Physica Scripta

PAPER

Quantum tomography and nonlocality

To cite this article: Evgeny V Shchukin and Stefano Mancini 2015 *Phys. Scr.* **90** 074019

IOP Publishing | Royal Swedish Academy of Sciences Physica Scripta
Phys. Scr. 90 (2015) 074019 (11pp) doi:10.1088/0031-8949/90/7/074019

Quantum tomography and nonlocality

Evgeny V Shchukin[1] **and Stefano Mancini**[2,3]

[1] Institute of Physics, Johannes-Gutenberg University of Mainz, Staudingerweg 7, 55128 Mainz, Germany
[2] School of Science and Technology, University of Camerino, 62032 Camerino, Italy & INFN Sezione di Perugia, I-06123 Perugia, Italy

E-mail: evgeny.shchukin@gmail.com and stefano.mancini@unicam.it

Received 28 October 2014, revised 4 January 2015
Accepted for publication 8 January 2015
Published 15 June 2015

Abstract

We present a tomographic approach to the study of quantum nonlocality in multipartite systems. Bell inequalities for tomograms belonging to a generic tomographic scheme are derived by exploiting tools from convex geometry. Then, possible violations of these inequalities are discussed in specific tomographic realizations providing some explicit examples.

Keywords: quantum tomography, entanglement and nonlocality, formalism

1. Introduction

The Bell inequalities [1] demonstrate paradigmatic difference of quantum and classical worlds. They were originally written for dichotomic (spin–$\frac{1}{2}$) variables [2]. Spin–$\frac{1}{2}$ operators realize the Lie algebra of the SU(2) group. For several spin particles their spin operators form the Lie algebra of the tensor product of the Lie algebras. Due to the algebraic equivalence of the operators satisfying commutation relations of the Lie algebra constructed from particle spin operators and constructed from creation and annihilation operators of a field, one can obtain Bell inequalities also for the case of continuous variables besides discrete ones [3]. Beyond the specific operators involved in the Bell inequalities, their possible violations obviously depend on the state under consideration.

For a (multipartite) classical system with fluctuations, the system state is described by means of a joint probability distribution function of random variables corresponding to the subsystems. In contrast, for a (multipartite) quantum system, the state is described by the density matrix. In view of this difference, the calculations of the system's statistical properties (including correlations) are accomplished differently in classical and quantum domains.

Recently, a probability representation of quantum mechanics has been suggested [4]. This representation, equivalent to all other well known formulations of quantum mechanics (see, e.g. [5]), goes back to *quantum tomography*, a technique used for quantum state reconstruction [6]. The approach makes use of a set of fair probabilities, *tomograms*, to 'replace' the notion of a quantum state. It has also been understood [7] that for classical statistical mechanics, the states with fluctuations can be described as well by tomograms related to standard probability distributions in classical phase-space. A comparison of classical and quantum tomograms can be found in [7, 8].

Thus, in the probability representation, tomograms turned out to be a unique tool to describe both classical and quantum states. As a consequence, they represent a natural bed where to place inequalities marking the border-line between quantum and classical worlds. Tomograms can be either continuous or discrete variable functions depending on the tomographic scheme (realization). In both cases they might be directly used to test nonlocality. This possibility was described for symplectic tomography [9] in a bipartite system [10] and spin tomography [11] still in a bipartite system [12].

Here we shall derive Bell inequalities for *multipartite* systems in terms of tomograms belonging to a *generic* tomographic scheme. Then, we shall discuss the possibility of violating such inequalities depending on the tomographic realization.

The layout of the paper is the following. In section 2 we formalize quantum tomography in a multipartite setting. Then, in section 3 we derive the Bell inequalities in terms of tomograms. In section 4 we provide some evidence of violations of such inequalities for spin–$\frac{1}{2}$ systems as well as for field modes and finally draw the conclusions in section 5.

[3] Author to whom any correspondence should be addressed.

2. Quantum tomography

Here we briefly review the general quantum tomography approach for a single system, by detailing three relevant cases (optical [13], spin [11] and photon-number tomography [14]) and then extend the formalism to multipartite systems.

The basic ingredients of any tomographic scheme are a Hilbert space $\mathcal{H}$ associated with the space of the system under consideration and a pair of measurable sets (X, Λ) with measures $\mu(x)$ and $\nu(\lambda)$ correspondingly. More precisely, the set of system states is the set $\mathcal{S}(\mathcal{H})$ of Hermitian non-negative trace-class operators on $\mathcal{H}$ with trace 1. Usually the set X is the spectrum of an observable of the system and the set Λ plays the role of transformations.

We use the notation $\mathcal{P}(X)$ for the set of probability distributions on X, i.e. the set of nonnegative measurable functions $p : X \to \mathbb{R}$ normalized to one in the following sense $\int p(x)\mathrm{d}\mu(x) = 1$.

Both sets $\mathcal{S}(\mathcal{H})$ and $\mathcal{P}(X)$ are closed with respect to the convex combinations: if $\hat{\varrho}, \hat{\sigma} \in \mathcal{S}(\mathcal{H})$ (resp. $p(x), q(x) \in \mathcal{P}(X)$) and $a \in [0, 1]$ then

$$a\hat{\varrho} + (1-a)\hat{\sigma} \in \mathcal{S}(\mathcal{H}) \quad (\text{resp. } ap(x) + (1-a)q(x) \in \mathcal{P}(X)).$$

Definition 1. A map $\mathcal{T} : \mathcal{S}(\mathcal{H}) \to \mathbb{R}^{X\times\Lambda}$ is called a tomographic map if the following three conditions are satisfied:

(i) for any $\hat{\varrho} \in \mathcal{S}(\mathcal{H})$ the image $\mathcal{T}(\hat{\varrho}) : X \times \Lambda \to \mathbb{R}$ restricted on the set $X \times \{\lambda\}$ is a probability density on X

$$\mathcal{T}_\lambda(\hat{\varrho}) \in \mathcal{P}(X) \quad \forall \lambda \in \Lambda, \quad \text{where}$$
$$\mathcal{T}_\lambda(\hat{\varrho}) = \mathcal{T}(\hat{\varrho})\big|_{X\times\{\lambda\}} : X \to \mathbb{R}.$$

(ii) the map $\mathcal{T}$ preserves convex combinations

$$\mathcal{T}(a\hat{\varrho} + (1-a)\hat{\sigma}) = a\mathcal{T}(\hat{\varrho}) + (1-a)\mathcal{T}(\hat{\sigma}),$$
$$\forall \hat{\varrho}, \hat{\sigma} \in \mathcal{S}(\mathcal{H}), a \in [0, 1].$$

(iii) the map $\mathcal{T}$ is one-to-one

$$\mathcal{T}(\hat{\varrho}) = \mathcal{T}(\hat{\sigma}) \Leftrightarrow \hat{\varrho} = \hat{\sigma}.$$

These conditions have simple meanings: (i) means that the tomogram $\mathcal{T}(\hat{\varrho})$ of any state $\hat{\varrho}$ is a probability distribution on X parameterized by the points of Λ, (ii) is the linearity condition, and (iii) requires that the tomogram of each state be unique, or, in other words, that any state can be unambiguously reconstructed from its tomogram.

In the present work we deal with tomographic maps of the following form

$$\mathcal{T}(\hat{\varrho})(x, \lambda) \equiv p_{\hat{\varrho}}(x, \lambda) = \mathrm{Tr}\big(\hat{\varrho}\hat{U}(x, \lambda)\big), \tag{1}$$

where $\hat{U}(x, \lambda)$ is a family of operators on $\mathcal{H}$ parameterized by points (x, λ) of the set $X \times \Lambda$. In the examples considered below the state $\hat{\varrho}$ can be reconstructed from its tomogram $p_{\hat{\varrho}}(x, \lambda)$ according to the formula

$$\hat{\varrho} = \iint_{X\times\Lambda} p(x, \lambda)\hat{D}(x, \lambda)\mathrm{d}\mu(x)\mathrm{d}\nu(\lambda), \tag{2}$$

for the appropriate (x, λ)-parameterized family of operators $\hat{D}(x, \lambda)$ on $\mathcal{H}$.

The set X is the spectrum of an observable $\hat{O}$ and the set Λ is a group equipped with a representation (in general projective) $\pi : \Lambda \to \mathcal{H}$ in $\mathcal{H}$. The operators $\hat{U}(x, \lambda)$ have the following form

$$\hat{U}(x, \lambda) = \pi(\lambda)|x\rangle\langle x|\,\pi^\dagger(\lambda), \tag{3}$$

where $|x\rangle$ is an eigenstate of the observable $\hat{O}$. For a group theoretical approach to quantum tomography see [15]. See also [16] for a relation to groupoids.

2.1. Spin tomography

Let us consider a system with spin j. In this case we have: $\mathcal{H} = \mathbb{C}^{2j+1}$, $X = \{-j, -j+1, \ldots, j-1, j\}$ and $\Lambda = \mathrm{SO}(3, \mathbb{R})$. We denote the elements of the sets X and Λ as s and Ω, respectively. The measure on X is equal to one on each element, so the corresponding integral is simply the finite sum over $2j+1$ terms. The measure on $\mathrm{SO}(3, \mathbb{R})$ is Haar's one. For the group $\mathrm{SO}(3, \mathbb{R})$, parameterized with Euler angles $\Omega \equiv (\varphi, \psi, \theta)$ the measure $\nu(\Omega)$ reads $\nu(\Omega) \equiv \nu(\varphi, \psi, \theta) = \sin\psi\, \mathrm{d}\varphi\, \mathrm{d}\psi\, \mathrm{d}\theta$ and the operator $\hat{U}$ of (3) takes the form

$$\hat{U}(s, \Omega) = \hat{K}(\Omega)|j, s\rangle\langle j, s|\,\hat{K}^\dagger(\Omega). \tag{4}$$

Here the vectors $|j, s\rangle$, $s = -j, -j+1, \ldots, j-1, j$ are the basis of the space $\mathbb{C}^{2j+1}$ (eigenvectors of the spin projection $\hat{s}_z$) and the operators $\hat{K}(\Omega)$ are the operators of the irreducible representation of $\mathrm{SO}(3, \mathbb{R})$ in $\mathbb{C}^{2j+1}$. Their matrix elements are given by

$$\langle j, s|\,\hat{K}(\Omega)|j, s'\rangle = \mathrm{e}^{\mathrm{i}(s\theta+s'\varphi)}\sqrt{\frac{(j+s')!(j-s')!}{(j+s)!(j-s)!}} \times \cos^{s+s'}(\psi/2)\sin^{s'-s}(\psi/2)P_{j-s'}^{(s'-s,s'+s)}(\cos\psi), \tag{5}$$

with $P_n^{(\alpha,\beta)}(x)$ the Jacobi polynomials.

Then the tomogram $p(s, \Omega) \equiv p(s, \varphi, \psi, \theta)$ of (1) is

$$p(s, \Omega) = \langle j, s|\,\hat{K}(\Omega)\hat{\varrho}\hat{K}^\dagger(\Omega)|j, s\rangle. \tag{6}$$

Due to the property $\langle j, s|\,\hat{K}(\Omega)|j, s'\rangle = (-1)^{s'-s}\langle j, -s|\,\hat{K}(\Omega)|j, -s'\rangle$, the tomogram does not depend on the angle θ, i.e. $p(s, \varphi, \psi, \theta) \equiv p(s, \varphi, \psi)$.

Finally, the operator $\hat{D}$ of (2) results

$$\hat{D}(s, \Omega) = \sum_{n,m=-j}^{j} \langle j, n|\,\hat{D}(s, \Omega)|j, m\rangle|j, n\rangle\langle j, m|,$$

where the matrix elements $\langle j, n|\,\hat{D}(s, \Omega)|j, m\rangle$ are given by

the following expression

$$\langle j, n|\,\hat{D}(s,\Omega)\,|j, m\rangle = \frac{(-1)^{s+m}}{8\pi^2}\sum_{j_3=0}^{2j}\left(2j_3+1\right)^2 \times \sum_{k=-j_3}^{j_3}\langle j, k|\,\hat{K}(\Omega)\,|j, 0\rangle \times \begin{pmatrix} j & j & j_3 \\ n & -m & k\end{pmatrix}\begin{pmatrix} j & j & j_3 \\ s & -s & k\end{pmatrix}, \tag{7}$$

in terms of Wigner $3j$-symbols.

2.2. Optical tomography

Here we have : $\mathcal{H} = L_2(\mathbb{R})$, $X = \mathbb{R}$ and $\Lambda = \{e^{i\theta}|\theta \in [0, 2\pi]\}$. The measures on X and Λ are Lebegue's ones. The operator corresponding to equation (3) reads

$$\hat{U}(X,\theta) = \hat{R}(\theta)|X\rangle\langle X|\,\hat{R}^{\dagger}(\theta), \tag{8}$$

where $\hat{R}(\theta)$ is the rotation operator

$$\hat{R}(\theta) = \exp\left(i\frac{\theta}{2}\left(\hat{x}^2+\hat{p}^2\right)\right),$$

acting on and the canonical position $\hat{x}$ and momentum $\hat{p}$ operators as

$$\hat{R}(\theta)\begin{pmatrix}\hat{x}\\ \hat{p}\end{pmatrix}\hat{R}^{\dagger}(\theta) = \begin{pmatrix}\cos\theta & -\sin\theta\\ \sin\theta & \cos\theta\end{pmatrix}\begin{pmatrix}\hat{x}\\ \hat{p}\end{pmatrix}.$$

In other words, $\hat{U}(X,\theta)$ of (8) is the projector on the rotated eigenvector $|X\rangle$ of the position operator $\hat{x}$. The tomogram $p(X,\theta)$ of (1) is the diagonal matrix element

$$p(X,\theta) = \langle X|\,\hat{R}(\theta)\hat{\varrho}\hat{R}^{\dagger}(\theta)|X\rangle. \tag{9}$$

Furthermore, the operator $\hat{D}$ of (2) results

$$\hat{D}(X,\theta) = \frac{1}{4\pi}\int |r|\,\exp\left(-ir\left(X-\cos\theta\hat{x}-\sin\theta\hat{p}\right)\right)dr.$$

2.3. Photon-Number tomography

Here we have: $\mathcal{H} = L_2(\mathbb{R})$, $X = \mathbb{Z}_+ = \{0, 1, \ldots\}$ and $\Lambda = \mathbb{C}$. We denote the elements of the sets X and $\mathbb{C}$ as n and α, respectively. The measure on X is equal to one on each element and the measure on $\mathbb{C}$ is $(1/\pi)d^2\alpha$, where $d^2\alpha = d\mathrm{Re}\,\alpha\, d\mathrm{Im}\,\alpha$ is Lebegue's measure on the real plane. Here, the operator $\hat{U}$ is the projector onto the displaced Fock state

$$\hat{U}(n,\alpha) = \hat{D}(\alpha)|n\rangle\langle n|\,\hat{D}^{\dagger}(\alpha), \tag{10}$$

with

$$D(\alpha) \equiv \exp\left[\frac{\alpha-\alpha^*}{\sqrt{2}}\hat{x} - i\frac{\alpha+\alpha^*}{\sqrt{2}}\hat{p}\right].$$

From (1) the tomogram $p(n,\alpha)$ reads

$$p(n,\alpha) = \langle n|\hat{D}(\alpha)\hat{\varrho}\hat{D}^{\dagger}(\alpha)|n\rangle. \tag{11}$$

Furthermore, the operator $\hat{D}$ of (2) becomes in this case

$$\hat{D}(n,\alpha) = 4(-1)^n\sum_{m=0}^{+\infty}(-1)^m\hat{D}(\alpha)|m\rangle\langle m|\,\hat{D}^{\dagger}(\alpha).$$

2.4. Tomography for multi-partite systems

The generalization for multi-partite systems is straightforward.

Definition 2. Consider an n-partite system with the state space $\mathcal{H}^{\otimes n}$ and n tomographic schemes, one for each part with sets (X_k, Λ_k) and operators $\hat{U}_k(x_k,\lambda_k)$ and $\hat{D}_k(x_k,\lambda_k)$, $k = 1, \ldots, n$. The tomographic scheme for the whole system is then constructed as the direct product of these schemes, by using

$$X \equiv \prod_{k=1}^{n} X_k, \quad \Lambda \equiv \prod_{k=1}^{n}\Lambda_k,$$
$$\hat{U}(\boldsymbol{x},\lambda) \equiv \bigotimes_{k=1}^{n}\hat{U}_k\left(x_k,\lambda_k\right),$$
$$\hat{D}(\boldsymbol{x},\lambda) \equiv \bigotimes_{k=1}^{n}\hat{D}_k\left(x_k,\lambda_k\right), \tag{12}$$

where $\boldsymbol{x} \equiv (x_1, \ldots, x_n)$, $\boldsymbol{\lambda} \equiv (\lambda_1, \ldots, \lambda_n)$ and the measures $\mu(\boldsymbol{x})$, $\nu(\boldsymbol{\lambda})$ on X, Λ are direct products of $\mu_1(x_1), \ldots, \mu_n(x_n)$ and $\nu_1(\lambda_1), \ldots, \nu_n(\lambda_n)$, respectively. The tomogram $p(\boldsymbol{x},\boldsymbol{\lambda})$ of a state $\hat{\varrho}$ (generalizing (1)) is

$$p(\boldsymbol{x},\boldsymbol{\lambda}) = \mathrm{Tr}\left(\hat{\varrho}\hat{U}(\boldsymbol{x},\boldsymbol{\lambda})\right). \tag{13}$$

For any $\boldsymbol{\lambda} \in \Lambda$ it is a probability distribution on X, thus $\int_X p(\boldsymbol{x},\boldsymbol{\lambda})d\mu(\boldsymbol{x}) = 1$.

Remark. From the definition (12) of the operator $\hat{U}(\boldsymbol{x},\boldsymbol{\lambda})$ it immediately follows that the tomogram $p(\boldsymbol{x},\boldsymbol{\lambda})$ of a factorized state

$$\hat{\varrho} = \hat{\varrho}_1 \otimes \ldots \otimes \hat{\varrho}_n \tag{14}$$

is also factorized, i.e.

$$p(\boldsymbol{x},\boldsymbol{\lambda}) = p_1\left(x_1,\lambda_1\right)\ldots p_n\left(x_n,\lambda_n\right), \tag{15}$$

where $p_k(x_k,\lambda_k)$ is the tomogram of the state $\hat{\varrho}_k$. More generally, the tomogram of a separable state

$$\hat{\varrho} = \sum_{i=0}^{+\infty}a_i\hat{\varrho}_1^{(i)} \otimes \ldots \hat{\varrho}_n^{(i)}, \qquad a_i \geqslant 0, \quad \sum_{i=0}^{+\infty}a_i = 1 \tag{16}$$

is also separable in the following sense

$$p(\boldsymbol{x},\boldsymbol{\lambda}) = \sum_{i=0}^{+\infty}a_i p_1^{(i)}\left(x_1,\lambda_1\right)\ldots p_n^{(i)}\left(x_n,\lambda_n\right), \tag{17}$$

where $p_k^{(i)}(x_k,\lambda_k)$ is the tomogram of the state $\hat{\varrho}_k^{(i)}$.

3. Bell inequalities for tomograms

Let us consider an n-partite system in tomographic representation, with each subsystem supplied by a tomographic map $\mathcal{T}_k$, $k = 1,\ldots, n$. The tomogram $p(\boldsymbol{x}, \boldsymbol{\lambda})$ of a state $\hat{\varrho}$ is a function of $2n$ arguments and with respect to one half of them, it is a probability distribution. We will show that in general it cannot be considered as a classical joint probability.

Definition 3. For any $k = 1, \ldots, n$ let Y_k and Z_k be two measurable sets such that

$$X_k = Y_k \bigcup Z_k, \quad Y_k \bigcap Z_k = \varnothing,$$

and for any $\lambda_k \in \Lambda_k$ let $A_k(\lambda_k)$ be a dichotomic random variable on $X = \prod_k X_k$ such that

$$\mathbf{P}\left(A_k(\lambda_k) = 1\right) = \int_{Y_k} \mathrm{Tr}\left(\hat{\varrho}\hat{U}_k(x_k, \lambda_k)\right)\mathrm{d}\mu_k(x_k),$$
$$\mathbf{P}\left(A_k(\lambda_k) = -1\right) = \int_{Z_k} \mathrm{Tr}\left(\hat{\varrho}\hat{U}_k(x_k, \lambda_k)\right)\mathrm{d}\mu_k(x_k). \tag{18}$$

Symbolically the variables $A_k(\lambda_k)$ can be written as

$$A_k(\lambda_k) = \begin{cases} 1 & \text{if } x_k \in Y_k, \\ -1 & \text{if } x_k \in Z_k \end{cases} \tag{19}$$

in the coordinate system deformed by the operator $\hat{U}_k(x_k, \lambda_k)$. The joint probability distribution of the random variables $A_1(\lambda_1), \ldots, A_n(\lambda_n)$, namely

$$p_{\varepsilon_1,\ldots,\varepsilon_n}(\lambda_1, \ldots, \lambda_n) = \mathbf{P}\left(A_1(\lambda_1) = \varepsilon_1, \ldots, A_n(\lambda_n) = \varepsilon_n\right),$$

where $\varepsilon_k = \pm 1$, is given by

$$p_{\varepsilon_1,\ldots,\varepsilon_n}(\lambda_1, \ldots, \lambda_n) = \int_{W_1} \ldots \int_{W_n} p(\boldsymbol{x}, \boldsymbol{\lambda})\mathrm{d}\boldsymbol{x}, \tag{20}$$

with

$$W_k = \begin{cases} Y_k & \text{if } \varepsilon_k = 1, \\ Z_k & \text{if } \varepsilon_k = -1. \end{cases}$$

The correlation function of $A_1(\lambda_1), \ldots, A_n(\lambda_n)$ results

$$E(\lambda_1, \ldots, \lambda_n) \equiv \left\langle A_1(\lambda_1)\ldots A_n(\lambda_n)\right\rangle = \sum_{\varepsilon_1,\ldots,\varepsilon_n=\pm 1} p_{\varepsilon_1,\ldots,\varepsilon_n}(\lambda_1, \ldots, \lambda_n)\varepsilon_1\ldots\varepsilon_n. \tag{21}$$

Definition 4. Let us fix two parameters $\lambda_k^{(1)}$ and $\lambda_k^{(2)}$ for $k = 1, \ldots, n$ and denote

$$E(j_1, \ldots, j_n) \equiv E\left(\lambda_1^{(j_1)}, \ldots, \lambda_n^{(j_n)}\right), \quad j_k = 1, 2. \tag{22}$$

Since each index j_k can take 2 values independently on all the other indices, there are 2^n correlation functions (22). Then we define by

$$\boldsymbol{e} \equiv \left(E(j_1, \ldots, j_n)\right) \in \mathbb{R}^{2^n} \tag{23}$$

the vector of these correlation functions with some order of multi-indices $(j_1, \ldots, j_n)$.

It is convenient to enumerate the functions $E(j_1, \ldots, j_n)$. For this purpose, we use the binary base with 'digits' 1 and 2 instead of 0 and 1. This means that we use the following one-to-one correspondence

$$\left\{1, \ldots, 2^n\right\} \ni j \leftrightarrow (j_1, \ldots, j_n), \quad j_k = 1, 2,$$

where j and $(j_1, \ldots, j_n)$ are related to each other according to

$$j = (j_1 - 1)2^{n-1} + \ldots + (j_n - 1)2 + j_n. \tag{24}$$

By virtue of such an ordering, the vector $\boldsymbol{e}$ (23) can be written as

$$\boldsymbol{e} = \left(E(1), \ldots, E(2^n)\right) = (E(1, \ldots, 1), \ldots, E(2, \ldots, 2)) \in \mathbb{R}^{2^n}. \tag{25}$$

What region $\Omega_n \subset \mathbb{R}^{2^n}$ fills the vector $\boldsymbol{e}$ of (25)? Due to the fact that each observable has only two outcomes ± 1, it follows that each correlation function (22) is bounded by one, by an absolute value and, so, the set Ω_n is a subset of 2^n-dimensional cube $[-1, 1]^{2^n}$.

Suppose that it is possible to model the result of the measurement by a random variable, $A_k(j_k)$, which can take two values ± 1. We assume that these random variables can be arbitrary correlated.

Definition 5. Let us define by

$$p\left(i_1(1), \ldots, i_n(2)\right) \equiv \mathbf{P} \times \left(A_1(1) = i_1(1), \ldots, A_n(2) = i_n(2)\right), \tag{26}$$

the joint probability distribution for random variables $A_k(j_k)$, with $i_k(j_k) = \pm 1$. Since each index $i_k(j_k)$ can take independently 2 values, we have 2^{2n} numbers (26) which completely describe statistical characteristics of the random variables under consideration. We enumerate them with a single number $i = 1, \ldots, 2^{2n}$ using the same rule as for the correlation functions $E(j)$, namely

$$\left\{1, \ldots, 2^{2n}\right\} \ni i \leftrightarrow \left(i_1(1), \ldots, i_n(2)\right),$$

where i and $(i_1(1), \ldots, i_n(2))$ are related to each other according to

$$i = \left(i_1(1) - 1\right)2^{2n-1} + \left(i_n(1) - 1\right)2 + i_n(2). \tag{27}$$

Enumerated in such a way the probabilities (26) form a 2^{2n}-dimensional vector

$$\boldsymbol{p} \equiv (p_1, \ldots, p_{2^{2n}}) \in \mathbb{R}^{2^{2n}}. \tag{28}$$

The point (28) lies in the standard simplex

$$S_{2^{2n}-1} = \left\{(p_1, \ldots, p_{2^{2n}}) \,\middle|\, \sum_{i=1}^{2^{2n}} p_i = 1, p_i \geqslant 0\right\} \subset \mathbb{R}^{2^{2n}}. \tag{29}$$

What region $\Omega_n \subset \mathbb{R}^{2^n}$ fills the vector $\boldsymbol{e}$ (25) when the point $\boldsymbol{p}$ (28) runs over the simplex $S_{2^{2n}-1}$ (29)? To answer this question, we are going to explicitly relate $\boldsymbol{e}$ and $\boldsymbol{p}$ assuming the former, expressed like classical joint probabilities. Then,

the correlation function $E(j)$ is intended as a simple linear combination of p_i with proper coefficients. Looking at (21) we consider such coefficients, $\mathcal{E}(j, i)$, given by the product

$$\mathcal{E}(j, i) = i_1(j_1) \dots i_n(j_n), \tag{30}$$

where j_k and $i_k(j_k)$ are 'digits' of the numbers j and i in the binary representations (24) and (27). The numbers $\mathcal{E}(j, i)$, $j = 1, \dots, 2^n$, $i = 1, \dots, 2^{2n}$ form a $2^n \times 2^{2n}$ matrix $\mathcal{E}_n$ and the relation between $\boldsymbol{e}$ and $\boldsymbol{p}$ can then be written as

$$\boldsymbol{e} = \mathcal{E}_n \boldsymbol{p}. \tag{31}$$

We see that the region Ω_n is the image of the standard simplex $S_{2^{2n}-1}$

$$\Omega_n = \mathcal{E}_n\left(S_{2^{2n}-1}\right), \tag{32}$$

where we do not distinguish the linear map $\mathcal{E}_n$: $\mathbb{R}^{2^{2n}} \to \mathbb{R}^{2^n}$ and its matrix $\mathcal{E}_n$ in the standard bases of $\mathbb{R}^{2^{2n}}$ and $\mathbb{R}^{2^n}$.

Thus, we have reduced the problem of finding Bell inequalities to find the set Ω_n. It means that the problem of finding Bell inequalities boils down to a standard problem of convex geometry, referred to as the convex hull problem: given points $\boldsymbol{c}_i$ find their convex hull, or facets of maximal dimension of the corresponding polytope (for notions of convex geometry see, e.g. [17]).

Now we will get the Bell inequalities explicitly. Note that permutations of the columns of the matrix $\mathcal{E}_n$ do not change their convex hull and that they correspond to permutations of the components of the vector $\boldsymbol{p}$ or different orderings of the probabilities (28), so one can safely permute columns of $\mathcal{E}_n$ without altering (31).

Theorem 1. *The set Ω_n is specified by the (Bell) inequalities for the vector of the correlation functions*

$$\left(\boldsymbol{e}, H_{2^n}\boldsymbol{c}\right) \leqslant 2^n, \quad \forall \boldsymbol{c} = (\pm 1, \dots, \pm 1). \tag{33}$$

The matrix H_2^n is the Hadamard matrix recurrently defined as

$$H_{2^n} = \underbrace{H_2 \otimes \dots \otimes H_2}_{n}, \quad H_2 = \begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix}.$$

Proof. The key fact in deriving the Bell inequality (33) is that the matrix $\mathcal{E}_n$ can be written in the following block form

$$\mathcal{E}_n = \left(\underbrace{H_{2^n} \quad -H_{2^n} \quad \dots \quad H_{2^n} \quad -H_{2^n}}_{2^n} \right) \tag{34}$$

after appropriate arrangement of its columns. One can rewrite the r.h.s. of (34) as the product of two matrices

$$\mathcal{E}_n = H_{2^n}\left(E_{2^n} \quad -E_{2^n} \quad \dots \quad E_{2^n} \quad -E_{2^n}\right) = H_{2^n} A_n,$$

which means that the linear map $\mathcal{E}_n : \mathbb{R}^{2^{2n}} \to \mathbb{R}^{2^n}$ can be decomposed into two maps

$$\mathcal{E}_n = H_{2^n} \circ A_n, \quad A_n : \mathbb{R}^{2^{2n}} \to \mathbb{R}^{2^n}, \quad H_{2^n} : \mathbb{R}^{2^n} \to \mathbb{R}^{2^n}.$$

According to this decomposition (31) reads

$$\boldsymbol{e} = H_{2^n}\boldsymbol{q}, \tag{35}$$

where the vector $\boldsymbol{q} = A_n \boldsymbol{p} \in \mathbb{R}^{2^n}$ is explicitly given by the following expression

$$\boldsymbol{q} = \begin{pmatrix} p_1 - p_{2^n+1} + \dots - p_{(2^n-1)2^n+1} \\ \vdots \\ p_{2^n} - p_{2\cdot 2^n} + \dots - p_{2^{2n}} \end{pmatrix}. \tag{36}$$

Define the following convex polytope $\mathcal{O}_N \subset \mathbb{R}^N$

$$\mathcal{O}_N = \left\{ \boldsymbol{x} \in \mathbb{R}^N \middle| (\boldsymbol{x}, \boldsymbol{c}) \leqslant 1, \ \forall \boldsymbol{c} = (\pm 1, \dots, \pm 1) \right\}. \tag{37}$$

As one can easily see the image of the standard simplex $S_{2^{2n}-1}$ is exactly the polytope $\mathcal{O}_{2^n}$, that is $A_n(S_{2^{2n}-1}) = \mathcal{O}_{2^n}$. From this fact, we have

$$\Omega_n = H_{2^n}\left(\mathcal{O}_{2^n}\right). \tag{38}$$

Now the Bell inequalities can be straightforwardly obtained from this relation. Just notice that a non-degenerate linear map $f : \mathbb{R}^N \to \mathbb{R}^N$ with the matrix F maps a half-space $\mathfrak{h} = \{\boldsymbol{x} \in \mathbb{R}^N | (\boldsymbol{x}, \boldsymbol{a}) \leqslant b\}$ to the half-space $f(\mathfrak{h}) = \{\boldsymbol{y} \in \mathbb{R}^N | (\boldsymbol{y}, (F^T)^{-1}\boldsymbol{a}) \leqslant b\}$. Taking into account the following representation of the polytope (38)

$$\mathcal{O}_{2^n} = \bigcap_{\boldsymbol{c}=(\pm 1, \dots, \pm 1)} \{\boldsymbol{q} | (\boldsymbol{q}, \boldsymbol{c}) \leqslant 1\}, \tag{39}$$

the symmetry of the Hadamard matrix H_{2^n} and the formula for its inverse $H_{2^n}^{-1} = \frac{1}{2^n} H_{2^n}$, we get the explicit form of the set Ω_n, i.e.

$$\Omega_n = \bigcap_{\boldsymbol{c}=(\pm 1, \dots, \pm 1)} \left\{ \boldsymbol{e} \middle| \left(\boldsymbol{e}, H_{2^n}\boldsymbol{c}\right) \leqslant 2^n \right\}. \tag{40}$$

Hence, the Bell inequalities (33). □

Remark. Explicitly (33) can be written as

$$\left| \sum_{j_1, \dots, j_n = 1}^{2} a_{j_1, \dots, j_n} E\left(j_1, \dots, j_n\right) \right| \leqslant 2^n, \tag{41}$$

where the coefficients $a_{j_1, \dots, j_n}$ are connected with the vector $\boldsymbol{c}$ by the following relation

$$a_{j_1, \dots, j_n} = \sum_{\varepsilon_1, \dots, \varepsilon_n = \pm 1} c(\varepsilon_1, \dots, \varepsilon_n) \varepsilon_1^{j_1 - 1} \dots \varepsilon_n^{j_n - 1}. \tag{42}$$

The number $c(\varepsilon_1, \dots, \varepsilon_n)$ here is the ith component of the vector $\boldsymbol{c}$, where the binary representation of i is $i = (\varepsilon_1 \dots \varepsilon_n)_2$ with digits +1 and −1 instead of 0 and 1.

One can easily see that there are 2^{n+1} inequalities of the form $\pm E(j_1, \dots, j_n) \leqslant 1$. They correspond to the functions $c(\varepsilon_1, \dots, \varepsilon_n)$ that are columns of either H_{2^n} or $-H_{2^n}$ and they are referred to as trivial inequalities.

Finally, notice that the well known CHSH inequality [18] is a particular instance of (41) corresponding to $n = 2$ and

$$c(-1,-1) = -1,$$
$$c(-1,+1) = c(+1,-1) = c(+1,+1) = +1.$$

Theorem 2. *Any separable state satisfies (33) with correlation functions (22).*

Proof. Let us start with a factorized state (14) whose tomogram (15) is also factorized. Due to this, the random variables $A_1(\lambda_1), \ldots, A_n(\lambda_n)$ are independent and the correlation function $E(\lambda_1, \ldots 0, \lambda_n)$ reads

$$E(\lambda_1, \ldots, \lambda_n) = q_1(\lambda_1)\ldots q_n(\lambda_n) \tag{43}$$

with

$$q_k(\lambda_k) = p_1^{(k)}(\lambda_k) - p_{-1}^{(k)}(\lambda_k),$$

where

$$p_{\varepsilon_k}^{(k)}(\lambda_k) = \mathbf{P}\big(A_k(\lambda_k) = \varepsilon_k\big), \quad \varepsilon_k = \pm 1.$$

Due to the fact that

$$p_1^{(k)}(\lambda_k) + p_{-1}^{(k)}(\lambda_k) = 1, \quad \forall k = 1, \ldots, n \quad \forall \lambda_k \in \Lambda_k,$$

it is clear that $-1 \leqslant q_k(\lambda_k) \leqslant 1$. The left hand side of the inequality (33) is a linear function of any $q_k(\lambda_k^{(j_k)})$ where all the $q_1(\lambda_1^{(j_1)}), \ldots, q_n(\lambda_n^{(j_n)})$, $j_k = 1, 2$ are considered as independent variables. A linear function defined on the convex set $[-1, 1]$ takes its maximum on a boundary point, ± 1 in this case, and so, the left hand side of (33) is maximal if $q_k(\lambda_k^{(j_1)}) = \pm 1$, $j_k = 1, 2$, $k = 1, \ldots, n$. In such a case the vector $\boldsymbol{e}$ of correlation functions is a column of either H_{2^n} or $-H_{2^n}$. Just note that due to (43), the vector $\boldsymbol{e}$ reads

$$\boldsymbol{e} = \begin{pmatrix} q_1(1) \\ q_1(2) \end{pmatrix} \otimes \ldots \otimes \begin{pmatrix} q_n(1) \\ q_n(2) \end{pmatrix},$$

where $q_k(j_k) = q_k(\lambda_k^{(j_k)})$. That is to say, $\boldsymbol{e} = \boldsymbol{c}_i$ is the ith column of H_{2^n} or $-H_{2^n}$, then

$$\begin{aligned}(\boldsymbol{e}, H_{2^n}\boldsymbol{c}) &= \pm(\boldsymbol{c}_i, H_{2^n}\boldsymbol{c}) = \pm(H_{2^n}\boldsymbol{c}_i, \boldsymbol{c}) \\ &= \pm(2^n\boldsymbol{e}_i, \boldsymbol{c}) = \pm 2^n \leqslant 2^n.\end{aligned} \tag{44}$$

Here we used the orthogonality of the columns of H_{2^n}: $H_{2^n}\boldsymbol{c}_i = 2^n\boldsymbol{e}_i$, where all the coordinates of $\boldsymbol{e}_i$ are zero except the ith, which is one. Hence, we have proved that all factorized states satisfy (33).

Let us now consider a general separable state (16). Since the correlation function $E(\lambda_1, \ldots, \lambda_n)$ is a linear function of the state, the vector $\boldsymbol{e}$ is a linear combination of the vectors $\boldsymbol{e}^{(i)}$ corresponding to the states $\hat{\varrho}^{(i)} = \hat{\varrho}_1^{(i)} \otimes \ldots \otimes \hat{\varrho}_n^{(i)}$, i.e.

$$\boldsymbol{e} = \sum_{n=0}^{+\infty} a_i \boldsymbol{e}^{(i)}.$$

As we have already shown each vector $\boldsymbol{e}^{(i)}$ satisfies all the Bell inequalities (33) or lies in the convex set Ω_n. Once all the vectors $\boldsymbol{e}^{(i)}$ are in Ω_n, so is their convex combination $\boldsymbol{e}$. This means that any separable state satisfies all the inequalities (33). □

4. Quantum violations

The Bell inequalities are of interest not because they are always valid but because they can be violated. One can ask if there was a mistake in the proof of theorem 1. The problem relies on the underlying hypothesis of locality when relating $\boldsymbol{e}$ with $\boldsymbol{p}$ in (31). In doing so we have implicitly assumed (21) as a classical joint probability, which is not generally true at the quantum level.

We follow Mermin [19] to derive the only Bell inequality whose maximal quantum violation is the largest among all others.

For an odd number n of systems, let us consider the following random variable

$$M_n = \mathrm{Im}\left[\prod_{k=1}^{n}\big(A_k(1) + iA_k(2)\big)\right]. \tag{45}$$

Since each A_k can only take values ± 1, each term in this product is equal to $\sqrt{2}$ by absolute value. Furthermore, since n is odd the whole product has the phase that is an integer multiplier of $\pi/4$. As a consequence, we have

$$|\langle M_n \rangle| \leqslant 2^{(n-1)/2}. \tag{46}$$

Explicitly, this inequality reads

$$\left|\sum_{(j_1,\ldots,j_n)\in J} (-1)^{\delta(j_1,\ldots,j_n)} E(j_1, \ldots, j_n)\right| \leqslant 2^{(n-1)/2}, \tag{47}$$

where the sum here runs over the set of multi-indices $(j_1, \ldots, j_n)$, which contain an odd number of 2

$$J = \Big\{(j_1, \ldots, j_n) \Big| \big|\{k \,|\, j_k = 2\}\big| = 2l + 1\Big\}$$

and

$$\delta(j_1, \ldots, j_n) = l, \quad \big|\{k \,|\, j_k = 2\}\big| = 2l + 1.$$

Multiplied by $2^{(n+1)/2}$ the inequality (47) takes the form (41) and it is easy to show that it is a Bell inequality, i.e. there is a vector $\boldsymbol{c}$ that gives the coefficients of (47) (multiplied by $2^{(n+1)/2}$) according to (42).

We now consider an even number n. Let us denote the expression (45) as $M_n(1, 2)$ and the similar expression with the variables $A_k(1)$ and $A_k(2)$ swapped as $M_n(2, 1)$. Consider

the following combination

$$\widetilde{M}_n = M_{n-1}(1,2)\big(A_n(1)+A_n(2)\big) + M_{n-1}(2,1)\big(A_n(1)-A_n(2)\big). \tag{48}$$

Since $M_{n-1}(1,2)$ is equal to $\pm 2^{n/2-1}$ and $A_n(j)=\pm 1$, we will have

$$|\langle \widetilde{M}_n\rangle| \leqslant 2^{n/2}. \tag{49}$$

Using the explicit form (47) for the odd number $n-1$, one can write (49) as

$$\left|\sum_{j_1,\ldots,j_n=1}^{2} (-1)^{\tilde{\delta}(j_1,\ldots,j_n)} E\left(j_1,\ldots,j_n\right)\right| \leqslant 2^{n/2}, \tag{50}$$

where

$$\tilde{\delta}\left(j_1,\ldots,j_n\right) = \begin{cases} 1 & \text{if } j_n=2 \text{ and } \left|\left\{k \,\middle|\, j_k=2\right\}\right| \text{ is nonzero and even} \\ 0 & \text{otherwise} \end{cases}.$$

One can see that it is a Bell inequality and multiplied by $2^{n/2}$ it takes the form (41). Furthermore, for $n=2$, equation (50) exactly reduces to the CHSH inequality [18].

Let us now see how the inequalities (47) or (50) can be violated in different tomographic realizations starting from the following entangled state

$$|\Psi\rangle = \frac{1}{\sqrt{2}}(|\mathbf{0}\rangle + |\mathbf{1}\rangle), \tag{51}$$

where $\mathbf{0}=(0,\ldots,0)$ and $\mathbf{1}=(1,\ldots,1)$. Below, for the sake of simplicity, the focus will mainly be to $n=2,3$.

4.1. Spin tomography

Using the notation $|0\rangle \equiv |-\frac{1}{2}\rangle$, $|1\rangle \equiv |+\frac{1}{2}\rangle$ for the spin projection along z, the state (51) becomes

$$|\Psi\rangle = \frac{1}{\sqrt{2}}\left(\left|-\frac{1}{2},\ldots,-\frac{1}{2}\right\rangle + \left|+\frac{1}{2},\ldots,+\frac{1}{2}\right\rangle\right),$$

whose tomogram, referring to equation (7), reads as

$$p\left(s_1,\ldots,s_n,\Omega_1,\ldots,\Omega_n\right) = \frac{1}{2}\left|\prod_{j=1}^{n}\left\langle s_j\middle|\hat{K}\left(\Omega_j\right)\middle|-\tfrac{1}{2}\right\rangle + \prod_{j=1}^{n}\left\langle s_j\middle|\hat{K}\left(\Omega_j\right)\middle|+\tfrac{1}{2}\right\rangle\right|^2. \tag{52}$$

For $n=2$, we immediately get

$$p\left(+\tfrac{1}{2},+\tfrac{1}{2},\Omega_1,\Omega_2\right) = p\left(-\tfrac{1}{2},-\tfrac{1}{2},\Omega_1,\Omega_2\right) = \frac{1}{4}\Big(1+\cos\psi_1\cos\psi_2 + \sin\psi_1\sin\psi_2\cos\left(\varphi_1+\varphi_2\right)\Big),$$

$$p\left(+\tfrac{1}{2},-\tfrac{1}{2},\Omega_1,\Omega_2\right) = p\left(-\tfrac{1}{2},+\tfrac{1}{2},\Omega_1,\Omega_2\right) = \frac{1}{4}\Big(1-\cos\psi_1\cos\psi_2 - \sin\psi_1\sin\psi_2\cos\left(\varphi_1+\varphi_2\right)\Big),$$

and the correlation function (21) becomes

$$E\left(\Omega_1,\Omega_2\right) = \cos\psi_1\cos\psi_2 + \sin\psi_1\sin\psi_2\cos\left(\varphi_1+\varphi_2\right). \tag{53}$$

The Bell inequality (50) reads in this case

$$\left|E\left(\Omega_1^{(1)},\Omega_2^{(1)}\right) + E\left(\Omega_1^{(1)},\Omega_2^{(2)}\right) + E\left(\Omega_1^{(2)},\Omega_2^{(1)}\right) - E\left(\Omega_1^{(2)},\Omega_2^{(2)}\right)\right| \leqslant 2, \tag{54}$$

for all $\Omega_k^{(j)} = (\varphi_k^{(j)},\psi_k^{(j)},\theta_k^{(j)})$, $j,k=1,2$. The maximum of the l.h.s. of (54) with (53) is $2\sqrt{2}$ and is attained by taking e.g. (the angles θ do not matter here)

$$\Omega_1^{(1)} = \left(\varphi_1,-\pi/8,0\right), \quad \Omega_2^{(1)} = \left(-\varphi_1,\pi/8,0\right),$$
$$\Omega_1^{(2)} = \left(\varphi_1,3\pi/8,0\right), \quad \Omega_2^{(2)} = \left(-\varphi_1,-3\pi/8,0\right).$$

In the case of $n=3$, from (52), we have (not to overload the notation, we omit the Ω's)

$$p\left(+\tfrac{1}{2},+\tfrac{1}{2},+\tfrac{1}{2}\right) = \frac{1}{8}\Big[1+\cos\psi_1\cos\psi_2 + \cos\psi_1\cos\psi_3+\cos\psi_2\cos\psi_3 - \sin\psi_1\sin\psi_2\sin\psi_3\cos\left(\varphi_1+\varphi_2+\varphi_3\right)\Big],$$

$$p\left(+\tfrac{1}{2},+\tfrac{1}{2},-\tfrac{1}{2}\right) = \frac{1}{8}\Big[1+\cos\psi_1\cos\psi_2-\cos\psi_1\cos\psi_3 - \cos\psi_2\cos\psi_3 + \sin\psi_1\sin\psi_2\sin\psi_3\cos\left(\varphi_1+\varphi_2+\varphi_3\right)\Big],$$

$$p\left(+\tfrac{1}{2},-\tfrac{1}{2},+\tfrac{1}{2}\right) = \frac{1}{8}\Big[1-\cos\psi_1\cos\psi_2 + \cos\psi_1\cos\psi_3-\cos\psi_2\cos\psi_3 + \sin\psi_1\sin\psi_2\sin\psi_3\cos\left(\varphi_1+\varphi_2+\varphi_3\right)\Big],$$

$$p\left(+\tfrac{1}{2},-\tfrac{1}{2},-\tfrac{1}{2}\right) = \frac{1}{8}\Big[1-\cos\psi_1\cos\psi_2 - \cos\psi_1\cos\psi_3+\cos\psi_2\cos\psi_3 - \sin\psi_1\sin\psi_2\sin\psi_3\cos\left(\varphi_1+\varphi_2+\varphi_3\right)\Big],$$

$$p\left(-\tfrac{1}{2},+\tfrac{1}{2},+\tfrac{1}{2}\right)=\frac{1}{8}\Big[1-\cos\psi_1\cos\psi_2 - \cos\psi_1\cos\psi_3+\cos\psi_2\cos\psi_3 + \sin\psi_1\sin\psi_2\sin\psi_3\cos\left(\varphi_1+\varphi_2+\varphi_3\right)\Big],$$

$$p\left(-\tfrac{1}{2},+\tfrac{1}{2},-\tfrac{1}{2}\right)=\frac{1}{8}\Big[1-\cos\psi_1\cos\psi_2 + \cos\psi_1\cos\psi_3-\cos\psi_2\cos\psi_3 - \sin\psi_1\sin\psi_2\sin\psi_3\cos\left(\varphi_1+\varphi_2+\varphi_3\right)\Big],$$

$$p\left(-\tfrac{1}{2},-\tfrac{1}{2},+\tfrac{1}{2}\right)=\frac{1}{8}\Big[1+\cos\psi_1\cos\psi_2 - \cos\psi_1\cos\psi_3-\cos\psi_2\cos\psi_3 - \sin\psi_1\sin\psi_2\sin\psi_3\cos\left(\varphi_1+\varphi_2+\varphi_3\right)\Big],$$

$$p\left(-\tfrac{1}{2},-\tfrac{1}{2},-\tfrac{1}{2}\right)=\frac{1}{8}\Big[1+\cos\psi_1\cos\psi_2 + \cos\psi_1\cos\psi_3+\cos\psi_2\cos\psi_3 + \sin\psi_1\sin\psi_2\sin\psi_3\cos\left(\varphi_1+\varphi_2+\varphi_3\right)\Big].$$

Thanks to these tomograms, the correlation function (21) results

$$E\left(\Omega_1,\Omega_2,\Omega_3\right)=-\sin\psi_1\sin\psi_2\sin\psi_3 \times\cos\left(\varphi_1+\varphi_2+\varphi_3\right). \tag{55}$$

Finally, the Bell inequality (47) in this case reads

$$\Big|E\left(\Omega_1^{(2)},\Omega_2^{(1)},\Omega_3^{(1)}\right)+E\left(\Omega_1^{(1)},\Omega_2^{(2)},\Omega_3^{(1)}\right) + E\left(\Omega_1^{(1)},\Omega_2^{(1)},\Omega_3^{(2)}\right)-E\left(\Omega_1^{(2)},\Omega_2^{(2)},\Omega_3^{(2)}\right)\Big|\leqslant 2. \tag{56}$$

Using (55) the maximum violation occurs when the l.h.s equals 4. This value can be attained by taking e.g. (again the angles θ do not matter here)

$$\psi_1^{(1)}=\psi_2^{(1)}=\psi_3^{(1)}=\pi/2,\qquad \varphi_1^{(1)}=\varphi_2^{(1)}=\varphi_3^{(1)}=5\pi/6,$$
$$\psi_1^{(2)}=\psi_2^{(2)}=\psi_3^{(2)}=\pi/2,\qquad \varphi_1^{(2)}=\varphi_2^{(2)}=\varphi_3^{(2)}=\pi/3.$$

4.2. Optical tomography

The tomogram of the state (51) accordingly to (9) is given by

$$p(\boldsymbol{X},\boldsymbol{\theta})=\frac{1}{2\sqrt{\pi^n}}\left[1+2^n\prod_{i=1}^{n}\left(X_i^2\right) + 2^{(n+2)/2}\prod_{i=1}^{n}\left(X_i\right)\cos\left(\theta_1+\ldots+\theta_n\right)\right] \times\exp\left[-\sum_{i=1}^{n}X_i^2\right], \tag{57}$$

where $\boldsymbol{X}=(X_1,\ldots,X_n)$ and $\boldsymbol{\theta}=(\theta_1,\ldots,\theta_n)$.

We take the sets Y_k and Z_k of definition 3 to be

$$Y_k=[x,+\infty),\ Z_k=(-\infty,x).$$

For such sets and tomogram (57), the correlation function (21) results

$$E(\boldsymbol{\theta})=2^{n-1}\left[\mathsf{a}_0^n(x)+\mathsf{a}_1^n(x)\right] + 2^n\mathsf{b}_0^n(x)\cos\left(\theta_1+\ldots+\theta_n\right), \tag{58}$$

where

$$\mathsf{a}_0(x)=-\frac{1}{2}\operatorname{erf}(x),$$
$$\mathsf{a}_1(x)=-\frac{1}{2}\operatorname{erf}(x)+\frac{1}{\sqrt{\pi}}xe^{-x^2},$$
$$\mathsf{b}_0(x)=\frac{1}{\sqrt{2\pi}}e^{-x^2}.$$

We have now to insert (58) into (47) or (50) to get an explicit version fo the Bell inequality. In doing so, we use a Lemma, reported in appendix, showing that the maximal value of

$$\sum_{j_1,\ldots,j_n=1}^{2}a_{j_1,\ldots,j_n}\cos\left(\theta_1^{(j_1)}+\ldots+\theta_n^{(j_n)}\right)$$

does not exceed $2^{n+(n-1)/2}$ and this value is attained with coefficients of (47) or (50). It then follows that the maximal value $f_n(x)$ of the l.h.s. of (47) and of (50) is

$$f_n(x)=\begin{cases}2^n\left|\mathsf{a}_0^n(x)+\mathsf{a}_1^n(x)\right|+2^{n+(n+1)/2}\left|\mathsf{b}_0^n(x)\right|, & n\text{ odd}\\ 2^n\left|\mathsf{a}_0^n(x)+\mathsf{a}_1^n(x)\right|+2^{n+n/2}\quad\left|\mathsf{b}_0^n(x)\right|, & n\text{ even}\end{cases}. \tag{59}$$

figure 1 illustrates the function $f_n(x)$ for $n=2,3$. A tiny violation of Bell inequality only occurs for $n=3$.

4.3. Photon-Number tomography

Considering the state (51) its number tomogram (11) can be computed as

$$p(m_1,\ldots m_n,\alpha_1,\ldots\alpha_n)=\prod_{i=1}^{n}\frac{|\alpha_i|^{2m_i-2}}{m_i!}\mathrm{e}^{-|\alpha_i|^2} \times\left|\prod_{i=1}^{n}\alpha_i+\prod_{i=1}^{n}\left(m_i-|\alpha_i|^2\right)\right|^2. \tag{60}$$

We further choose the sets of definition 3 as $Z_1=\ldots=Z_n=\{0\}$, $Y_1=\ldots=Y_n=\{1,2,3,\ldots\}$.

The corresponding correlation function (21) for $n=2$ is

$$E(\alpha_1,\alpha_2)=\mathrm{e}^{-|\alpha_1|^2-|\alpha_2|^2}\Big[2+4\Re(\alpha_1\alpha_2) + 2\,|\alpha_1|^2|\alpha_2|^2-\left(1+|\alpha_2|^2\right)\mathrm{e}^{|\alpha_1|^2} - \left(1+|\alpha_1|^2\right)\mathrm{e}^{|\alpha_2|^2}+\mathrm{e}^{|\alpha_1|^2+|\alpha_2|^2}\Big]. \tag{61}$$

Furthermore, the Bell inequality for the number tomogram

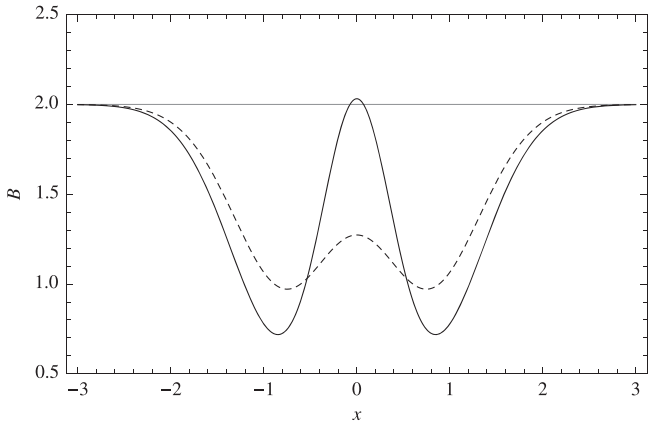

Figure 1. Function f_n of equation (59) versus x for $n = 2$ (dashed line) and $n = 3$ (solid line).

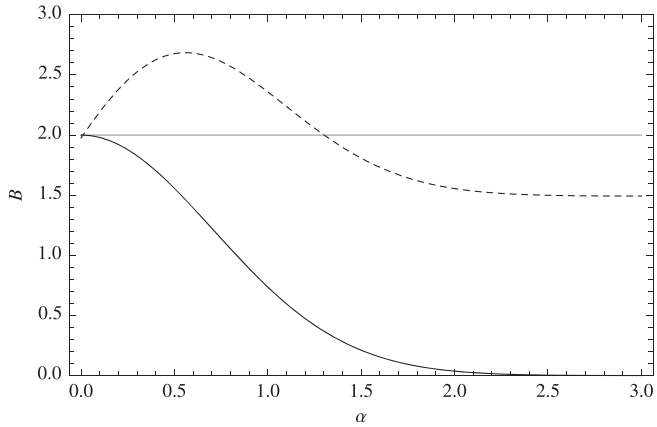

Figure 2. The left hand side of (62) as a function of $\alpha_2^{(2)}$ (solid line); the other parameters are given by $\alpha_1^{(1)} = 0.165$, $\alpha_2^{(1)} = -0.165$, $\alpha_1^{(2)} = -0.559$. The left hand side of (64) as a function of $\alpha_2^{(2)}$ (dashed line); the other parameters are given by $\alpha_1^{(1)} = \alpha_2^{(1)} = 0$, $\alpha_3^{(1)} = 5.936$, $\alpha_1^{(2)} 4.767$, $\alpha_3^{(2)} = 4$.

with $n = 2$ is from (50)

$$\left|E\left(\alpha_1^{(1)}, \alpha_2^{(1)}\right) + E\left(\alpha_1^{(1)}, \alpha_2^{(2)}\right) + E\left(\alpha_1^{(2)}, \alpha_2^{(1)}\right) - E\left(\alpha_1^{(2)}, \alpha_2^{(2)}\right)\right| \leqslant 2, \tag{62}$$

for all $\alpha_1^{(j)}, \alpha_2^{(j)} \in \mathbb{C}$, $j = 1, 2$. Figure 2 illustrates that this inequality can be violated using (61).

Analogously, from (60) it follows that the correlation function (21) for $n = 3$ is

$$\begin{aligned} E(\alpha_1, \alpha_2, \alpha_3) = \mathrm{e}^{-|\alpha_1|^2-|\alpha_2|^2-|\alpha_3|^2}[&-4 \\ &+ 8\Re(\alpha_1\alpha_2\alpha_3) - 4|\alpha_1|^2|\alpha_2|^2|\alpha_3|^2 \\ &+ 2\left(\mathrm{e}^{|\alpha_1|^2} + \mathrm{e}^{|\alpha_2|^2} + \mathrm{e}^{|\alpha_3|^2}\right) \\ &+ 2|\alpha_1|^2|\alpha_2|^2\mathrm{e}^{|\alpha_3|^2} + 2|\alpha_1|^2|\alpha_3|^2\mathrm{e}^{|\alpha_2|^2} \\ &+ 2|\alpha_2|^2|\alpha_3|^2\mathrm{e}^{|\alpha_1|^2} \\ &- \left(1 + |\alpha_1|^2\right)\mathrm{e}^{|\alpha_2|^2+|\alpha_3|^2} \\ &- \left(1 + |\alpha_2|^2\right)\mathrm{e}^{|\alpha_1|^2+|\alpha_3|^2} \\ &- \left(1 + |\alpha_3|^2\right)\mathrm{e}^{|\alpha_1|^2+|\alpha_2|^2} \\ &+ \mathrm{e}^{|\alpha_1|^2+|\alpha_2|^2+|\alpha_3|^2}\Big]. \end{aligned} \tag{63}$$

This time the Bell inequality for the number tomogram reads from (47)

$$\left|E\left(\alpha_1^{(1)}, \alpha_2^{(1)}, \alpha_3^{(2)}\right) + E\left(\alpha_1^{(1)}, \alpha_2^{(2)}, \alpha_2^{(1)}\right) + E\left(\alpha_1^{(2)}, \alpha_2^{(1)}, \alpha_2^{(1)}\right) - E\left(\alpha_1^{(2)}, \alpha_2^{(2)}, \alpha_2^{(2)}\right)\right| \leqslant 2. \tag{64}$$

This inequality, by numerical checking, results never violated with (63) and an example of the behavior of the l.h.s. is shown in figure 2.

By also choosing $Z_1 = \ldots = Z_n = \{0, \ldots, m\}$, $Y_1 = \ldots = Y_n = \{m + 1, m + 2, \ldots\}$, with $m > 0$, neither (62) nor (64) will result (by numerical checking) ever violated by using (61) and (63) respectively.

5. Concluding remarks

As we have seen from the previous examples, the use of a finite (namely 2^n) number of tomograms within a tomographic realization may lead to the evidence of nonlocality. Actually, it results that finite dimensional systems by means of spin tomograms allow for the best evidence of nonlocality. In contrast, violations of Bell inequalities seem much harder to uncover in infinite dimensional systems where $\mathcal{H} = L_2(\mathbb{R})$. Given that we have considered in both cases the same (entangled) state (51), this difference, according to [20], must be ascribed to the diversity of observables employed (from which the tomograms stem). However, we argue that also the way the spectrum of an observable is binned could play a role. As matter of fact, the choices made in sections 4.2 and 4.3 for Y_k and Z_k do not exhaust all possibilities of these measurable sets. Unfortunately, looking at Bell inequalities violations using optical tomograms (resp. photon number tomograms) by scanning the possible sets Y_k and Z_k appears a daunting task.

All in all, the advantage of the tomographic approach is to allow one to find the large violations of Bell inequalities typical of spin systems also in infinite dimensional systems. In fact, introducing in $L_2(\mathbb{R})^{\otimes n}$ the following local pseudo-spin operators [21]

$$\hat{S}_x^{(k)} = \sum_{n_k=0}^{+\infty} \left(|2n_k\rangle\langle 2n_k + 1| + |2n_k + 1\rangle\langle 2n_k|\right),$$

$$\hat{S}_y^{(k)} = -i \sum_{n_k=0}^{+\infty} \left(|2n_k\rangle\langle 2n_k + 1| - |2n_k + 1\rangle\langle 2n_k|\right),$$

$$\hat{S}_z^{(k)} = \sum_{n_k=0}^{+\infty} (-1)^{n_k}|n_k\rangle\langle n_k|,$$

where $|n_k\rangle$ are Fock states of the kth subsystem, we can derive the tomograms of the spin tomography realized with the above operators from those of any other tomographic scheme (see e.g. [22]). The price one ought to pay in such a case is the

completeness of the set of starting tomograms, (i.e. a number of tomograms much greater than 2^n).

Acknowledgments

This work was planned some years ago after an interesting discussion with V I Man'ko. We affectionately dedicate its completion to him on the occasion of his 75th birthday.

Appendix

Lemma 1. *For any coefficients $a_{\boldsymbol{j}} = a_{j_1,\ldots,j_n}$ ($\boldsymbol{j} = (j_1, \ldots, j_n)$) of (42) and for any angles $\theta_k^{(1)}$, $\theta_k^{(2)}$ ($k = 1, \ldots, n$) we have*

$$\frac{1}{2^n}\left|\sum_{j=1}^{2} a_j \cos\left(\theta_1^{(j_1)} + \ldots + \theta_n^{(j_n)}\right)\right| \leqslant 2^{(n-1)/2}. \tag{A.1}$$

The equality is attained with coefficients from (47), (50).

Proof. To estimate the l.h.s. of (A.1) note that

$$\left|\sum_{j=1}^{2} a_j \cos\left(\theta_1^{(j_1)} + \ldots + \theta_n^{(j_n)}\right)\right| \leqslant \left|\sum_{j=1}^{2} a_j \mathrm{e}^{\mathrm{i}\left(\theta_1^{(j_1)}+\ldots+\theta_n^{(j_n)}\right)}\right|, \tag{A.2}$$

so we need to estimate the last sum. To this end, we use (42) obtaining

$$\left|\sum_{j=1}^{2} a_j \mathrm{e}^{\mathrm{i}\left(\theta_1^{(j_1)}+\ldots+\theta_n^{(j_n)}\right)}\right| = \sum_{\varepsilon_1,\ldots,\varepsilon_n=\pm1} c(\varepsilon_1, \ldots, \varepsilon_n) \prod_{k=1}^{n} \left(\mathrm{e}^{\mathrm{i}\theta_k^{(1)}} + \varepsilon_k \mathrm{e}^{\mathrm{i}\theta_k^{(2)}}\right). \tag{A.3}$$

Next we define

$$\theta_k = \frac{\theta_k^{(1)} - \theta_k^{(2)}}{2}, \qquad \varphi_k = \frac{\theta_k^{(1)} + \theta_k^{(2)}}{2} \tag{A.4}$$

so that the r.h.s. of (A.3) simplifies to

$$2^n \mathrm{e}^{\mathrm{i}(\varphi_1+\ldots+\varphi_n)} \sum_{\varepsilon_1,\ldots,\varepsilon_n=\pm1} c(\varepsilon_1, \ldots, \varepsilon_n) \prod_{k=1}^{n} a_k(\varepsilon_k),$$

where $a_k(+1) = \cos\theta_k$ and $a_k(-1) = i\sin\theta_k$. Taking into account that we use an absolute value in (A.2) and divide by 2^n in (A.1), we have to prove the following inequality

$$\left|\sum_{\varepsilon_1,\ldots,\varepsilon_n=\pm1} c(\varepsilon_1, \ldots, \varepsilon_n) \prod_{k=1}^{n} a_k(\varepsilon_k)\right| \leqslant 2^{(n-1)/2}, \tag{A.5}$$

for any ±1-valued function $c(\varepsilon_1, \ldots, \varepsilon_n)$. We employ the induction method. For $n = 1$, we simply have

$$\left|c(+1)\cos\theta_1 + c(-1)i\sin\theta_1\right| = 1 = 2^{(1-1)/2}.$$

Then, we can write the sum in (A.5) as

$$\begin{aligned}
&\sum_{\varepsilon_1,\ldots,\varepsilon_n=\pm1} c(\varepsilon_1, \ldots, \varepsilon_n) \prod_{k=1}^{n} a_k(\varepsilon_k) \\
&\quad \equiv A_{n-1}\cos\theta_n + iB_{n-1}\sin\theta_n, \\
&\quad = \sum_{\varepsilon_1,\ldots,\varepsilon_{n-1}=\pm1} c(\varepsilon_1, \ldots, \varepsilon_{n-1}, +1) \prod_{k=1}^{n-1} a_k(\varepsilon_k)\cos\theta_n \\
&\quad + i \sum_{\varepsilon_1,\ldots,\varepsilon_{n-1}=\pm1} c(\varepsilon_1, \ldots, \varepsilon_{n-1}, -1) \prod_{k=1}^{n-1} a_k(\varepsilon_k)\sin\theta_n
\end{aligned}$$

where, according to the induction assumption, we have

$$|A_{n-1}|, \ |B_{n-1}| \leqslant 2^{(n-2)/2}. \tag{A.6}$$

The sum in (A.5) can be estimated in the following way

$$\begin{aligned}
&\left|\sum_{\varepsilon_1,\ldots,\varepsilon_n=\pm1} c(\varepsilon_1, \ldots, \varepsilon_n) \prod_{k=1}^{n} a_k(\varepsilon_k)\right| \\
&\quad = \left|A_{n-1}\cos\theta_n + iB_{n-1}\sin\theta_n\right| \\
&\quad \leqslant \sqrt{|A_{n-1}|^2 + |B_{n-1}|^2} \leqslant 2^{(n-1)/2}.
\end{aligned}$$

Now we show that with the coefficients of (47) or (50), the maximal value $2^{(n-1)/2}$ is attained. Due to (A.2), we need to estimate the sum

$$S_n = \frac{1}{2^{n+(n-1)/2}} \sum_{j_1,\ldots,j_n=1}^{2} a_{j_1,\ldots,j_n} \mathrm{e}^{\mathrm{i}\left(\theta_1^{(j_1)}+\ldots+\theta_n^{(j_n)}\right)} \tag{A.7}$$

and show that it can be equal to one by absolute value. First, let us consider the case of an odd n. From (50) we have

$$a_{j_1,\ldots,j_n} = 2^{(n+1)/2}(-1)^{\delta(j_1,\ldots,j_n)}, \qquad (j_1, \ldots, j_n) \in J.$$

Furthermore, from (45) it is

$$S_n = \frac{1}{2^n i}\left[\prod_{k=1}^{n}\left(\mathrm{e}_k^{\mathrm{i}\theta^{(1)}} + i\mathrm{e}_k^{\mathrm{i}\theta^{(2)}}\right) - \prod_{k=1}^{n}\left(\mathrm{e}_k^{\mathrm{i}\theta^{(1)}} - i\mathrm{e}_k^{\mathrm{i}\theta^{(2)}}\right)\right].$$

Taking into account that each term in these products can be written as

$$\mathrm{e}_k^{\mathrm{i}\theta^{(1)}} \pm \mathrm{e}_k^{\mathrm{i}\tilde{\theta}^{(2)}}, \qquad \tilde{\theta}_k^{(2)} = \theta_k^{(2)} + \pi/2,$$

and using the relations (A.4), S_n can be simplified to

$$S_n = \frac{1}{i}\left(\prod_{k=1}^{n}\cos\theta_k' \pm i\prod_{k=1}^{n}\sin\theta_k'\right)\mathrm{e}^{\mathrm{i}(\varphi_1'+\ldots+\varphi_n')},$$

where $\theta_k' = \theta_k - \pi/4$, $\varphi_k' = \varphi_k + \pi/4$. It is clear that the imaginary part of the sum S_n takes its maximal absolute value 1 when, for example, $\theta_k = \varphi_k = 0$ for $k = 1, \ldots, n$.

Now we consider the case of an even n. The coefficients $a_{j_1,\ldots,j_n}$ in this case come from (47)

$$a_{j_1,\ldots,j_n} = 2^{n/2}(-1)^{\tilde{\delta}(j_1,\ldots,j_n)},$$

and the sum S_n (A.7) becomes

$$S_n = \frac{1}{i2^n\sqrt{2}}\left(\mathrm{e}_n^{\mathrm{i}\theta^{(1)}} + \mathrm{e}_n^{\mathrm{i}\theta^{(2)}}\right)\left[\prod_{k=1}^{n-1}\left(\mathrm{e}_k^{\mathrm{i}\theta^{(1)}} + \mathrm{e}_k^{\mathrm{i}\tilde{\theta}^{(2)}}\right) - \prod_{k=1}^{n-1}\left(\mathrm{e}_k^{\mathrm{i}\theta^{(1)}} - \mathrm{e}_k^{\mathrm{i}\tilde{\theta}^{(2)}}\right)\right] + \frac{1}{i2^n\sqrt{2}}\left(\mathrm{e}_n^{\mathrm{i}\theta^{(1)}}\mathrm{e}_n^{\mathrm{i}\theta^{(2)}}\right)\left[\prod_{k=1}^{n-1}\left(\mathrm{e}_k^{\mathrm{i}\tilde{\theta}^{(1)}} + \mathrm{e}_k^{\mathrm{i}\theta^{(2)}}\right) - \prod_{k=1}^{n-1}\left(\mathrm{e}_k^{\mathrm{i}\tilde{\theta}^{(1)}} - \mathrm{e}_k^{\mathrm{i}\theta^{(2)}}\right)\right].$$

According to (A.4), S_n can be simplified to

$$S_n = \frac{\mathrm{e}^{\mathrm{i}\varphi}}{\sqrt{2}}\left[\left(i\prod_{k=1}^{n-1}\cos\theta_k' \mp \prod_{k=1}^{n-1}\sin\theta_k'\right)\cos\theta_n + \left(\prod_{k=1}^{n-1}\cos\theta_k'' \pm i\prod_{k=1}^{n-1}\sin\theta_k''\right)\sin\theta_n\right],$$

where $\theta_k' = \theta_k - \pi/4$, $\theta_k'' = \theta_k + \pi/4$ and $\varphi = \varphi_1 + \ldots \varphi_n + (n-1)\pi/4$. The imaginary part of S_n is (when $\varphi = 0$)

$$\left|\mathrm{Im}\left(S_n\right)\right| = \frac{1}{\sqrt{2}}\left|\cos\theta_n\prod_{k=1}^{n-1}\cos\left(\theta_k - \pi/4\right) \pm \sin\theta_n\prod_{k=1}^{n-1}\sin\left(\theta_k + \pi/4\right)\right|.$$

It is clear that this expression takes its maximal value 1 when $\theta_k = \pi/4$, $k = 1, \ldots, n-1$, and $\theta_n = \pm\pi/4$. This completes the proof. □

References

[1] Bell J S 1964 *Physics* **1** 195
Bell J S 1987 *Speakable and Unspeakable in Quantum Mechanics* (Cambridge: Cambridge University Press)
[2] Bohm D 1951 *Quantum Theory* (Englewood Cliffs, NJ: Prentice Hall)
[3] Bell J S 1986 *Ann. N.Y. Acad. Sci.* **480** 263
Banaszek K and Wodkiewicz K 1998 *Phys. Rev.* A **58** 4345
Banaszek K and Wodkiewicz K 1999 *Phys. Rev. Lett.* **82** 2009
Chen Z B, Pan J W, Hou G and Zhang Y D 2002 *Phys. Rev. Lett.* **88** 040406
Wenger J, Hafezi M, Grosshans F, Tualle-Brouri R and Grangier Ph 2003 *Phys. Rev.* A **67** 012105
[4] Mancini S, Manko V I and Tombesi P 1996 *Phys. Lett.* A **213** 1
Mancini S, Manko V I and Tombesi P 1997 *Found. Phys.* **27** 81
Mancini S, Manõko O V, Manko V I and Tombesi P 2001 *J. Phys.* A **34** 3461
[5] Styer D F *et al* 2002 *Am. J. Phys.* **70** 288
[6] see e.g., Special Issue: Quantum State Preparation and Measurement 1997 *J. Mod. Opt.* **44** N.11/12
Welsch D G, Vogel W and Opatrny T 1999 *Progress in Optics* **39** 63
[7] Manko O V and Manko V I 1997 *J. Russ. Laser Res.* **18** 407
Manko O V and Manko V I 2000 *J. Russ. Laser Res.* **21** 411
Manko O V and Manko V I 2004 *J. Russ. Laser Res.* **25** 477
[8] Manko V I and Mendes R V 2000 *Physica* D **145** 222
[9] Mancini S, Manko V I and Tombesi P 1995 *Quantum Semiclass Opt.* **7** 615
D'Ariano G M, Mancini S, Manko V I and Tombesi P 1996 *Quantum Semiclass Opt.* **8** 1017
[10] Mancini S, Manko V I, Shchukin E V and Tombesi P 2003 *J. Opt.* B **5** S333
[11] Dodonov V V and Manko V I 1997 *Phys. Lett.* A **229** 335
Manko V I and Manko O V 1997 *JETP* **85** 430
Manko V I and Safonov S S 1998 *Yad. Fisika* **61** 658
Andreev V A and Manko V I 1998 *JETP* **87** 239
Amiet J P and Weigert S 1999 *J. Phys.* A **32** L269
D'Ariano G M, Macone L and Paini M 2003 *Quantum Semiclass Opt.* **5** 77
[12] Leggio B, Man'ko V I, Man'ko M A and Messina A 2009 *Phys. Lett.* A **373** 4101
[13] Vogel K and Risken H 1989 *Phys. Rev.* A **40** 2847
D'Ariano G M, Leonhardt U and Paul H 1995 *Phys. Rev.* A **52** R1801
[14] Banaszek K and Wodkiewicz K 1996 *Phys. Rev. Lett.* **76** 4344
Wallentowitz S and Vogel W 1996 *Phys. Rev.* A **53** 4528
Mancini S, Man'ko V I and Tombesi P 1997 *Europhys. Lett.* **37** 79
[15] Cassinelli G, D'Ariano G M, de Vito E and Levrero A 2000 *J. Math. Phys.* **41** 7940
[16] Ibort A, Man'ko V I, Marmo G, Simoni A and Stornaiolo C 2013 *Phys. Scr.* **88** 055003
[17] Gruber P M 2007 *Convex and Discrete Geometry* (New York: Springer-Verlag)
[18] Clauser J F, Horne M A, Shimony A and Holt R A 1969 *Phys. Rev. Lett.* **23** 880
[19] Mermin N D 1990 *Phys. Rev. Lett.* **65** 1838
[20] Braunstein S L, Mann A and Revzen M 1992 *Phys. Rev. Lett.* **68** 3259
[21] Mista L, Filip R and Fiurasek J 2002 *Phys. Rev.* A **65** 062315
Chen Z B, Pan J W, Hou G and Zhang Y D 2002 *Phys. Rev. Lett.* **88** 040406
[22] Mancini S, Man'ko V I and Tombesi P 1997 *Quant. Semiclass. Opt.* **9** 987