R}^2$.

Proof. We prove that there is a sequence $\boldsymbol{\mu}_n \in \mathcal{M}_{\rm PR}^{(1)}$ that converges to an element $\boldsymbol{\mu}$ that is outside of $\mathcal{M}_{\rm PR}^{(1)}$. Let $\boldsymbol{\mu}$ be the set of measures where the two outcomes are always perfectly correlated for all settings: $\mu^{x,y}(a,b) = \delta(a-b)$. $\boldsymbol{\mu}$ is clearly no-signaling, but not extreme.

We define $\boldsymbol{\mu}_n$ as follows:

$$\mu_n^{x,y} = \begin{cases} \frac{1}{n}\sum_{k=0}^{n}\delta_{\frac{k}{n},\frac{k}{n}}, & \text{for } x\cdot y = 0\\ \frac{1}{n}\left[\sum_{k=0}^{n-1}\delta_{\frac{k}{n},\frac{k+1}{n}} + \delta_{1,0}\right], & \text{for } x\cdot y = 1,\end{cases} \tag{C1}$$

which yields

$$\iint_{[0,1]^2} f(a,b)\mu_n^{x,y}(a,b) \tag{C2}$$

$$= \begin{cases} \frac{1}{n}\left[\sum_{k=0}^{n} f\left(\frac{k}{n},\frac{k}{n}\right)\right], & \text{for } x\cdot y = 0\\ \frac{1}{n}\left[\sum_{k=0}^{n-1} f\left(\frac{k}{n},\frac{k+1}{n}\right) + f(1,0)\right], & \text{for } x\cdot y = 1,\end{cases} \tag{C3}$$

where $f \in C_b([0,1]^2)$. Now, by a applying standard integration theory it follows that $[\frac{1}{n}\sum_{k=0}^{n-1} f(\frac{k}{n},\frac{k+1}{n}) + f(1,0)] \to \int_{[0,1]} f(a,a) = \iint_{[0,1]^2} f(a,b)\mu^{x,y}(a,b)$. We thus proved that $\boldsymbol{\mu}_n$ converges to an element that is outside of $\mathcal{M}_{\rm PR}^{(1)}$ (since $\boldsymbol{\mu}$ has an infinite number of outcomes contrary to all elements of $\mathcal{M}_{\rm PR}^{(1)}$). ■

[1] J. S. Bell, On the Einstein-Podolsky-Rosen paradox, Phys. **1**, 195 (1964).
[2] J. F. Clauser, M. A. Horne, A. Shimony, and R. A. Holt, Proposed Experiment to Test Local Hidden-Variable Theories, Phys. Rev. Lett. **23**, 880 (1969).
[3] S. Popescu and D. Rohrlich, Quantum nonlocality as an axiom, Found. Phys. **24**, 379 (1994).
[4] B. S. Cirel'son, Quantum generalizations of Bell's inequality, Lett. Math. Phys. **4**, 93 (1980).
[5] A. Elitzur, S. Popescu, and D. Rohrlich, Quantum nonlocality for each pair in an ensemble, Phys. Lett. A **162**, 25 (1992).
[6] J. Barrett, A. Kent, and S. Pironio, Maximally Nonlocal and Monogamous Quantum Correlations, Phys. Rev. Lett. **97**, 170409 (2006).
[7] W. van Dam, Ph.D. thesis, University of Amsterdam, 2002.
[8] W. van Dam, Implausible Consequences of Superstrong Nonlocality, Nat Comput. **12**, 9 (2013).
[9] G. Brassard, H. Buhrman, N. Linden, A. A. Méthot, A. Tapp, and F. Unger, Limit on Nonlocality in a World in which Communication Complexity is not Trivial, Phys. Rev. Lett. **96**, 250401 (2006).
[10] M. Pawlowski, T. Paterek, D. Kaszlikowski, V. Scarani, A. Winter, and M. Zukowski, Information causality as a physical principle, Nature (London) **461**, 1101 (2009).
[11] M. Navascués and H. Wunderlich, A glance beyond the quantum model, Proc. R. Soc. London, Ser. A **466**, 881 (2009).
[12] J. Oppenheim and S. Wehner, The uncertainty principle determines the non-locality of quantum mechanics, Science **330**, 1072 (2010).
[13] R. Gallego, L. E. Würflinger, A. Acín, and M. Navascués, Quantum Correlations Require Multipartite Information Principles, Phys. Rev. Lett. **107**, 210403 (2011).
[14] T. Fritz, A. B. Sainz, R. Augusiak, J. B. Brask, R. Chaves, A. Leverrier, and A. Acín, Local orthogonality as a multipartite principle for quantum correlations, Nat. Commun. **4**, 2263 (2013).
[15] J. Barrett, N. Linden, S. Massar, S. Pironio, S. Popescu, and D. Roberts, Nonlocal correlations as an information-theoretic resource, Phys. Rev. A **71**, 022101 (2005).
[16] J. Barrett and S. Pironio, Popescu-Rohrlich Correlations as a Unit of Nonlocality, Phys. Rev. Lett. **95**, 140401 (2005).
[17] M. Almeida, J.-D. Bancal, N. Brunner, A. Acín, N. Gisin, and S. Pironio, Guess your Neighbor's Input: a Multipartite Non-Local Game with no Quantum Advantage, Phys. Rev. Lett. **104**, 230404 (2010).

[18] A. B. Sainz, N. Brunner, D. Cavalcanti, P. Skrzypczyk, and T. Vértesi, Postquantum Steering, Phys. Rev. Lett. **115**, 190403 (2015).

[19] N. Brunner, D. Cavalcanti, S. Pironio, V. Scarani, and S. Wehner, Bell nonlocality, Rev. Mod. Phys. **86**, 419 (2014).

[20] P. Grangier, M. J. Potasek, and B. Yurke, Probing the phase coherence of parametrically generated photon pairs: A new test of Bell's inequalities, Phys. Rev. A **38**, 3132(R) (1988).

[21] S. M. Tan, D. F. Walls, and M. J. Collett, Nonlocality of a Single Photon, Phys. Rev. Lett. **66**, 252 (1991).

[22] A. Gilchrist, P. Deuar, and M. D. Reid, Contradiction of Quantum Mechanics with Local Hidden Variables for Quadrature Phase Amplitude Measurements, Phys. Rev. Lett. **80**, 3169 (1998).

[23] W. J. Munro and G. J. Milburn, Characterizing Greenberger-Horne-Zeilinger Correlations in Nondegenerate Parametric Oscillation via Phase Measurements, Phys. Rev. Lett. **81**, 4285 (1998).

[24] S.-W. Lee and D. Jaksch, Maximal violation of tight Bell inequalities for maximal high-dimensional entanglement, Phys. Rev. A **80**, 010103(R) (2009).

[25] E. G. Cavalcanti, C. J. Foster, M. D. Reid, and P. D. Drummond, Bell Inequalities for Continuous-Variable Correlations, Phys. Rev. Lett. **99**, 210405 (2007).

[26] E. Shchukin and W. Vogel, Quaternions, octonions, and Bell-type inequalities, Phys. Rev. A **78**, 032104 (2008).

[27] R. García-Patrón, J. Fiurášek, N. J. Cerf, J. Wenger, R. Tualle-Brouri, and Ph. Grangier, Proposal for a Loophole-Free Bell Test Using Homodyne Detection, Phys. Rev. Lett. **93**, 130409 (2004).

[28] J. Wenger, M. Hafezi, F. Grosshans, R. Tualle-Brouri, and P. Grangier, Maximal violation of Bell inequalities using continuous-variable measurements, Phys. Rev. A **67**, 012105 (2003).

[29] A. Salles, D. Cavalcanti, and A. Acín, Quantum Nonlocality and Partial Transposition for Continuous-Variable Systems, Phys. Rev. Lett. **101**, 040404 (2008).

[30] A. Salles, D. Cavalcanti, A. Acin, D. Perez-Garcia, and M. M. Wolf, Bell inequalities from multilinear contractions, Quantum Inf. Comput. **10**, 0703 (2010).

[31] S. L. Braunstein and P. van Loock, Quantum information with continuous variables, Rev. Mod. Phys. **77**, 513 (2005).

[32] A. Ferraro, S. Olivares, and M. G. A. Paris, *Gaussian States in Continuous Variable Quantum Information* (Bibliopolis, Napoli, 2005).

[33] C. Weedbrook, S. Pirandola, R. García-Patron, N. J. Cerf, T. C. Ralph, J. H. Shapiro, and S. Lloyd, Gaussian quantum information, Rev. Mod. Phys. **84**, 621 (2012).

[34] M. Zukowski, D. Kaszlikowski, A. Baturo, and J.-A. Larsson, Strengthening the Bell Theorem: conditions to falsify local realism in an experiment, arXiv:quant-ph/9910058.

[35] D. Kaszlikowski, P. Gnaciński, M. Żukowski, W. Miklaszewski, and A. Zeilinger, Violations of Local Realism by Two Entangled N-Dimensional Systems Are Stronger than for Two Qubits, Phys. Rev. Lett. **85**, 4418 (2000).

[36] L. Masanes, Tight Bell inequality for d-outcome measurements correlations, Quantum Info. Comput. **3**, 345 (2002).

[37] P. Busch, P. J. Lahti, J. P. Pellonpää, and K. Ylinen, *Quantum Measurement* (Springer, Basel, 2016).

[38] S. Pironio, Lifting Bell inequalities, J. Math. Phys. **46**, 062112 (2005).

[39] B. Simon, *Convexity: An Analytic Viewpoint* (Cambridge University Press, 2011).

[40] E. Haapasalo, M. Sedlák, and M. Ziman, Distance to boundary and minimum-error discrimination, Phys. Rev. A **89**, 062303 (2014).

[41] M. Kleinmann and A. Cabello, Quantum Correlations Are Stronger Than All Nonsignalling Correlations Produced by n-Outcome Measurements, Phys. Rev. Lett. **117**, 150401 (2016).

[42] A. S. Holevo, Statistical definition of observable and the structure of statistical models, Rep. Math. Phys. **22**, 385 (1985).

[43] T. Heinosaari and J.-P. Pellonpää, Canonical phase measurement is pure, Phys. Rev. A **80**, 040101(R) (2009).

[44] J.-P. Pellonpää, Complete characterization of extreme quantum observables in infinite dimensions, J. Phys. A: Math. Theor. **44**, 085304 (2011).