An elaborated review with proofs of Schmidt canonical decomposition of any bipartite state vector is approached through general subsystem basis expansion. The upgraded forms of Schmidt decomposition in terms of correlation operator and twin observables are presented in detail. The discussion is extended to distant measurement, Einstein–Podolsky–Rosen states and Schrödinger’s steering. All claims and proofs are given in standard form unlike in the previous articles of the author where all results were obtained utilizing the very rarely used antilinear Hilbert–Schmidt maps of one subsystem state space into the other. For practical reasons the formalism of partial traces with their rules and reduced density operators together with correlation operator are used.

1 Introduction

The aim of this extended and elaborated review is to present a detailed exposition of Schmidt decomposition with almost all claims proved by arguments that mostly do not coincide with those in previous articles [1–7]. Namely, since much work has been done so far in the Belgrade school, the present-day views are more mature, hence they differ from the originally perceived ones and make possible simpler proofs. This fact alone should justify writing most of this review as if it were done for the first time. In addition, many of our research results have been previously presented in the formalism in which bipartite state vectors are written as antilinear Hilbert–Schmidt operators mapping one subsystem state space into the other (cf [1]), which is not well known, and it is very rarely used. Eventually, this approach has been found replaceable by standard basis-independent treatment. The basic aim is an in-depth study of the Schmidt decomposition. Its various forms are presented with its underlying foundations in three layers.

We assume that a completely arbitrary bipartite state vector, i.e., that of a two-subsystem composite system, \(|\Psi_{AB}\rangle\) is given. It is an arbitrary normalized vector in the composite-system state space \(\mathcal{H}_A \otimes \mathcal{H}_B\), where the factor spaces are finite- or infinite-dimensional complex separable Hilbert spaces, the state spaces of the subsystems
A and B. The statements are, as a rule, asymmetric in the roles of the two factor spaces. But, as it is well known, for every general asymmetric statement, also its analogous symmetric counterpart, obtained by exchanging the roles of subsystems A and B, is valid.

Having in mind local, i.e. subsystem measurement, we choose arbitrarily that it is performed on subsystem B. We call subsystem B the nearby one, and the opposite subsystem A, which is not affected dynamically by the local measurement, we call distant one. This is not a synonym for far away. But the suggestion of the latter may help to picture the lack of dynamical influence on subsystem A.

The basic mathematical tools in the analysis are the partial scalar product (elaborated in §8.1) and the rules of the partial trace (presented and proved in §8.2).

Hermitian operators, i.e., observables and subsystem density operators will be given, unless otherwise stated, in their so-called unique spectral forms, which are defined by lack of repetition in the eigenvalues. For instance, 

\[ O = \sum_k o_k \rho_k, \quad k \neq k' \Rightarrow o_k \neq o_{k'}, \]

where \( \Rightarrow \) denotes logical implication. Then \( \rho_k \) is said to be the eigen-projector of \( O \) that corresponds to the eigenvalue \( o_k \), and its range \( \mathcal{R}(\rho_k) \) is the corresponding eigen-subspace. We consider only Hermitian operators that have a purely discrete spectrum. They are called discrete operators.

Vectors that are not necessarily of norm one are written overlined throughout. Besides, when a number multiplies from the left a vector or an operator, the multiplication symbol \( \times \) may be put between them for clarity. One should keep in mind the convention that if a term in a sum has two or more factors and the first is zero, the rest need not be defined; it is understood that the entire term is zero.

In tensor products of vectors we put only occasionally the tensor multiplication sign \( \otimes \) when more clarity is required. By basis, we mean a complete orthonormal set of elements throughout.

The reader will not find, hopefully, the abundant use of mathematical structure (theorems, propositions, lemmata, corollaries, remarks and definitions) annoying. They are important for the many cross-references in the present paper, as well as for references in future articles. Besides, they reveal the logical status of the claim they contain.

### 2 Expansion in a subsystem basis

The natural framework for Schmidt (or biorthogonal) decomposition is decomposition in a factor-space basis, or, as we shall call it, expansion in a subsystem basis.

**Theorem 1.** (A) Let \( |\Psi_{AB}\rangle \in \mathcal{H}_A \otimes \mathcal{H}_B \) be any bipartite state vector. Let further \( \langle n|_B : \forall n \rangle \) be an arbitrary basis in the state space \( \mathcal{H}_B \). Then there exists a unique expansion in the subsystem basis

\[ |\Psi_{AB}\rangle = \sum_n \overline{|n\rangle}_A \langle n|_B. \tag{1} \]

The expansion coefficients \( \overline{|n\rangle}_A \) have the form

\[ \forall n : \overline{|n\rangle}_A = \sum_m \langle m|_A \langle n|_B |\Psi_{AB}\rangle \times |m\rangle_A, \tag{2} \]

where \( \langle n|_B : \forall m \rangle \) is an arbitrary basis in \( \mathcal{H}_A \). The expansion coefficients \( \overline{|n\rangle}_A \) in (1) are elements in \( \mathcal{H}_A \), and they are not necessarily of norm one. They depend only on \( |\Psi_{AB}\rangle \) and the corresponding basis elements \( |n\rangle_B \), and not on the choice of the rest of basis elements in the basis \( \{\langle n'|_B : \forall n'\} \).

The sums in (1) and (2), if infinite, are absolutely convergent, and one has

\[ |||\Psi_{AB}\rangle||^2 = \sum_n ||\overline{|n\rangle}_A||^2, \tag{3} \]

as well as

\[ \forall n : ||\overline{|n\rangle}_A||^2 = \sum_m \langle m|_A \langle n|_B |\Psi_{AB}\rangle^2. \tag{4} \]

In case of infinity, each of the sums is an absolutely convergent series as inherited from the absolutely convergent series

\[ |\Psi_{AB}\rangle = \sum_{m,n} \langle m|_A \langle n|_B |\Psi_{AB}\rangle \times |m\rangle_A |n\rangle_B. \tag{5} \]

Further, one can suitably write \( \forall n : \overline{|n\rangle}_A = ||\overline{|n\rangle}_A|| \times |n\rangle_A \) (definition of the norm-one elements \( |\langle n|_B : \forall n\rangle \rangle \)), and replace these in (1). Relation (1) then becomes

\[ |\Psi_{AB}\rangle = \sum_n ||\overline{|n\rangle}_A|| \times |n\rangle_A |n\rangle_B. \tag{6} \]

This is an expansion in the orthonormal set of elements \( \{\langle n|_B |n\rangle_B : \forall n\} \) in \( \mathcal{H}_A \otimes \mathcal{H}_B \). Actually, it is normal in both factors, but orthogonal, in general, only in the second one. Some norm-one elements \( |n\rangle_A \) may not exist, when \( \overline{|n\rangle}_A = 0 \) (depending on \( |\Psi_{AB}\rangle \)).

(B) The expansion coefficients can be evaluated utilizing the partial scalar product

\[ \forall n : \overline{|n\rangle}_A = \langle n|_B |\Psi_{AB}\rangle_A \tag{7}. \]

**Proof.** (A) is straightforward, but, on account of the importance of the theorem (see end of the section), it is presented as easy reading.
Let \(|m\rangle_A : \forall m\) be an arbitrary basis in \(\mathcal{H}_A\). Then one can perform the expansion (5). As it is well known, if the double-sum is infinite, the series is absolutely convergent allowing any change of order in which the terms are written (any permutation). Hence we can group together all terms around each \(|n\rangle_B\) tensor factor and rewrite (5) as

\[
|\Psi\rangle_{AB} = \sum_n \left( \sum_m \left( \langle m|_A \langle n|_B |\Psi\rangle_{AB} \right) \times |m\rangle_A \right) |n\rangle_B. \tag{8}
\]

Thus one obtains (1) and (2).

In this way we have established that the claimed expansion exists. Now we show that the expansion coefficients \(|m\rangle_A\) in (1) do not depend on the choice of the basis \(|m\rangle_A : \forall m\). Let \(|\langle k\rangle_A : \forall k\) be any other basis in \(\mathcal{H}_A\), and let \(\forall m : |m\rangle_A = \sum_k U_{m,k} |k\rangle_A\) be the unitary transition matrix. Then, starting with the expansion coefficient evaluated in the first basis, we find out its form in the second basis:

\[
|\bar{m}\rangle_A = \sum_n \left( \langle m|_A \langle n|_B |\Psi\rangle_{AB} \times |m\rangle_A \right) = \sum_k \sum_k' U_{m,k}^* U_{m,k'} \langle k|_A \langle n|_B |\Psi\rangle_{AB} \times |k\rangle_A.
\]

Since \(\sum_m U_{m,k}^* U_{m,k'} = \delta_{k,k'}\) is valid for the unitary transition matrix elements, one is further led to

\[
|\bar{m}\rangle_A = \sum_k \langle k|_A \langle n|_B |\Psi\rangle_{AB} \times |k\rangle_A. \tag{9}
\]

If the expansion coefficient were evaluated in the other basis, it would give the same element of \(\mathcal{H}_A\). The additional claims in (A) are obvious.

(B) Proof of (2) is given in §8.1 where the partial scalar product is defined in three and a half ways; one of them consisting precisely in equating RHS of (2) and RHS of (7).

**Corollary 1.** If the nearby state is pure, i.e., a state vector, e.g., \(|\bar{n}\rangle_B\), then also the distant state is necessarily pure, but it can be arbitrary (depending on \(|\Psi\rangle_{AB}\)).

**Proof.** By assumption \(\rho_B = tr_A (|\Psi\rangle_{AB} \langle \Psi|_{AB}) = |\bar{n}\rangle_B \langle \bar{n}|_B\).

Choosing a nearby subsystem basis \(|\bar{n}\rangle_B : \forall n\) so that it contains \(|\bar{n}\rangle_B\), one obtains \(|\bar{n}\rangle_B \rho_B |\bar{n}'\rangle_B = \delta_{n\bar{n},n'\bar{n}} |\bar{n}\rangle_B \langle \bar{n}|_B\).

On the other hand, expansion (1) implies

\[
\langle n|_B \rho_B |n'\rangle_B = \langle n|_B tr_A (|\Psi\rangle_{AB} \langle \Psi|_{AB}) |n'\rangle_B = |\bar{n}\rangle_A \langle n'|_A.
\]

Altogether,

\[
|\bar{n}\rangle_A \langle n'|_A = \delta_{n\bar{n},n'\bar{n}} |\bar{n}\rangle_B \langle \bar{n}|_B.
\]

i.e., \(\forall n : |\bar{n}\rangle_A = \delta_{n\bar{n}} |\bar{n}\rangle_A\). Hence, \(|\Psi\rangle_{AB} = |\bar{n}\rangle_A \langle \bar{n}|_B\). Note that \(|\bar{n}\rangle_A\) is of norm one because so are \(|\Psi\rangle_{AB}\) and \(|\bar{n}\rangle_B\).

As an alternative proof of Corollary 1 one may consider the canonical Schmidt decomposition (cf Definition 3 and relation (14) together with (15), (16) below). Then the claim in Corollary 1 is obvious, but the burden of the proof lies on Theorem 3.

**Definition 1.** (Purification) If \(\rho\) is an arbitrary mixed state (density operator that is not a rewritten state vector) of a quantum system in the state space (Hilbert space) \(\mathcal{H}\), then one can isomorphically map \(\mathcal{H}\) onto the subsystem state space \(\mathcal{H}_A\) of a bipartite quantum system the state of which is in \(\mathcal{H}_A \otimes \mathcal{H}_B\), and find a state vector \(|\Psi\rangle_{AB}\) such that its first subsystem state operator (reduced density operator) \(\rho_A\) is isomorphic to the initially given \(\rho\). This procedure is called purification.

**Theorem 2.** (Purifiability) Any mixed state \(\rho\) can be purified if it is written as any mixture

\[
\rho = \sum_n |\bar{n}\rangle \langle \bar{n}|
\]

by writing down a bipartite state vector \(|\Psi\rangle_{AB}\) in the form (1) with any basis in \(\mathcal{H}_B\), denoted as \(|\langle n|_B : \forall n\), with expansion coefficients \(|\bar{n}\rangle\) given by (3) with added index \(A\). The subsystem state operator (reduced density operator) \(\rho_A\) is then isomorphic to \(\rho\).

**Proof.** Evaluating \(\rho_A = tr_B (|\Psi\rangle_{AB} \langle \Psi|_{AB})\) and keeping in mind that \(tr(|\bar{n}\rangle_A \langle n'|_A) = \delta_{n\bar{n},n'\bar{n}}\), one obtains \(\rho_A = \sum_n |\bar{n}\rangle_A \langle \bar{n}|_A\).

To understand the importance of subsystem-basis decomposition (1), one must realize that expansion (1) is a crossroad. A number of different paths lead from it:

(i) Definition of the partial scalar product. Von Neumann in his seminal book [9], in which he gave the mathematical grounding of quantum mechanics in case of infinite-dimensional state spaces, did not encompass partial scalar product and partial trace. Therefore, a careful mathematical exposition of these concepts is given, together with the basic properties, in §3.

(ii) The expansion at issue leads to purification (cf Theorem 2 and relation (4)).

(iii) It is the framework for Schmidt decomposition (see §3 and further).

(iv) Remark 6 and relation (29) below open the way for a more fruitful application of (1), particularly for Schrödinger’s important concept of steering (cf §6.3).

(v) Expansion (1) gives a new angle on the concept of erasure (cf Remark 23 below).

(vi) A theory of preparation in quantum experiments can be based on (1): If the preparator is subsystem B, and the object on which the experiment is conducted is subsystem A, and if \(|\Psi\rangle_{AB}\) is the state after interaction,
then $|n_B\rangle$ is the state of the preparator that the experimenter sees at the end of the preparation, and simultaneously $|n\rangle_A$ is then the state of the experimental object (at the beginning of the experiment). This will be elaborated in future work.

(vii) Expansion (1) can play a crucial role in Everett’s relative-states interpretation of quantum mechanics: The state $|n\rangle_A$ is the relative state of subsystem A with respect to the state $|n\rangle_B$ of subsystem B in the composite-system state $|\Psi\rangle_{AB}$. A detailed discussion of this and its ramifications is left for future work.

Subsystem-basis expansion (1), and the enumerated paths (i) and (iv)-(vii) that lead away from it were not analyzed in previous work. This material is new in this article.

3 Schmidt decomposition

Now we define Schmidt (or biorthogonal) decomposition. It is well known and much used in the literature.

Definition 2. If besides the basis elements $|n\rangle_A$ also the expansion coefficients $|n\rangle_A$ are orthogonal in expansion (1), then one speaks of a Schmidt or biorthogonal decomposition. It is usually written in terms of subsystem state vectors $|n\rangle_A : \forall n$ that are not only orthogonal, but also normalized:

$$|\Psi\rangle_{AB} = \sum_n \alpha_n |n\rangle_A |n\rangle_B,$$

(11)

where $\forall n : \alpha_n$ are complex numbers, and $\forall n : |n\rangle_A$ and $|n\rangle_B$ for the same $n$ value are referred to as partners in a pair of Schmidt states.

The term Schmidt decomposition can be replaced by Schmidt expansion or Schmidt form. To avoid confusion, we will stick to the first term throughout (as it is usually done in the literature).

Theorem 3. Expansion (1) is a Schmidt decomposition if and only if the second-tensor-factor-space basis $|n\rangle_B : \forall n$ is an eigenbasis of the corresponding reduced density operator $\rho_B = tr_A (|\Psi\rangle_{AB}\langle\Psi|_{AB})$:

$$\forall n : \rho_B |n\rangle_B = \rho B |n\rangle_B, \quad 0 \leq \rho B.$$

(12)

Proof. Let us evaluate $|n\rangle_A |n\rangle_B$ making use of (2).

$$\langle n\rangle_A |n\rangle_B = (\langle n\rangle_A |\Psi\rangle_{AB}) (\langle n\rangle_B |\Psi\rangle_{AB})$$

$$= \langle n\rangle_A |\Psi\rangle_{AB} |\Psi\rangle_{AB}$$

$$= tr (|\Psi\rangle_{AB}\langle\Psi|_{AB})(|n\rangle_B\langle n|_B)$$

$$= tr_B (tr_A (|\Psi\rangle_{AB}\langle\Psi|_{AB})(|n\rangle_B\langle n|_B))$$

$$\langle n\rangle_B |n\rangle_B = (\langle n|_B |\rho\rangle_B (|n\rangle_B |n\rangle_B).$$

The third equality in the above derivation, where the expectation value is rewritten as a suitable trace, is a standard, textbook step. (Evaluating the trace in a basis in which the relevant state vector is one of the basis elements, the equality becomes obvious.) In the fourth equality, the first partial-trace rule (cf §8.2) was used.

We have obtained

$$\langle n\rangle_A |n\rangle_B = \rho B |n\rangle_B.$$

(13)

It is clear from relation (13) that the vectors $|n\rangle_A : \forall n$ are orthogonal if and only if $\rho B$ is diagonal, and this is the case if and only if the eigenvalue relations (12) are valid as claimed.

Corollary 2. If one expands $|\Psi\rangle_{AB}$ in a second-subsystem basis like in (1), then the subsystem state (reduced density operator) $\rho_A$ is given as a mixture (3). If, in addition, the B-subsystem basis is an eigenbasis of $\rho_B$, then (3) is simultaneously also a spectral decomposition of $\rho_A$ (in terms of its eigenvectors).

Now we turn to a special form of Schmidt decomposition that is often more useful. It is called canonical Schmidt decomposition. It is due to the fact that the non-trivial phase factors of the non-zero coefficients $\alpha_n$ in (11) can be absorbed either in the basis elements in $H_A$ or in those in $H_B$ (or partly in the former and partly in the latter).

Definition 3. If in a Schmidt expansion (11) all $\alpha_n$ are non-negative real numbers, then we write the expansion in the following way:

$$|\Psi\rangle_{AB} = \sum_i r_i^{1/2} |\hat{n}\rangle_A |\hat{n}\rangle_B,$$

(14)

and the sum is confined to non-zero terms (one is reminded of all this by the replacement of the index $n$ by $i$ in this notation). Relation (14) is called canonical Schmidt decomposition. (The term canonical reminds of the form of (14), i.e., of $\forall i : r_i^{1/2} > 0$.)

Needless to say that every $|\Psi\rangle_{AB}$ can be written in the form of a canonical Schmidt decomposition, and it is, of course, non-unique.

Corollary 3. Every canonical Schmidt decomposition (14) is accompanied by the spectral forms of the reduced density operators:

$$\rho A = \sum_i r_i |\hat{n}\rangle_A \langle \hat{n}|_A,$$

(15)

$$\rho B = \sum_i r_i |\hat{n}\rangle_B \langle \hat{n}|_B.$$

(16)
Note that the same eigenvalues \( r_i \) appear in \((14)\) and in the two spectral forms \((15)\) and \((16)\). Note also that \((15)\) is the same as \((3)\) if the RHS of \((3)\) is determined by \((1)\), and \(|n\rangle_B : \forall n\rangle\) is an eigenbasis of \(\rho_B\).

Proof. The Schmidt canonical decomposition \((14)\) allows the straightforward evaluation

\[
\rho_A = \text{tr}_B (|\Psi\rangle\langle \Psi|_{AB}) = \sum_{i,j} \frac{1}{r_i} \frac{1}{r_j} \text{tr}_B (|\hat{b}_A \rangle |\hat{b}_j \rangle\langle \hat{b}_i \rangle\langle \hat{b}_j \rangle_{AB}) \\
= \sum_{i,j} \frac{1}{r_i} \frac{1}{r_j} \langle \hat{b}_i | \langle \hat{b}_j | (|\hat{b}_A \rangle |\hat{b}_j \rangle\langle \hat{b}_i \rangle_{AB}) \\
= \sum_i r_i |\hat{b}_A \rangle |\hat{b}_i \rangle_{A}
\]

(the first partial-trace rule in \((8.2)\) was used). Relation \((16)\) is proved symmetrically. \(\square\)

Remark 1. The ranges \(\mathcal{R}(\rho_s)\), \(s = A, B\), of the reduced density operators \(\rho_s\), \(s = A, B\) are equally dimensional. The common dimension is the number of terms in a canonical Schmidt decomposition \((14)\). It is sometimes called the Schmidt rank of the given bipartite state vector.

We denote the range-projectors of the reduced density operators \(\rho_s\), \(s = A, B\) by \(Q_s\), \(s = A, B\). It is seen from \((15,16)\) that

\[
Q_A = \sum_i |\hat{d}_A \rangle |\hat{d}_i \rangle_{A}, \\
Q_B = \sum_i |\hat{d}_B \rangle |\hat{d}_i \rangle_{B}.
\]

The reduced density operators have equal positive eigenvalues \(|r_i > 0 : \forall i\rangle\) (implying equality of the multiplicities of the distinct ones among them). The possible zero eigenvalues may differ arbitrarily (cf \((15,16)\)).

The Schmidt canonical decomposition was studied in a previous work \((2)\).

Corollary 4. The following relations are always valid:

\[
|\Psi\rangle_{AB} = Q_s |\Psi\rangle_{AB}, \quad s = A, B
\]

Proof. Since \(Q_s = \sum_i |\hat{d}_s \rangle |\hat{d}_i \rangle_{s}\), \(s = A, B\), the claim is obvious when \(|\Psi\rangle_{AB}\) is written as a canonical Schmidt decomposition \((14)\). \(\square\)

Corollary 5. One always has

\[
|\Psi\rangle_{AB} \in \mathcal{R}(Q_A Q_B)
\]

Remark 2. If we enumerate by \(j\) the distinct positive eigenvalues \(|r_j > 0 : \forall j\rangle\) of \(\rho_s\), \(s = A, B\), and by \(Q_j\), \(s = A, B\) the corresponding eigenprojectors, then one has the relations

\[
\rho_s = \sum_j r_j Q_j, \quad s = A, B
\]

\[
\mathcal{R}(\rho_s) = \mathcal{R}(Q_s) = \sum_j \mathcal{R}(Q_j) \quad s = A, B
\]

\[
\forall j : \text{dim} (\mathcal{R}(Q_j)) < \infty.
\]

As to \((22)\), one should note that if and only if \(\text{dim} (\mathcal{R}(\rho_s)) = \infty\), \(s = A, B\), then the range \(\mathcal{R}(\rho_s)\) is a linear manifold that is not equal but only dense in its topological closure \(\bar{\mathcal{R}}(\rho_s)\), \(s = A, B\). The symbol \(\oplus\) denotes orthogonal sum of subspaces.

One should also note that all positive-eigenvalue eigensubspaces \(\mathcal{R}(Q_j)\) are necessarily always finite dimensional \((23)\) because \(\sum_i r_i = 1\) (a consequence of the normalization of \(|\Psi\rangle_{AB}\), and hence no positive-eigenvalue can have infinite degeneracy. But there may be denumerably infinitely many distinct positive eigenvalues \(r_j\).

We refer to \((21,22,23)\) as the subsystem picture of \(|\Psi\rangle_{AB}\). It serves as a first layer of an underlying grounding for Schmidt decomposition.

Remark 3. One can say that one has a canonical Schmidt decomposition \((14)\) if and only if the expansion is biorthonormal and all expansion coefficients are positive.

Remark 4. A canonical Schmidt decomposition \((14)\) of any bipartite state vector \(|\Psi\rangle_{AB}\) is non-unique because the eigen-sub-basis \(|\hat{b}_B \rangle : |\hat{v}_i\rangle\) of \(\rho_B\) spanning its range \(\mathcal{R}(\rho_B)\) is non-unique. Even if \(\rho_B\) is non-degenerate in all its positive eigenvalues, there is the non-uniqueness of the phase factors of \(\hat{b}_B\).

4 Correlated Schmidt decomposition

We investigate further the mentioned non-uniqueness (see end of the preceding section). In the canonical Schmidt decomposition \((14)\) it is clear that the entanglement in \(|\Psi\rangle_{AB}\) boils down to the choice of the partner in the terms of the decomposition.

We introduce explicitly this choice of a partner keeping in mind the subsystem picture (cf \((21,22,23)\)). It turns out that the best thing to do is to define an antiunitary map that takes the (topologically) closed range \(\bar{\mathcal{R}}(\rho_B)\) onto the symmetrical entity \(\bar{\mathcal{R}}(\rho_A)\).

The map is called the correlation operator, and it is denoted by the symbol \(U_s\) \((2,4)\).
Definition 4. If a canonical Schmidt decomposition \([14]\) is given, then the two orthonormal sub-bases of equal power \(|\tilde{\mathbf{0}}_B : \forall i\rangle\) and \(|\tilde{\mathbf{0}}_A : \forall i\rangle\) appearing in it determine an antilinear, i.e., antilinear and unitary, operator \(U_a\), the correlation operator, which is a correlation entity inherent in the given state vector \(|\Psi_{AB}\rangle\):

\[
\forall i : |\tilde{\mathbf{0}}_A \equiv (U_a |\tilde{\mathbf{0}}_B\rangle_B).
\]

The correlation operator \(U_a\), mapping \(\tilde{\mathcal{R}}(\rho_B)\) onto \(\tilde{\mathcal{R}}(\rho_A)\), is well defined by \((24)\) and by the additional requirements of antilinearity (complex conjugation of numbers, coefficients in any linear combination on which the operator may act) and continuity (if the bases are infinite). Both these requirements follow from that of antiunitarity. Preservation of every scalar product up to complex conjugation, which, by definition, makes \(U_a\) antilinear, is easily seen to follow from \((24)\) and the requirements of antilinearity and continuity because \(U_a\) takes a basis into another one.

Definition 5. On account of Definition 4 and \((24)\), any canonical Schmidt decomposition \((14)\) of any bipartite state vector \(|\Psi_{AB}\rangle\) can be written in the form

\[
|\Psi_{AB}\rangle = \sum_i r_i^{\frac{1}{2}} (U_a |\tilde{\mathbf{0}}_B\rangle_B) \otimes |\tilde{\mathbf{0}}_B\rangle.
\]

This form is called a correlated canonical Schmidt decomposition. (In [5] §2, instead of the term correlated, the term strong was used.)

One should note that \((25)\) contains all the entities that appear in \((14)\) plus (explicitly) the correlation operator \(U_a\), which is implicitly contained in \((14)\). Expansion \((25)\) makes explicit the fact that the opposite-subsystem eigen-sub-basis \(|\tilde{\mathbf{0}}_A : \forall i\rangle\) in \((14)\) is not just any such set of vectors once the eigen-sub-basis \(|\tilde{\mathbf{0}}_B : \forall i\rangle\) is chosen.

Theorem 4. The correlation operator \(U_a\) is uniquely determined by the given (arbitrary) bipartite state vector \(|\Psi_{AB}\rangle\).

Proof. Let \(|\tilde{j},k\rangle_B : \forall k, \forall j\rangle\) and \(|\tilde{l},j\rangle_B : \forall l, \forall j\rangle\) be two arbitrary eigen-sub-bases of \(\rho_B\) spanning \(\tilde{\mathcal{R}}(\rho_B)\). The vectors are written with two indices, \(j\) denoting the eigen-subspace \(\tilde{\mathcal{R}}(Q_B^j)\), \(\forall j\rangle : Q_B^j = \sum_k |\tilde{j},k\rangle_B\langle j,k|_B\) corresponding to the eigenvalue \(r_j\) of \(\rho_B\) to which the vector belongs, and the other index \(k\) \((l\rangle\) enumerates the vectors within the eigen-subspace \(\tilde{\mathcal{R}}(Q_B^j)\) in case the eigenvalue \(r_j\) of \(\rho_B\) is degenerate, i.e., if its multiplicity is more than 1.

The proof goes as follows. Let \(
\forall j : |j,k\rangle_B = \sum_{l_j} U_{k,l_j}^{(j)} |j,l\rangle_B.,
\)

where \(U_{k,l_j}^{(j)}\) are unitary sub-matrices. Then, keeping \(U_a\) one and the same, we can start out with the correlated Schmidt decomposition in the \(k_j\)-eigen-sub-basis, and after a few simple steps (utilizing the antilinearity of \(U_a\) and the unitarity of the transition sub-matrices), we end up with the correlated Schmidt decomposition (of the same \(|\Psi_{AB}\rangle\)) in the \(l_j\)-eigen-sub-basis. Complex conjugation is denoted by asterisk.

\[
|\Psi_{AB}\rangle = \sum_j r_j^{\frac{1}{2}} \sum_{k_j} (U_a |j,k_j\rangle_B) \otimes |j,k_j\rangle_B
\]

\[
= \sum_j r_j^{\frac{1}{2}} \sum_{k_j} (U_{k,l_j}^{(j)})^\dagger (U_a |j,l_j\rangle_B) \otimes |j,l_j\rangle_B
\]

\[
= \sum_j r_j^{\frac{1}{2}} \sum_{l_j} \sum_{k_j} (U_{k,l_j}^{(j)})^\dagger (U_a |j,l_j\rangle_B) \otimes |j,l_j\rangle_B
\]

\[
= \sum_j r_j^{\frac{1}{2}} \sum_{l_j} \delta_{l_j,j} (U_a |j,l_j\rangle_B) \otimes |j,l_j\rangle_B
\]

\[
= \sum_j r_j^{\frac{1}{2}} (U_a |j,l_j\rangle_B) \otimes |j,l_j\rangle_B.
\]

It may seem that the uniqueness of \(U_a\) when \(|\Psi_{AB}\rangle\) is given is a poor compensation for the trouble one has treating an antilinear operator. But the difficulty is more psychological than practical, because all that distinguishes an antilinear operator from a unitary one is

(i) its antilinearity, i.e., it complex-conjugates the numbers in any linear combination on which it acts, and

(ii) its property that it complex-conjugates every scalar product (preserving its absolute value):

\[
\langle \psi | \phi \rangle = \left( \langle \psi | U_a^\dagger \rho_B U_a | \phi \rangle \right)^{\frac{1}{2}}.
\]

The full compensation comes, primarily from the insight in entanglement that \(U_a\) furnishes, from its practical usefulness, and, at last but not at least, from its important physical meaning.

The physical meaning of the correlation operator \(U_a\) is best discussed in the context of Schrödinger’s steering (see three passages beneath relation (70) in §6.3 below). One should realize that in quantum mechanics comes always heavily packed in mathematics. One must discern the physics in the haze of the formalism. This is attempted below.

Remark 5. The correlated Schmidt canonical expansion \((25)\) can be viewed in two opposite ways:

(i) as a given bipartite state vector \(|\Psi_{AB}\rangle\) determining its two inherent entities, the reduced density operator \(\rho_B\) in spectral form (cf \((16)\)) and the correlation operator \(U_a\) (cf \((24)\)), both relevant for the entanglement in the state vector (and one can read them in the given expansion);
(ii) as a given pair \((\rho_B, U_a)\) (\(U_a\) mapping antiunitarily \(R(\rho_B)\) onto some equally dimensional subspace of \(\mathcal{H}_A\)) determining a bipartite state vector \(|\Psi\rangle_{AB}\).

The second view of correlated Schmidt expansion allows a systematic generation and classification of all state vectors in \(\mathcal{H}_A \otimes \mathcal{H}_B\) (cf [6]).

**Theorem 5.** The expansion coefficients \(\langle \overline{n} \rangle_A : \forall n \rangle\) in any subsystem-basis expansion \(|\Psi\rangle = \sum_n (U_a \rho_B^{1/2} |n\rangle_B) \otimes |n\rangle_B\) (1) can be evaluated, besides by (2), also utilizing the reduced density operator \(\rho_B\) and the correlation operator \(U_a\) as follows:

\[
\forall n : \quad \langle \overline{n} \rangle_A = \left( U_a \rho_B^{1/2} |n\rangle_B \right) .
\]

Proof. We substitute a canonical Schmidt decomposition of \(|\Psi\rangle_{AB}\) in (2) for an arbitrary \(n\) value:

\[
\langle \overline{n} \rangle_A = \langle n |_B \langle \Psi |_{AB} = \langle n |_B \left( \sum_i r_i^2 |i_A \rangle_B \right) = \sum_i r_i^2 \langle n |_B \langle i |_B \rangle_A .
\]

On the other hand, evaluating the RHS of (26) making use of the spectral form (16) of \(\rho_B\) and of (24), we obtain:

\[
U_a \rho_B^{1/2} |n\rangle_B = U_a \sum_i \frac{1}{r_i^2} |i \rangle_B \langle i |_B \langle n |_B = \sum_i \frac{1}{r_i^2} \langle i |_B \langle n |_B \times \langle i |_B .
\]

The asterisk denotes complex conjugation. It is required by the antilinearity property of the correlation operator. Comparing (27) and (28), we see that they are equal. \(\Box\)

Theorem 5 as it stands, is new with respect to previous work. Though, in [2, Eq. 34] an analogous result was obtained, the derivation was formulated and presented in the approach in which bipartite states are written as antilinear Hilbert–Schmidt mappings of \(\mathcal{H}_B\) into \(\mathcal{H}_A\). This approach is almost never used in the literature.

**Remark 6.** Substituting (26) in (1) one obtains

\[
|\Psi\rangle_{AB} = \sum_n \left( U_a \rho_B^{1/2} |n\rangle_B \right) \otimes |n\rangle_B .
\]

This can be called a generalized correlated canonical Schmidt decomposition. Note that the nearby subsystem basis \(|n \rangle_B : \forall n\rangle\) is not necessarily an eigenbasis of \(\rho_B\); it is arbitrary. This is how it is a generalization. Form (29) of expansion in a subsystem basis is relevant for Schrödinger’s steering discussed in detail in §6.3 below.

**Remark 7.** Theorem 5 and relation (26) enables one to prove the uniqueness of the correlation operator \(U_a\) independently of Theorem 4. Namely, this uniqueness is a consequence of the uniqueness of the partial scalar product (proved in §8.1).

**Remark 8.** When a pair of orthonormal sub-bases \(|i\rangle_B : \forall i\rangle\) and \(|i \rangle_A : \forall i\rangle\) appearing in a canonical Schmidt decomposition (14) is given, one can extend \(U_a\) to the entire \(\mathcal{H}\), denote the extended operator as \(U_a\), and write

\[
U_a = \sum_i |i \rangle_A K_i |i \rangle_B .
\]

where \(K\) is complex conjugation (denoted by asterisk when acting on numbers). Definition (30) is actually symbolical. Its true meaning consists in the following.

\[
\forall \langle \phi |_B \in \mathcal{H}_B : \quad \hat{U}_a \langle \phi |_B = \sum_i \langle i |_A K_i |i \rangle_B |\phi \rangle_B = \sum_i \langle \langle i |_B | \phi \rangle_B |^* \langle i |_A .
\]

The extended operator \(\hat{U}_a\) acts as \(U_a\) in the range \(\mathcal{R}(\rho_B)\), and it acts as zero in the null space of \(\rho_B\). In other words, one can write

\[
\hat{U}_a = U_a Q_B .
\]

where \(Q_B\) is the range projector of \(\rho_B\). Since \(Q_B\) projects onto the range, it does not matter that \(U_a\) is defined only on the range.

**Remark 9.** As one can easily see, utilizing complete orthonormal eigen-bases of \(\rho_s\), \(s = A, B\) (cf [15,16,17,18]), one has

\[
\rho_A = U_a \rho_B U_a^{-1} \rho_A ,
\]

\[
\rho_B = U_a^{-1} \rho_B U_a Q_B .
\]

Thus, the reduced density operators are, essentially, images of each other via the correlation operator. The term essentially points to the fact that the dimensions of the null spaces are independent of each other. Property (33,34) is called twin operators. More will be said about such pairs of operators below, cf Definition 6 below.

In terms of subspaces, to (33,34) correspond the image-relations

\[
\mathcal{R}(\rho_A) = U_a \mathcal{R}(\rho_B) ,
\]

\[
\mathcal{R}(\rho_B) = U_a^{-1} \mathcal{R}(\rho_A) .
\]

One obtains an even more detailed view when one takes into account the eigen-subspaces \(\mathcal{R}(\rho_s)\) of \(\rho_s\), corresponding to (the common) distinct positive eigenvalues \(r_j\) of \(\rho_s\), where \(Q^\perp_s\) projects onto the \(r_j\)-eigen-subspace, \(s = A, B\) (cf the subsystem picture 21,22,23). Then one obtains a
view of the *entanglement* in a given composite state \( |\Psi\rangle_{AB} \) in terms of the so-called *correlated subsystem picture* [3]:

\[
\rho_s = \sum_j r_j Q_s^j, \quad s = A, B, \tag{37}
\]

and in terms of subspaces

\[
\mathcal{R}(\rho_s) = \sum_j \mathcal{R}(Q_s^j), \quad s = A, B, \tag{38}
\]

where \( \oplus \) denotes an orthogonal sum of subspaces.

Further, as it is also straightforward to see in eigenbases of \( \rho_s \), \( s = A, B \),

\[
\forall j : \quad \mathcal{R}(Q_s^j) = U_a \mathcal{R}(Q_s^j), \quad \mathcal{R}(Q_s^j) = U_a^{-1} \mathcal{R}(Q_s^j). \tag{39}
\]

The correlation operator makes not only the ranges of the reduced density operators images of each other, but also all positive-eigenvalue eigen-subspaces of the reduced density operators. In other words, the correlation operator \( U_a \), making the reduced density operators \( \rho_s \), \( s = A, B \) images of each other, makes also the eigen-decompositions of the ranges \( \mathcal{R}(\rho_s) \), \( s = A, B \) images of each other.

The relations (38) and (37) constitute the *correlated subsystem picture* of the given state vector \( |\Psi\rangle_{AB} \) in terms of operators and corresponding subspace state entities. This is the second layer in the underlying grounding of the (correlated) Schmidt decomposition.

### 5 Twin-correlated Schmidt decomposition

In the correlated subsystem picture of a given bipartite state vector \( |\Psi\rangle_{AB} \) (in the preceding section) we have searched for a comprehension of entanglement and its canonical form, but doing so we have investigated only state entities \( \rho_A, \rho_B, U_a \). Now, we introduce observables that can contribute to the theory by enriching and broadening our understanding.

**Lemma 1.** Let \( |\Psi\rangle_{AB} \) be a bipartite state vector, \( \rho_A \) its first-subsystem reduced density operator, \( Q_A \) the range projector of the latter, and \( O_A = \sum_k o_k P_A^k \) a first-subsystem observable in spectral form. Let further \( P_s^{\tau} \) be the sum of all those eigen-projectors \( P_A^k \) of \( O_A \) that do not nullify \( |\Psi\rangle_{AB} \). Then, \( P_s^{\tau} Q_A = Q_A \), i.e., \( P_s^{\tau} \geq Q_A \), or, in words, \( P_s^{\tau} \) is larger than \( Q_A \) or equivalently \( \mathcal{R}(P_s^{\tau}) \supseteq \mathcal{R}(Q_A) \).

**Proof.** One can write

\[
\rho_A = \text{tr}_B (|\Psi\rangle_{AB} \langle \Psi|_{AB})
\]

\[
= \text{tr}_B \left( \sum_k P_A^k |\Psi\rangle_{AB} \langle \Psi|_{AB} \right)
\]

\[
= P_s^{\tau} \rho_A
\]

where the second partial-trace rule in [8, 2] was used.

Taking an eigen-sub-basis \( \{|\tilde{\rho}_A : vi\} \rangle \rho_A \) spanning its range, one can further write \( \rho_A \) in spectral form and one obtains

\[
\sum_i r_i |\tilde{\rho}_A⟩⟨\tilde{\rho}_A| = \sum_i r_i P_s^{\tau} |\tilde{\rho}_A⟩⟨\tilde{\rho}_A|, \quad \forall i : \quad r_i > 0.
\]

Applying this to an eigenvector \( |\tilde{\tilde{\rho}}_A \rangle \) corresponding to \( r_i > 0 \), one obtains \( r_i |\tilde{\tilde{\rho}}_A⟩ = r_i P_s^{\tau} |\tilde{\tilde{\rho}}_A⟩ \). Finally, since \( Q_A = \sum_i |\tilde{\rho}_A⟩ ⟨\tilde{\rho}_A| \), the claimed relation follows. \( \square \)

**Definition 6.** Let \( O_A \equiv \sum_k a_k P_A^k \) and \( O_B \equiv \sum_l b_l P_B^l \) be opposite-subsystem Hermitian operators (observables) in spectral form. If one can reenumerate all eigen-projectors \( P_A^k \) and \( P_B^l \) that do not nullify the given composite state vector \( |\Psi\rangle_{AB} \) by a common index, e.g. \( m \), so that

\[
\forall m : \quad P_A^m |\Psi\rangle_{AB} = P_B^m |\Psi\rangle_{AB} \tag{40}
\]

is valid, then the operators \( O_A \) and \( O_B \) are said to be twin operators or twin observables in \( |\Psi\rangle_{AB} \). Twin projectors will also be called twin events.

In [7] twin observables were called physical twins, and also algebraic twins, were mentioned. They were defined by \( O_A |\Psi\rangle_{AB} = O_B |\Psi\rangle_{AB} \).

**Remark 10.** Introducing \( P_s^{\tau} \equiv \sum_m P_s^m, \quad s = A, B \), Lemma 1 implies \( P_s^{\tau} Q_s = Q_s \), i.e., that \( Q_s \) is a subprojector of \( P_s^{\tau} \). \( Q_s \leq P_s^{\tau} \), or equivalently \( \mathcal{R}(Q_s) \subseteq \mathcal{R}(P_s^{\tau}) \), \( s = A, B \). Further, we can define \( P_s^{\tau} \equiv \sum_s P_s^{\tau} \), \( s = A, B \) as the sum of all nullifying eigen-projectors: \( P_s^{\tau} \equiv \sum_k P_A^k \), \( k \neq P_A^k \), \( \mathcal{R}(P_s^{\tau}) \subseteq \mathcal{R}(Q_s) \), \( s = A, B \). Thus, in the case \( k = A \), \( k \neq P_A^k \), \( \mathcal{R}(P_s^{\tau}) \subseteq \mathcal{R}(Q_s) \), \( s = A, B \). Further, it follows that \( \forall k' : \quad P_k' \subseteq P_s^{\tau} \), \( s = A, B \).

**Proposition 1.** The corresponding results \( a_m \) and \( b_m \) of subsystem measurements of twin observables are equally probable and ideal measurement causes equal change of the bipartite state:

\[
\forall m : \quad \langle \Psi|_{AB} P_A^m |\Psi\rangle_{AB} = \langle \Psi|_{AB} P_B^m |\Psi\rangle_{AB}, \tag{41}
\]

\[
|\Psi\rangle_{AB} \langle \Psi|_{AB} \rightarrow \sum_m P_A^m |\Psi\rangle_{AB} \langle \Psi|_{AB} P_B^m
\]

\[
= \sum_m P_A^m P_B^m |\Psi\rangle_{AB} \langle \Psi|_{AB} P_B^m. \tag{42}
\]
Proof. Follows directly from Definition 6 and 40. □

Theorem 6. If $O_A$ and $O_B$ are twin operators (cf Definition 6), then each of their non-nullifying eigen-projects $P_{m}^{O}$, $s = A, B$ commutes with the corresponding reduced density operator

$$\forall m: \{P_{m}^{O_A}, \rho_A\} = 0, \quad (43)$$

$$\forall m: \{P_{m}^{O_B}, \rho_B\} = 0. \quad (44)$$

Proof. Straightforward evaluation, utilizing 40 and both partial-trace rules from 8.2 gives:

$$P_{m}^{O_A} \rho_A = \rho_A \text{tr}_B(\langle \Psi_{AB} | \Psi_{AB} \rangle)$$

$$= \text{tr}_B \left( (P_{m}^{O_B} | \Psi_{AB} \rangle \langle \Psi_{AB} | P_{m}^{O_B} \rangle \right)$$

$$= \text{tr}_B \left( (P_{m}^{O_B} | \Psi_{AB} \rangle \langle \Psi_{AB} | P_{m}^{O_B} \rangle \right)$$

$$= (\text{tr}_B (| \Psi_{AB} \rangle \langle \Psi_{AB} |)) P_{m}^{O_A}$$

$$= \rho_A P_{m}^{O_A}$$

The symmetrical claim is proved similarly. □

We now state and prove (for the reader’s convenience) a basic claim of quantum mechanics that is crucial for our further development of the correlated subsystem picture (elaborated in the preceding section).

Lemma 2. Let $O = \sum_k \epsilon_k P_k$ and $\bar{O} = \sum_{\bar{k}} \epsilon_{\bar{k}} \bar{P}_1$ be two commuting Hermitian operators (each with a purely discrete spectrum) in spectral form. Then also

$$\forall k, \forall \bar{k}: \{P_k, \bar{P}_\bar{k}\} = 0. \quad (45)$$

Proof. Let $| k, \epsilon_k \rangle$ be a complete orthonormal eigenbasis of $O$: $\forall k, \epsilon_k : | k, \epsilon_k \rangle = \epsilon_k | k, \epsilon_k \rangle$. Then $O (\bar{O} | k, \epsilon_k \rangle) = (\bar{O} | k, \epsilon_k \rangle = \epsilon_k (\bar{O} | k, \epsilon_k \rangle)$. Hence, $P_k (\bar{O} | k, \epsilon_k \rangle) = (\bar{O} | k, \epsilon_k \rangle = \bar{O}P_k | k, \epsilon_k \rangle$. Further, for $k \neq \bar{k}$, $P_k (\bar{O} | k', \epsilon_{k'} \rangle = P_k P_{\bar{k'}} (\bar{O} | k', \epsilon_{k'} \rangle) = 0 = \bar{O}P_k | k', \epsilon_{k'} \rangle$. Thus, $\forall k, \bar{k}: \{P_k, \bar{P}_{\bar{k}}\} = 0$. Applying this result to the last commutation itself, one finally obtains $\forall k, \bar{k}, k': \{P_k, \bar{P}_{\bar{k}}\} = 0$ as claimed. □

Definition 7. Let $O_B = \sum_k \epsilon_k P_k^B$ be a nearby-subsystem observable that commutes with the corresponding reduced density operator $\rho_B$ of a given bipartite state vector $| \Psi_{AB} \rangle$. Also, let us index the non-nullifying eigen-projects of $O_B$ by $m$. Then, according to Lemma 2, each eigen-projector $P_{m}^{O_B}$ of $O_B$ commutes with $Q_c^B$, the null-projector of $\rho_B$ (cf Remark 10), because it is also the eigen-projector of $\rho_B$ corresponding to its zero eigenvalue. This implies that it also commutes with $Q_B^c$ because the latter is ortho-complementary to $Q_B^c$. Hence, for each value of $m$, we can define the minimal sub-projector $P_{m}^{B_{min}}$ that acts on $| \Psi_{AB} \rangle$ equally $\rho_B$.

$$\forall m: \{P_{m}^{B_{min}}, | \Psi_{AB} \rangle \} = 0,$$

$$\forall m: \{P_{m}^{B_{min}}, \rho_B \} = 0. \quad (46)$$

Naturally,

$$\forall m: \{P_{m}^{B_{min}} = P_B Q_B = Q_B P_{m}^{B_{min}} Q_B. \quad (47)$$

Finally, we can define

$$O_B^{min} = \sum_{m} a_{m} P_{m}^{B_{min}} \quad (48)$$

and call it the minimal part of $O_B$.

Proposition 2. If $O_B = \sum_k \epsilon_k P_k^B$ commutes with $\rho_B$, then the corresponding minimal operator $O_B^{min}$ can be obtained as follows:

$$O_B^{min} \equiv O_B Q_B. \quad (49)$$

Proof. We write $O_B = (\sum_{m} a_{m} P_{m}^{B_{min}}) + \sum_{k'} \epsilon_{k'} P_{k'}^B$. Here by $P_{k'}^B$ are denoted the nullifying eigen-projects of $O_B$ (cf Remark 10). Then 47, 48, and Remark 10 imply

$$O_B Q_B = \left( \sum_{m} a_{m} P_{m}^{B_{min}} + \sum_{k'} \epsilon_{k'} P_{k'}^B \right) Q_B$$

$$= \left( \sum_{m} a_{m} P_{m}^{B_{min}, m} + \sum_{k'} \epsilon_{k'} P_{k'}^B Q_B \right) Q_B$$

$$= \sum_{m} a_{m} P_{m}^{B_{min}} = O_B^{min}. \quad (50)$$

Remark 11. Commutation 44 and Remark 10 which claims that $Q_B = Q_B \sum_{m} P_{m}^{B_m}$, and since $\forall j : \sum_{m} P_{m}^{B_m} = Q_B$, the former relation implies $\forall j : \sum_{m} P_{m}^{B_{min}} = Q_B$, in conjunction with 47, lead to the following spectral operator decomposition:

$$\rho_B = \sum_{j} \rho_B Q_B = \sum_{m} a_{m} P_{m}^{B_{min}}. \quad (50)$$

or in terms of the corresponding subspaces

$$R(Q_B) = \sum_{j} \sum_{m} \left( R(Q_B) \cap R(P_{m}^{B_{min}}) \right). \quad (51)$$

Naturally, the RHS of 50 may contain zero operator terms, and on the RHS of 51 may appear corresponding zero subspaces.
Remark 12. As it is well known, the commutation relations (44) and (43) imply that there exist common eigen-bases of $\rho_B$ and $O_B^{\text{min}}$ in $\mathcal{R}(Q_B)$ as well as of $\rho_A$ and $O_A^{\text{min}}$ in $\mathcal{R}(Q_A)$. We are primarily interested in the former. Let by $(jm)'$ be denoted a pair of indices for which $Q_B^{j}O_B^{\text{min},m} \neq 0$. We introduce a third index $q_{(jm)'}$ to enumerate the orthonormal vectors in the corresponding non-zero subspaces $\mathcal{R}(Q_B^j) \cap \mathcal{R}(P_B^{\text{min},m})$.

Remark 13. The decomposition without zero terms is

$$Q_B = \sum_{(jm)'} Q_B^{j}P_B^{\text{min},m}$$

$$= \sum_{(jm)'} \sum_{q_{(jm)'}} |(jm)'q_{(jm)'}\rangle_B \langle (jm)'q_{(jm)'}|_B. \quad (52)$$

Definition 8. Expanding a given bipartite state $|\Psi\rangle_{AB}$ in the subsystem sub-basis appearing in (52), we obtain, what we call the twin-correlated canonical Schmidt decomposition:

$$|\Psi\rangle_{AB} = \sum_{(jm)'} \sum_{q_{(jm)'}} r_{(jm)'}^{j} |(jm)'q_{(jm)'}\rangle_A \langle (jm)'q_{(jm)'}|_B, \quad (53)$$

with

$$\forall (jm)'q_{(jm)'}: |(jm)'q_{(jm)'}\rangle_A = (U_a |(jm)'q_{(jm)'}\rangle_B)_A \quad (54)$$

(cf the correlated canonical Schmidt decomposition (14)). If the role of the correlation operator $U_a$ is not made explicit in (53) or, equivalently, if (54) is not joined to it, i.e., (53) itself (as it stands) we call twin-adapted canonical Schmidt decomposition.

As a consequence of (54), one has

$$\forall (jm)': Q_A^{j}P_A^{\text{min},m} = U_a (Q_B^{j}P_B^{\text{min},m}) U_a^{-1}Q_A. \quad (55)$$

$$\forall (jm)': Q_B^{j}P_B^{\text{min},m} = U_a^{-1} (Q_A^{j}P_A^{\text{min},m}) U_a Q_B. \quad (56)$$

The following result is another consequence of (54).

Theorem 7. If $O_A^{\text{min}}$, $s = A, B$ are minimal twin observables for $|\Psi\rangle_{AB}$, then

$$\forall m: P_B^{\text{min},m} = \sum_{j} \sum_{q_{(jm)'}} |(jm)'q_{(jm)'}\rangle_B \langle (jm)'q_{(jm)'}|_B. \quad (57)$$

$$\forall m: P_A^{\text{min},m} = \sum_{j} \sum_{q_{(jm)'}} |(jm)'q_{(jm)'}\rangle_A \langle (jm)'q_{(jm)'}|_A. \quad (58)$$

where the prime sum over $j$ denotes restriction to those terms in which $j$ with the given $m$ gives a non-zero subspace in (51). Further, 

$$\forall m: P_A^{\text{min},m} = U_a P_B^{\text{min},m} U_a^{-1}Q_A. \quad (59)$$

$$\forall m: P_B^{\text{min},m} = U_a^{-1} P_A^{\text{min},m} U_a Q_B. \quad (60)$$

Relations (52,54,55,56,59,60) constitute the twin-correlated subsystem picture. It is the third and most intricate layer of the underlying foundation of Schmidt decomposition. It completes the correlated subsystem picture (cf (33,34,35,36) and (37,38,39)) by the pair of minimal twin observables $O_A^{\text{min}}$, $O_B^{\text{min}}$, and the latter picture was, in turn, a completion of the subsystem picture (cf (21,22,23)) by the correlation operator.

The original articles [1-7], which have been reviewed here, did not present the third layer of foundation sufficiently precisely and transparently. Therefore, a completely new derivation is given in this section.

One may wonder if there may exist two different observables $O_A$ and $\tilde{O}_A$ both twins with one and the same opposite-subsystem observable $O_B$ in a given $|\Psi\rangle_{AB}$.

Proposition 3. If $O_A$ and $\tilde{O}_A$ are both twin observables with one and the same opposite-subsystem observable $O_B$, then

$$O_A^{\text{min}} = \tilde{O}_A^{\text{min}}. \quad \square$$

Proof. Follows immediately from (59).

Remark 14. One can have $O_A \neq O_A$ if $\rho_A$ is singular, and then the only difference is in the terms $P_A^{m}Q_A^{c}$, where $Q_A^{c} \equiv I - Q_A$ is the null-space projector of $\rho_A$. The operators $P_A^{m}Q_A^{c}$ are sub-projectors of $Q_A^{c}$. These terms in the projectors $P_m = P_A^{m}Q_A + P_A^{m}O_A$ nullify $|\Psi\rangle_{AB}$. Taking $O_A$ or $\tilde{O}_A$ means no difference for the entanglement in $|\Psi\rangle_{AB}$ because the latter takes place between $\mathcal{R}(Q_B)$ and $\mathcal{R}(Q_A)$ (with no regard to the null spaces of $\rho_s$, $s = A, B$).

The minimal form of a discrete subsystem Hermitian operator that commutes with the corresponding reduced density operator of the given bipartite state vector $|\Psi\rangle_{AB}$ (cf Definition 7 and Proposition 2) was not defined explicitly in previous work. Hence, the presentation there of this last and most intricate form of Schmidt decomposition and its underlying entanglement foundation was not so transparent. In the present exposition there is new insight and there are new results.

One may wonder which observables $O_B$ do have a twin observable in the given bipartite state.

Theorem 8. Let $|\Psi\rangle_{AB}$ be any bipartite state vector and $O_B \equiv \sum b_i P_B^i$ be an observable for the nearby subsystem B. $O_B$ has a twin observable $O_A$ if and only if

(A) $O_B$, as an operator, commutes with the corresponding reduced density operator: $\rho_B = \text{tr}_A (|\Psi\rangle_{AB}\langle\Psi|_{AB})$, $[O_B, \rho_B] = 0$. Then there exists a unique minimal twin observable $O_A^{\text{min}}$.

(B) If the bipartite state is expanded in an eigenbasis $|l,q\rangle_B : \forall l,|l,q\rangle_B$ of $O_B$

$$|\Psi\rangle_{AB} = \sum_{l} \sum_{q} |l,q\rangle_A \langle l,q|_B.$$
the expansion coefficients satisfy the orthogonality conditions: \( \langle q_l | q_{l'} \rangle_A = 0 \) whenever \( l \neq l' \).

\[ \text{Proof.} \] (A) follows in a straightforward way from (52), for which the commutation of \( O_B \) with \( \rho_B \) is sufficient (cf Lemma 2). Then, with the help of (54), the eigenprojectors \( P_{A,m}^{\text{min}} \) are defined by (58).

(B) Obvious. \( \square \)

**Remark 15.** One may further wonder whether, if \( [O_B, \rho_B] = 0 \) and \( |\Psi_{AB}\rangle \) is expanded in the common eigenbasis of these two operators, it can happen that one does not obtain a twin-adapted Schmidt decomposition of the bipartite state. The answer is: No, it cannot happen! One necessarily obtains a twin-adapted Schmidt decomposition in terms of \( O_A^{\text{min}} \) and \( O_A^{\text{min}} = \sum_m a_m (U_B P_{B,m}^{\text{min}} U_B^{-1}) Q_A \) (cf Definition 7 and Proposition 2), where \( Q_A = \sum_i (U_B |i\rangle_B A \langle i| B U_B^{-1}) A \) and the eigenvalues \( \{a_m : \forall m\} \) are arbitrary distinct non-zero real numbers (they are irrelevant).

One may also wonder if there exits a bipartite state that has no twin observables. The answer is again: No! Formally, the reduced density operators \( \rho_s, s = A, B \) themselves are twin operators, as obvious in the canonical Schmidt decomposition (cf (14)). They, or any other Hermitian operators with the same eigen-projectors, can be viewed as minimal (in the sense of Definition 7) twin observables.

### 6 Distant measurement and Einstein–Podolsky–Rosen states

The approach based on correlation operator as an entanglement entity furnished a specific view of a historically important notion: the Einstein–Podolsky–Rosen paradox.

#### 6.1 Distant measurement

Let any bipartite state vector \( |\Psi_{AB}\rangle \) be given, and let \( O_A = \sum_m a_m P_{A,m} \) and \( O_B = \sum_m b_m P_{B,m} \) be twin observables in it (cf Definition 6). The relations \( O_A^\dagger |\Psi_{AB}\rangle = 0 = O_B^\dagger |\Psi_{AB}\rangle \) are valid.

The change of state in non-selective (9) (when no definite-result sub-ensemble is selected) ideal measurement (10) is

\[
|\Psi_{AB}\rangle\langle\Psi_{AB}| \rightarrow \sum_m P_{B,m} |\Psi_{AB}\rangle\langle\Psi_{AB}| P_{B,m}
\]  

(61)

can be caused, in principle, by *direct measurement* on the nearby subsystem \( B \). Further, this composite-system change of state implies the ideal-measurement change of state

\[
\rho_B \rightarrow \sum_m P_{B,m} \rho_B P_{B,m}
\]  

(62)
on the nearby subsystem \( B \) (obtained when the partial trace over subsystem \( A \) is taken). In this case, by the very definition of subsystem measurement, there is no interaction between the measuring instrument and the distant subsystem \( A \).

**Proposition 4.** In spite of lack of interaction with the distant subsystem \( A \) in the composite-system change-of-state (61), this subsystem nevertheless undergoes the ideal-measurement change

\[
\rho_A \rightarrow \sum_m P_{A,m} \rho_A P_{A,m}
\]

(63)
due to the entanglement in \( |\Psi_{AB}\rangle \).

\[ \text{Proof.} \] The change is implied by (42), and seen by taking the partial trace over subsystem \( B \). \( \square \)

**Definition 9.** Change (63) is said to be due to distant measurement (on the distant subsystem \( A \)) (2).

**Remark 16.** It has been proved in [13] that the ideal change (63) on the distant subsystem \( A \) can be caused by any exact subsystem measurement of the twin observable on the nearby subsystem \( B \). The entanglement in \( |\Psi_{AB}\rangle \) does not distinguish, as far as influencing the distant subsystem is concerned, ideal measurement, non-ideal non-demolition (synonyms: predictive, first-kind, repeatable) measurement and even demolition (synonyms: retrodictive, second-kind, non-repeatable) measurements on the nearby subsystem as long as they are exact measurements.

**Remark 17.** Distant measurement is always ideal measurement. Moreover, the non-selective version does not change the state of the opposite distant subsystem \( A \) at all. Namely, on account of the commutation \( \forall m : \ [P_{A,m}^n, \rho_A] = 0 \) (cf (43) in Theorem 6), one has

\[
\sum_m P_{A,m}^n \rho_A P_{A,m}^n = \sum_m \rho_A P_{A,m}^n = \rho_A \sum_m P_{A,m}^n = \rho_A \left( \sum_m P_{A,m}^n + \sum_k P_{A,k}^k \right) = \rho_A
\]  

(cf Remark 10). Hence, only the selective version of distant measurement may change the distant state.
Remark 18. One may further write

$$\rho_A = \sum_m P^m_A \rho A_P^m$$

and view mathematically $\rho_A$ as an orthogonal mixture of substates (selected subensembles empirically) each predicting a definite value of $O_A$. The selective distant measurements reduce $\rho_A$ to the corresponding state term. Since non-selective measurement is actually the entirety of all selective measurements, the true physical meaning of the change (63) is in making the term states available to selective measurement.

Remark 19. Let $\rho_A = \sum_m w_m P^m_A$ ($w_m$ being statistical weights: $w_m \geq 0$, $\sum_m w_m = 1$) be an arbitrary orthogonal decomposition of the distant state $\rho_A$. It can be realized by non-selective distant measurement caused by a suitable subsystem measurement on the nearby subsystem $A$. Namely, the range projectors $P^m_A$ of the term states $\rho_A^m$ are orthogonal. Defining $\forall m: P^m_A = Q^m_A$ and $O_A = \sum_m a_m P^m_A$, any term measure on $O_A$ of the opposite subsystem. Its measurement gives rise to the distant measurement of $O_A$, and hereby to the orthogonal state decomposition that we have started with.

Let us for the moment forget about twin observables, and consider more general ones.

Remark 20. Non-selective measurement of any nearby-subsystem observable $O_B = \sum_i b_i P^l_B$ gives rise to a distant state decomposition

$$\rho_A = \text{tr}_B (|\Psi\rangle_A \langle \Psi|_A)$$

and view mathematically $\rho_A$ as an orthogonal mixture of substates (selected subensembles empirically) each predicting a definite value of $O_A$. The selective distant measurements reduce $\rho_A$ to the corresponding state term. Since non-selective measurement is actually the entirety of all selective measurements, the true physical meaning of the change (63) is in making the term states available to selective measurement.

Remark 19. Let $\rho_A = \sum_m w_m P^m_A$ ($w_m$ being statistical weights: $w_m \geq 0$, $\sum_m w_m = 1$) be an arbitrary orthogonal decomposition of the distant state $\rho_A$. It can be realized by non-selective distant measurement caused by a suitable subsystem measurement on the nearby subsystem $A$. Namely, the range projectors $P^m_A$ of the term states $\rho_A^m$ are orthogonal. Defining $\forall m: P^m_A = Q^m_A$ and $O_A = \sum_m a_m P^m_A$, any term measure on $O_A$ of the opposite subsystem. Its measurement gives rise to the distant measurement of $O_A$, and hereby to the orthogonal state decomposition that we have started with.

Let us for the moment forget about twin observables, and consider more general ones.

Remark 20. Non-selective measurement of any nearby-subsystem observable $O_B = \sum_i b_i P^l_B$ gives rise to a distant state decomposition

$$\rho_A = \text{tr}_B (|\Psi\rangle_A \langle \Psi|_A)$$

and view mathematically $\rho_A$ as an orthogonal mixture of substates (selected subensembles empirically) each predicting a definite value of $O_A$. The selective distant measurements reduce $\rho_A$ to the corresponding state term. Since non-selective measurement is actually the entirety of all selective measurements, the true physical meaning of the change (63) is in making the term states available to selective measurement.

Remark 19. Let $\rho_A = \sum_m w_m P^m_A$ ($w_m$ being statistical weights: $w_m \geq 0$, $\sum_m w_m = 1$) be an arbitrary orthogonal decomposition of the distant state $\rho_A$. It can be realized by non-selective distant measurement caused by a suitable subsystem measurement on the nearby subsystem $A$. Namely, the range projectors $P^m_A$ of the term states $\rho_A^m$ are orthogonal. Defining $\forall m: P^m_A = Q^m_A$ and $O_A = \sum_m a_m P^m_A$, any term measure on $O_A$ of the opposite subsystem. Its measurement gives rise to the distant measurement of $O_A$, and hereby to the orthogonal state decomposition that we have started with.

Let us for the moment forget about twin observables, and consider more general ones.

Remark 20. Non-selective measurement of any nearby-subsystem observable $O_B = \sum_i b_i P^l_B$ gives rise to a distant state decomposition

$$\rho_A = \text{tr}_B (|\Psi\rangle_A \langle \Psi|_A)$$

and view mathematically $\rho_A$ as an orthogonal mixture of substates (selected subensembles empirically) each predicting a definite value of $O_A$. The selective distant measurements reduce $\rho_A$ to the corresponding state term. Since non-selective measurement is actually the entirety of all selective measurements, the true physical meaning of the change (63) is in making the term states available to selective measurement.

Remark 19. Let $\rho_A = \sum_m w_m P^m_A$ ($w_m$ being statistical weights: $w_m \geq 0$, $\sum_m w_m = 1$) be an arbitrary orthogonal decomposition of the distant state $\rho_A$. It can be realized by non-selective distant measurement caused by a suitable subsystem measurement on the nearby subsystem $A$. Namely, the range projectors $P^m_A$ of the term states $\rho_A^m$ are orthogonal. Defining $\forall m: P^m_A = Q^m_A$ and $O_A = \sum_m a_m P^m_A$, any term measure on $O_A$ of the opposite subsystem. Its measurement gives rise to the distant measurement of $O_A$, and hereby to the orthogonal state decomposition that we have started with.

Let us for the moment forget about twin observables, and consider more general ones.

Remark 20. Non-selective measurement of any nearby-subsystem observable $O_B = \sum_i b_i P^l_B$ gives rise to a distant state decomposition

$$\rho_A = \text{tr}_B (|\Psi\rangle_A \langle \Psi|_A)$$

and view mathematically $\rho_A$ as an orthogonal mixture of substates (selected subensembles empirically) each predicting a definite value of $O_A$. The selective distant measurements reduce $\rho_A$ to the corresponding state term. Since non-selective measurement is actually the entirety of all selective measurements, the true physical meaning of the change (63) is in making the term states available to selective measurement.

Remark 19. Let $\rho_A = \sum_m w_m P^m_A$ ($w_m$ being statistical weights: $w_m \geq 0$, $\sum_m w_m = 1$) be an arbitrary orthogonal decomposition of the distant state $\rho_A$. It can be realized by non-selective distant measurement caused by a suitable subsystem measurement on the nearby subsystem $A$. Namely, the range projectors $P^m_A$ of the term states $\rho_A^m$ are orthogonal. Defining $\forall m: P^m_A = Q^m_A$ and $O_A = \sum_m a_m P^m_A$, any term measure on $O_A$ of the opposite subsystem. Its measurement gives rise to the distant measurement of $O_A$, and hereby to the orthogonal state decomposition that we have started with.

6.2 Einstein–Podolsky–Rosen states

Definition 10. If a bipartite state vector $\langle \Psi \rangle_{AB}$ allows distant measurement of two mutually incompatible observables (non-commuting operators) $O_A$ and $\bar{O}_A$, then we say that we are dealing with an Einstein–Podolsky–Rosen state (14).

Theorem 9. A state $\langle \Psi \rangle_{AB}$ is an Einstein–Podolsky–Rosen state if and only if at least one of the positive eigenvalues $r_j$ of $\rho_B = \text{tr}_A \langle \Psi \rangle_{AB}$ (if $\rho_{AB}$ is degenerate, i.e., has multiplicity at least two. This amounts to some repetition in the expansion coefficients $r_j^2$ in the canonical Schmidt decomposition (14).

Proof. Considering the twin-correlated subsystem picture (cf (32)), it is straightforward to see that if at least one non-zero subspace $R(\rho^m_{min})$, indexed by $(jm)'$, is two or more dimensional, then, and only then, one can have two different eigen-bases $\{(jm)' q_{(jm)'} : \forall (jm)', q_{(jm)'}\}$ and $\{(jm)' r_{(jm)'} : \forall (jm)', r_{(jm)'}\}$ so that the correlation operator $U_B$ can determine the corresponding (also different) eigen-bases $\{(U_a |(jm)' q_{(jm)'}_B : \forall (jm)', q_{(jm)'}\}$ and $\{(U_a |(jm)' r_{(jm)'}_B : \forall (jm)', r_{(jm)'}\}$ of distant incompatible minimal observables $O_{min}^A$ and $\bar{O}_A^A$.

The original Einstein–Podolsky–Rosen paper (14) discussed the two-particle state $\langle \Psi \rangle_{AB}$ defined by a fixed eigenvalue $\beta$ of the total linear momentum

$$\hat{r}_A + \hat{r}_B \langle \Psi \rangle_{AB} = \beta \langle \Psi \rangle_{AB},$$

where $\hat{r}_A$ and $\hat{r}_B$ are the particle linear momentum vector operators, and a fixed eigenvalue $\beta$ of the relative radius vector

$$\hat{r}_{AB} \langle \Psi \rangle_{AB} = \beta \langle \Psi \rangle_{AB}.$$

For clarity, the operators are denoted with hats to distinguish them from fixed eigenvalues of vectors.
The discussion went essentially as follows: If one performs a position measurement of the nearby particle B and obtains the value \( \hat{r}_B \), then \textit{ipso facto} the distant particle A acquires without interaction, via distant measurement, the value \( \hat{r}_A = \hat{r} + \hat{r}_B \). On the other hand, as an alternative, one can perform a linear momentum measurement of the nearby particle with a result \( \hat{p}_B \) and also obtain, by distant measurement, a definite value of the linear momentum \( \hat{p}_A = \hat{p} - \hat{p}_B \) of the distant particle without interaction.

The authors found this paradoxical in view of the contention that quantum mechanics was complete, and \( |\Psi\rangle_{AB} \) did not contain the mentioned values obtained without interaction (with a \textit{spooky action} as Einstein liked to say), and, moreover, it could not contain the two incompatible values simultaneously as valid for one and the same pair of particles because position and linear momentum are incompatible.

As a slight formal objection, one may notice that the mentioned fixed values of the total linear momentum and the relative radius vector belong to continuous spectra, and the corresponding state is of infinite norm (a generalized vector). Bohm pointed out \([15]\) that one can easily escape this formal difficulty by taking for \( |\Psi\rangle_{AB} \) not the original Einstein–Podolsky–Rosen state described above, but the well known singlet two-particle spin state

\[
|\Psi\rangle_{AB} = \frac{1}{\sqrt{2}} (|\uparrow\rangle_A |\downarrow\rangle_B - |\downarrow\rangle_A |\uparrow\rangle_B),
\]

where \( \uparrow \) and \( \downarrow \) denote spin-up and spin-down respectively along any axis. For the same \( |\Psi\rangle_{AB} \) given by \((64)\) one can choose either the z-axis or the x-axis, and make an argument in complete analogy with the Einstein–Podolsky–Rosen state described above. Then it is fully within the quantum formalism.

It appears that the authors of \([14]\) consider that the paradoxical nature of an Einstein–Podolsky–Rosen state lies in its contradiction with completeness of the quantum-mechanical description of an individual bipartite system (which was claimed by the Copenhagen interpretation). Actually, this contradiction may be viewed to be present in every entangled bipartite state \( |\Psi\rangle_{AB} \) because it has at least one pair of twin observables (cf the final parts of the preceding section). They make possible selective \textit{distant measurement}, and it creates (or finds) a definite value of the distant twin observable that was not a sharp value in \( |\Psi\rangle_{AB} \).

One can find articles in the literature in which all entangled bipartite states are called Einstein–Podolsky–Rosen states. It might be due to realization of this point. The more so, since Schrödinger’s view of distant correlations, discussed in the next subsection, brings home this point.

Let us return to the singlet state given by \((64)\). \( |\Psi\rangle_{AB} \) is a suitable realization of Mermin’s Ithaca mantra “the correlations, not the correlata” \([19]\). Let us choose to measure the spin component of the nearby particle B along the z-axis. Let further the \textit{measuring instrument} be in the initial or ready-to-measure state \( |0\rangle_m \), and the \textit{experimenter} in the ready-to-watch the result state \( |0\rangle_e \). The entire four-partite system is in the initial state

\[
|\Psi\rangle_{AB} |0\rangle_m |0\rangle_e = \frac{1}{\sqrt{2}} (|\uparrow\rangle_A |\downarrow\rangle_B - |\downarrow\rangle_A |\uparrow\rangle_B) \otimes |0\rangle_m |0\rangle_e.
\]

At the end of the measurement, the four-partite system is, e.g., in the state

\[
|\uparrow\rangle_A |\downarrow\rangle_B \otimes |z, \uparrow_A, \downarrow_B \rangle_m |z, \uparrow_A, \downarrow_B \rangle_e,
\]

where \( |z, \uparrow_A, \downarrow_B \rangle_m \) is the state of the measuring instrument in which the so-called \textit{pointer position} show the results \( \downarrow \) for subsystem B, and \( \uparrow \) for the distant subsystem A, and \( |z, \uparrow_A, \downarrow_B \rangle_e \) is the analogous state of the experimenter in which the counterpart of the pointer position is the corresponding contents of consciousness.

Einstein, Podolsky and Rosen were troubled by the idea that, in transition from \((65)\) to \((66)\), the result \( \uparrow_A \) was brought about in a distant action without interaction (a \textit{spooky action}), which could not be reconciled with basic physical ideas that reigned outside quantum mechanics. It seems to me that the father of relativity ideas in physics has fallen victim to the Bohrian (or Copenhagen) suggestion that \((66)\) describes \textit{absolute reality}. But no wonder; this was more than two decades before Everett’s relative-state ideas appeared \([16]\).

In previous work \([17, \S 7C]\) I have adopted, what I call humorously, a \textit{pocket edition} of Everett’s relative-state interpretation of quantum mechanics. (I was sticking to the idea of a laboratory, forgetting about parallel worlds in a multiverse \([13]\).) I have called the approach \textit{relative reality of unitarily evolving states} (RRUES).

Let me apply RRUES to the above direct measurement on subsystem B, and to the simultaneous distant measurement on subsystem A.

If the unitary evolution of the system does not change spin projections, then the above initial four-partite state \((65)\) evolves into the state

\[
|\Psi\rangle_{AB} = \frac{1}{\sqrt{2}} (|\uparrow\rangle_A |\downarrow\rangle_B \otimes |z, \uparrow_A, \downarrow_B \rangle_m |z, \uparrow_A, \downarrow_B \rangle_e,
\]

Here the state \((66)\) is one of the components, one of the \textit{branches} in Everett’s terminology. The point is that the result \( |z, \uparrow_A, \downarrow_B \rangle_m \) is \textit{relative to the state} \( |z, \uparrow_A, \downarrow_B \rangle_e \) of the \textit{observer}. Reality of the measurement results are only relative to the branch in which the \textit{observer} finds himself. I think this is a suitable realization of Mermin’s Ithaca mantra “the correlations, not the correlata” \([19]\).
One might object that replacing absolute reality of the description of a quantum state by its relative reality is unacceptable. It is well known that for some time the same objection was raised when Einstein replaced absolute motion by relative motion. Nowadays we find no difficulty with it.

Thus, in RRUES there is no spooky action in distance without interaction. One might wonder if RRUES as a new term is justified, when it is pure Everett’s relative-state theory. Actually, the new term serves the sole purpose of emphasizing (via the two R’s) the new relativity idea introduced by Everett in his seminal work.

As it was pointed out in Remark 20 any non-selective or selective measurement on the nearby subsystem B gives rise to distant state decomposition or distant state preparation respectively on the distant subsystem A. One can easily see that two choices of distinct non-selective direct measurements on subsystem B can induce state decompositions on subsystem A that do not have a common continuation (finer decomposition), and hence are actually incompatible. This might be viewed as a kind of a generalized Einstein–Podolsky–Rosen phenomenon.

Realizations of Einstein–Podolsky–Rosen states in thought and real experiments are pointed out in the second and third passage of §7 below.

6.3 Schrödinger’s steering

Relation (29) introduces explicitly the correlation operator into investigations of the effects on the distant subsystem A caused by measurement performed on the nearby subsystem B. This enabled the Belgrade school to have an original angle and elaborate Schrödinger’s approach to distant correlations.

The role of the correlation operator in studying distant nearby-subsystem measurement effects has thus led to the articles [1, 20, 21]. But they were written partly in the antilinear Hilbert–Schmidt operators approach, which has been abandoned in this review.

For the reader’s convenience we rewrite (and renumberate) relation (29):

$$\langle \Psi \rangle_{AB} = \sum_{n'} (U_{a} \rho_{B}^{\frac{1}{2}} |n'\rangle_{B})_{A} \otimes |n'\rangle_{B}. \quad (68)$$

Inserting $U_{a}^{-1} U_{a} Q_{B} = Q_{B}$ between $\rho_{B}^{\frac{1}{2}}$ and $|n'\rangle_{B}$ in (68), which can be done because $\rho_{B}^{\frac{1}{2}} Q_{B} = Q_{B}$, one obtains the equivalent formula

$$\langle \Psi \rangle_{AB} = \sum_{n'} \rho_{A}^{\frac{1}{2}} (U_{a} Q_{B} |n'\rangle_{B})_{A} \otimes |n'\rangle_{B} \quad (69)$$

due to $U_{a} \rho_{B}^{\frac{1}{2}} U_{a}^{-1} Q_{A} = \rho_{A}^{\frac{1}{2}}$ (cf (33) etc).

If a nearby-subsystem observable

$$O_{B} = \sum_{n''} b_{n''} |n''\rangle_{B} \langle n''|_{B}, \quad \text{is measured ideally and selectively having, e.g., the result } b_{j} \text{ in mind, then } | \Psi \rangle_{AB} \text{ is hereby converted into the uncorrelated bipartite state } \left(U_{a} \rho_{B}^{\frac{1}{2}} |\bar{n}\rangle_{B}\right)_{A} \otimes |\bar{n}\rangle_{B}. \text{ This implies the fact that the distant subsystem } A \text{ is brought into the state (cf (6))}

$$|\bar{n}\rangle_{A} = \frac{\left(U_{a} \rho_{B}^{\frac{1}{2}} |\bar{n}\rangle_{B}\right)_{A}}{|| \left(U_{a} \rho_{B}^{\frac{1}{2}} |\bar{n}\rangle_{B}\right)_{A}||} = \frac{\left(U_{a} \rho_{B}^{\frac{1}{2}} Q_{B} |\bar{n}\rangle_{B}\right)_{A}}{|| \left(U_{a} \rho_{B}^{\frac{1}{2}} Q_{B} |\bar{n}\rangle_{B}\right)_{A}||}. \quad (70)$$

The fact that $\rho_{B}^{\frac{1}{2}} = \rho_{B}^{\frac{1}{2}} Q_{B}$ is always valid was utilized (cf Corollary 4).

The nearby-subsystem measurement that leads to (70) was called steering by Schrödinger [22, 23] and distant steering in previous work of the present author [20, 21]. It is also called distant preparation of a state. It is part of a distant state decomposition (cf Remark 20) that is brought about by the ideal non-selective measurement of the nearby observable $O_{B}$ mentioned above.

Schrödinger pointed out [22, 23] the paradoxical fact that a skilful experimenter can steer, without any interaction, a distant particle (that is correlated with a nearby one on account of past interactions) into any of a wide set of states.

The basic steering formula (70) makes clear what the physical meaning of the correlation operator $U_{a}$ is. It plays an essential role in determining into which state the distant subsystem is steered. Since this determination takes place jointly with $\rho_{B}$, the physical meaning of $U_{a}$ is much more clear when the action of $\rho_{B}$ is simplified. This is the case when $| \bar{n}\rangle_{B} = |i\rangle_{B}$ (cf (15), i.e., when $\rho_{B} | \bar{n}\rangle_{B} = r_{i} | \bar{n}\rangle_{B}$. Then $\rho_{B}^{\frac{1}{2}}$ amounts in (70) to multiplication with $r_{i}$, and this has no effect on steering; it affects only the probability (see below). Then $| \bar{n}\rangle_{B}$ is steered into the state $\left(U_{a} |i\rangle_{B}\right)_{A} = |\bar{i}\rangle_{A}$. If the eigenvalue $r_{i}$ is degenerate, i.e., if $\mathcal{R}(Q_{B})$ for $r_{j}$ is at least two dimensional, then the action of $U_{a}$ in mapping $\mathcal{R}(Q_{B})$ onto $\mathcal{R}(U_{a} Q_{B} U_{a}^{-1}) = \mathcal{R}(Q_{B})$ (cf the correlated subsystem picture (33, 34, 35, 36) and (37, 38, 39)), is non-trivial. Otherwise, it determines the phase factor of $|\bar{i}\rangle_{A}$.

Viewing all this in analogy with classical probability theory, one can say that the occurrence of $| \bar{n}\rangle_{B} |\bar{n}\rangle_{B}$ is the condition in the conditional probability, which is the state vector given by the LHS of (70).
Remark 22. From (70) follows that all choices of $|\tilde{n}\rangle_B$ that have the same projection in $\mathcal{R}(\rho_B)$ give the same distant state, and if two choices of nearby state vectors differ only by a phase factor, so do the corresponding distant states.

**Proposition 5.** (A) All states $|\phi\rangle_A$ that belong to $\mathcal{R}(\rho_A^{\perp})$ and no other states can be brought about by distant steering.

(B) A given state $|\phi\rangle_A \in \mathcal{R}(\rho_A^{\perp})^{\perp}$ can be steered into, i.e., it can be given rise to by selective direct measurement of $|\tilde{n}\rangle_B|\tilde{n}\rangle_B$ in $|\Psi\rangle_{AB}$ (cf passage below (68)), if and only if

$$0 \neq \frac{Q_B |\tilde{n}\rangle_B}{\|Q_B |\tilde{n}\rangle_B\|} = \frac{\rho_B^{-1/2}U_A^{-1}|\phi\rangle_A}{\|\rho_B^{-1/2}U_A^{-1}|\phi\rangle_A\|}.$$  (71)

**Proof.** (A) follows immediately from relation (69) and (B).

(B) Relation (68) is seen to imply the claim if one has in mind the fact that in $\mathcal{R}(\rho_B)\rho_B^{\perp}$ is non-singular and it maps $\hat{\mathcal{R}}(\rho_B)$ onto $\mathcal{R}(\rho_B^{\perp})$ in a one-to-one way (cf (75)) below.

Proposition 5B implies the Lemma of Hadjisavvas [24]: For any given density operator $\rho$ a state vector $|\phi\rangle$ can appear in a decomposition $\rho = w |\phi\rangle\langle\phi| + \sum_k w_k \rho_k$ (where $w + \sum w_k = 1$, each $\rho_k$ is a density operator, and the sum is finite or countably infinite) if and only if $|\phi\rangle \in \rho^{\perp}$ (let us call it suitability).

That every suitable state vector can appear in a decomposition follows from Proposition 5B by performing purification transforming by isomorphism $\rho$ into $\rho_A = \text{tr}_B(|\Psi\rangle_{AB}\langle\Psi|_{AB})$ in any way (cf Theorem 2), and then taking a basis in $\mathcal{H}_B$ that contains the final state vector in the relation in Proposition 5B. Clearly, this will give a pure-state decomposition of $\rho_A$ in which $|\phi\rangle\langle\phi|$ will appear.

That no state vector outside $\rho^{\perp}$ can appear in a decomposition can be seen by writing down such a decomposition, then using it for purification (cf Theorem 2), and getting into contradiction with Proposition 5.

**Remark 23.** A well-known special case of steering is quantum erasure [25]. For instance, the well-known two-slit interference disappears when linear polarizers, a vertical and a horizontal one, are put on the respective slits [26] because entanglement with the polarization (internal degree of freedom) suppresses the coherence. But a 45° polarization analyzer can restore (or revive) the interference. (The suppressing entanglement is erased.) Here choice of the analyzer is actually choice of the state $|\tilde{n}\rangle_B$ in Proposition 5B.

One should note that steering is not a deterministic operation. As it follows from (68), the state (70) comes about with the probability $p(b_n) = \|\rho_B^{1/2} |\tilde{n}\rangle_B\|^2$ (because a unitary operator does not change the norm). As easily seen, one actually has

$$p(b_n) = \|Q_B |\tilde{n}\rangle_B\| \times \|\rho_B^{1/2} Q_B |\tilde{n}\rangle_B\|^2.$$  (72)

Relation (72) implies that all choices of $|\tilde{n}\rangle_B$ the projections in $\mathcal{R}(\rho_B)$ of which differ only by a phase factor have the same probability.

Since, on account of the positive-eigenvalue eigensubspaces $\mathcal{R}(Q_B)$ of $\rho_B$, one has $Q_B = \sum_j Q_B^j$, and (71) can be further rewritten as

$$p(b_n) = \|Q_B |\tilde{n}\rangle_B\|^2 \times \sum_j r_j \times \|Q_B^j |\tilde{n}\rangle_B\|^2.$$  (74)

**Remark 24.** One can see in (72) that the probability of successful steering (occurrence of $|\tilde{n}\rangle_B |\tilde{n}\rangle_B$) is larger if:

(i) $|\tilde{n}\rangle_B$ has a larger projection in the range $\mathcal{R}(\rho_B)$ (if it is more in the range than in the null space), and

(ii) the projection is more favorably positioned in the range (if it grabs larger eigenvalues $r_j$).

On account of Remark 24, it is practical to restrict oneself to state vectors from the range

$$|\tilde{n}\rangle_B = Q_B |\tilde{n}\rangle_B.$$  (73)

Choice (73) implies

$$p(b_n) = \|Q_B^j |\tilde{n}\rangle_B\|^2 = \sum_j r_j \times \|Q_B^j |\tilde{n}\rangle_B\|^2.$$  (74)

(cf (37)).

In my previous work [21], Lemmata 1–3 give a detailed mathematical account of the fine structure of $\mathcal{R}(\rho_B)$ concerning the action of $\rho_B^{1/2}$. Neither the approach of writing bipartite state vectors in terms of antilinear Hilbert–Schmidt operators that is adopted in the article nor the results of Lemmata 1–3 do I consider physically sufficiently important (at the time of writing this review). Hence it is not reproduced here. All that should be pointed out is that one always has

$$\mathcal{R}(\rho) \subseteq \mathcal{R}(\rho^{1/2}) \subseteq \hat{\mathcal{R}}(\rho),$$  (75)

and if $\dim(\mathcal{R}(\rho)) < \infty$, then one has equality throughout in (75), and if $\dim(\mathcal{R}(\rho)) = \infty$, then both inclusion relations are proper. It is also worth pointing out that the mentioned Lemmata 1–3, unlike the rest of the article, are stated and proved in terms of standard quantum-mechanical arguments.
Remark 25. In case of infinite-dimensional range $\mathcal{R}(\rho_B)$, the distant states in $\tilde{\mathcal{R}}(\rho_A) \ominus \mathcal{R}(\rho_A^\perp)$, where $\ominus$ denotes set-theoretical subtraction (of a subset), are a kind of irrationals concerning steering: one cannot steer the distant subsystem into these states exactly, but one can achieve this arbitrarily closely (because $\mathcal{R}(\rho_A^\perp)$ is dense in $\tilde{\mathcal{R}}(\rho_A)$, cf [75]).

Remark 26. As it was pointed out in Remark 20, one can perform measurement of an incomplete observable $O_B$, i.e., one that has degenerate eigenvalues, on the nearby subsystem and obtain distant state decomposition in the non-selective version, or state preparation in the selective version. In the latter case one has generalized steering, which results, in general, in a mixed state of the distant subsystem.

Schrödinger’s steering has recently drawn much attention. For example, steering was generalized to mixed states in [27]. Asymmetric steering was studied in [28]. See also the review article in [29].

7 Concluding remarks

In the article “On bipartite pure-state entanglement structure in terms of disentanglement” [1], Schrödinger’s disentanglement, i.e., distant state decomposition, as a physical way to study entanglement, is carried one step further with respect to previous work in investigating the qualitative side of entanglement in any bipartite state vector. Distant measurement or, equivalently, distant orthogonal state decomposition from previous work (cf Remark 18 and Remark 19) is generalized to distant linearly independent complete state decomposition both in the non-selective and the selective versions (cf Remark 20). The results are displayed in terms of commutative square diagrams, which show the power and beauty of the physical meaning of the antiunitary correlation operator $U_a$, inherent in any given bipartite state vector $|Ψ_{AB}\rangle$. It is shown that linearly independent distant pure-state preparation, which is caused by selective measurement of an observable $O_B$ on the nearby system that does not commute with its state operator $ρ_B$ (cf Theorem 6), carries the highest probability of occurrence among distant preparations that are not obtained by selective distant measurement.

In the two-part article “On Einstein–Podolsky–Rosen-type entanglement in the experiments of Scully et al.”, “I. The micromaser case and delayed-choice quantum erasure” and “II. Insight in the real random delayed-choice erasure experiment” [17, 50], intricate realizations of Einstein–Podolsky–Rosen states in a thought experiment and a real experiment respectively are discussed.

In the preprint “Quantum correlations in multipartite states. Study based on the Wootters–Mermin theorem” [31] a nice example of an Einstein–Podolsky–Rosen state is given (cf §7, Eq.(15) there).

In the article “The role of coherence entropy of physical twin observables in entanglement” [32], the concept of twin observables for bipartite quantum states is simplified. The relation of observable and state is studied in detail from the point of view of coherence entropy.

In the article “Irrelevance of the Pauli principle in distant correlations between identical fermions” [33], it was shown that the Pauli non-local correlations do not contribute to distant correlations between identical fermions. In distant correlations a central role is played by distant measurement (cf §6, A). A negentropy measure of distant correlations is introduced and discussed. It is demonstrated that distant correlations are necessarily of dynamical origin.

In the article “How to define systematically all possible two-particle state vectors in terms of conditional probabilities” [6], all bipartite state vectors of given subsystems were systematically generated using the state operator $ρ_B$ of the nearby subsystem and the correlation operator $U_a$ (cf §3 and §4).

In the article “Complete Born rule from environment-assisted invariance in terms of pure-state twin unitaries” [5], the concept of twin observables was extended to twin unitaries. It was shown that the latter are the other face of Zurek’s envariance concept.

In the article “Mixed-state twin observables” [34], the twin-observables notion was extended to bipartite mixed states (density operators) $ρ_{AB}$. It was shown that commutation of the twin observables with the corresponding state operators $[O_A, ρ_A] = 0$ and $[O_B, ρ_B] = 0$ are necessary conditions also for mixed states, but these relations are no longer sufficient.

In the article “Hermitian Schmidt decomposition and twin observables of bipartite mixed states” [35], it was shown that every mixed bipartite state (density operator) $ρ_{AB}$ has a Schmidt decomposition in terms of Hermitian subsystem operators. This result is due to the fact that $ρ_{AB}$ is an element in the Hilbert space of all linear Hilbert–Schmidt operators in $\mathcal{H}_A ⊗ \mathcal{H}_B$.

In the article “On statistical and deterministic quantum teleportation” [36], it was shown that use of correlation operators gives insight in teleportation (cf §6, Fig. 2 there).

In the preprint “Delayed twin observables: are they a fundamental concept in quantum mechanics?” [37], the twin-observables concept is generalized to the case when unitary time evolution takes place.

Finally, it is worth reemphasizing that all results presented in [21, 22, 23, 25, 26] apply to every bipartite state vector. For instance, in $|Ψ⟩_{AB}$ subsystem $A$ can be the or-
bital, and subsystem B the spin degree of freedom of one electron, but it can also describe a many-particle system in which A contains some of the particles and B contains the rest.

The correlation operator provides us with a way to comprehend entanglement in a bipartite pure state. It primarily serves to give insight. For most practical purposes the canonical Schmidt decomposition or its stronger form, a twin-adapted canonical Schmidt decomposition, suffice. The correlation operator is implicit in it.

The elaborated systematic and comprehensive analysis presented should, hopefully, enable researchers to utilize Schmidt decomposition as a scalpel in surgery to derive new results. At least I was myself enabled by it to work out a detailed theory of exact quantum-mechanical measurement, which will be presented elsewhere.

8 Appendix

8.1 Partial scalar product

It will be shown that partial scalar product can be defined in three and a half ways, i.e., in three equivalent ways and incompletely in a fourth way.

We still write arbitrary ket or bra vectors with a bar, those without a bar are norm-one vectors. In each of the definitions below, we define the partial scalar product only for norm-one elements of the Hilbert spaces. If the norm of any (or both) of the factors in the product is not one, the final element is, by part of the definition, multiplied by this norm (or by both norms).

Definition 11. (Partial scalar product in terms of subsystem-basis expansion) We define partial scalar product by essentially equating RHS of (7) and RHS of (2). More precisely, for any norm-one element \(|n\rangle_B \in \mathcal{H}_B\) and any norm-one element \(|\Psi\rangle_{AB} \in \mathcal{H}_A \otimes \mathcal{H}_B\), we write:

\[
\langle n |_{B} \langle \Psi |_{AB} \rangle_{A} = \sum_{m} \langle m |_{A} \langle n |_{B} | \Psi \rangle_{AB} \rangle_{A} | m \rangle_{A}.
\]  

(76)

Note that the resulting element in \(\mathcal{H}_A\) is expanded in an arbitrary basis \(|m\rangle_{A} : \forall m\).

Next, we derive two basic properties of partial scalar product from the definition.

Property (i). If the bipartite element is uncorrelated \(|\Psi\rangle_{AB} = |\psi\rangle_A \otimes |\phi\rangle_B\), then partial scalar product reduces to ordinary scalar product:

\[
\langle n |_{B} (|\psi\rangle_A \otimes |\phi\rangle_B)_{A} = (\langle n |_{B} |\phi\rangle_B) \times |\psi\rangle_A.
\]  

(77)

This directly follows from (76).

Property (ii). If the bipartite element is expanded in an absolutely convergent orthogonal series \(|\Psi\rangle_{AB} = \sum_{k} |\Psi_{AB}^{k}\rangle_{A}\) (it can be a double series), then the partial scalar product has the property of extended linearity:

\[
\left(\langle n |_{B} \sum_{k} |\Psi_{AB}^{k}\rangle_{A}\right)_{A} = \sum_{k} \langle n |_{B} |\Psi_{AB}^{k}\rangle_{A}.
\]  

(78)

This also follows directly from (76) if one takes into account the fact that two absolutely converging series (or double series) can exchange order.

One can evaluate the form of the partial scalar product in the representation of arbitrary bases \(|m\rangle_{A} : \forall m\) in \(\mathcal{H}_A\) and \(|q\rangle_{B} : \forall q\) in \(\mathcal{H}_B\):

\[
\langle m |_{A} \langle q |_{B} | \Psi \rangle_{AB} \rangle_{A} = \sum_{m} \langle m |_{A} \langle n |_{B} | \Psi \rangle_{AB} \rangle_{A} | m \rangle_{A}
\]

\[
= \sum_{m} \sum_{q} \langle m |_{A} \langle q |_{B} | \Psi \rangle_{AB} \rangle_{A} | m \rangle_{A}.
\]

Thus, partial scalar product in the representation in the basis \(|q\rangle_{B} : \forall q\) (the q-representation) is

\[
\langle m |_{A} \langle q |_{B} | \Psi \rangle_{AB} \rangle_{A} = \sum_{q} \langle q |_{B} | \Psi \rangle_{AB} \rangle_{A} \times \langle m |_{A} \langle q |_{B} | \Psi \rangle_{AB} \rangle_{A},
\]

(79)

where the asterisk denotes complex conjugation.

The q-representation can be also purely continuous (as the coordinate or linear momentum representations). Then (79) has the form

\[
\langle m |_{A} \langle q |_{B} | \Psi \rangle_{AB} \rangle_{A} = \int \langle q |_{B} | \Psi \rangle_{AB} \rangle_{A} \times \langle m |_{A} \langle q |_{B} | \Psi \rangle_{AB} \rangle_{A} \ dq.
\]

(80)

Definition 12. (Partial scalar product in terms of properties (i) and (ii)) If we assume the validity of the two basic properties from above, then, substituting the suitable general expansion (3) for \(|\Psi\rangle_{AB}\) in \((\langle n |_{B} | \Psi \rangle_{AB} \rangle_{A})\) one recovers (2), and one is back to the subsystem-basis-expansion in Definition 11. Hence, both definitions are equivalent.

Definition 13. (Partial scalar product in representation) We define the partial scalar product by (79). Reading the above derivation of (79) backwards, we recover the subsystem-basis-expansion in Definition 11. Hence, both definitions are equivalent.

Definition 14. (Partial scalar product in terms of the partial trace up to a phase factor) See Proposition 6 in §8.3 below.

Remark 27. The partial scalar product \(|\Psi_{AB}^{k}\rangle_{A}\) can be evaluated also by expressing \(|\Psi_{AB}^{k}\rangle_{A}\) as any (finite) linear combination of tensor products of tensor-factor vectors.
8.2 The partial-trace and its rules

The partial trace

\[ \langle m \rangle_A \tr_B O_{AB} \mid m' \rangle_A = \sum_n \langle m \rangle_A \langle n \mid B \rho_{AB} \mid m' \rangle_A \mid n \rangle \]  

(81)

was explained in von Neumann’s book \([8, p. 425]\) as far as \(O_{AB} = \rho_{AB}\), a composite-system density operator was concerned. The so-called reduced entity (on the LHS) is defined by \((81)\) in bases in an apparently basis-dependent way. But the resulting positive operator \(\rho_A = \tr_B (\rho_{AB})\) of finite trace is basis independent.

The very concept of a partial trace comes from the fact that one can have a state operator (density operator; generalization of state vector) describing a subsystem as follows. For every first-subsystem observable \(O_A \otimes I_B\) one obtains

\[ \langle O_A, \rho_{AB} \rangle = \tr (\rho_{AB} (O_A \otimes I_B)) = \tr_A (\tr_B (\rho_{AB}) O_A) = \tr_A (\rho_A O_A). \]  

(82)

The second partial-trace rule (cf below) was used. (Note that in the full trace \(\tr = \tr_A \tr_B\) the indices are usually omitted as superfluous.)

First Rule: Commutation under the partial trace

If \(\mathcal{H}_A \otimes \mathcal{H}_B\) is a two-subsystem (complex and separable) composite Hilbert space, if, further, \(O_A\) is an operator that acts non-trivially only in \(\mathcal{H}_A\) and \(O_{AB}\) is any operator in the composite Hilbert space, then the following partial-trace rule is valid

\[ \tr_A (O_A O_{AB}) = \tr_A (O_{AB} O_A). \]  

(83)

Naturally, \(O_A\) is actually \(O_A \otimes I_B\) when acting in \(\mathcal{H}_A \otimes \mathcal{H}_B\).

Symmetrically,

\[ \tr_B (O_B O_{AB}) = \tr_B (O_{AB} O_B). \]  

(84)

Rules \((83)\) and \((84)\) are analogous to commutation under a full trace.

Proof. Let \(|r\rangle_A : \forall r\rangle\) and \(|s\rangle_B : \forall s\rangle\) be any complete orthonormal bases in the factor spaces. Then, in view of \(\langle s \mid_B I_B \mid s' \rangle_B = \delta_{s,s'}\), one can write

\[ \langle s \rangle_B \tr_A (O_A O_{AB}) \mid s' \rangle_B = \sum_{r,r',r'',s'} \langle s \rangle_B \langle r \mid_A \langle r' \mid_A \langle s' \mid_B O_{AB} \mid r'' \rangle_A \mid s'' \rangle_B \times \langle r'' \rangle_A \langle s'' \mid_B O_{AB} \mid r' \rangle_A \mid s \rangle_B = \sum_{r,r'} \langle s \rangle_B \langle r \rangle_A \langle r' \rangle_A \langle s \rangle_B  O_{AB} \mid r' \rangle_A \mid s \rangle_B. \]  

(85)

On the other hand,

\[ \langle s \rangle_B \tr_A (O_A O_{AB}) \mid s' \rangle_B = \sum_{r,r',r'',s'} \langle s \rangle_B \langle r \rangle_A \langle r' \rangle_A \langle s' \rangle_B O_{AB} \mid r'' \rangle_A \mid s'' \rangle_B \times \langle r'' \rangle_A \langle s'' \rangle_B O_{AB} \mid r' \rangle_A \mid s \rangle_B = \sum_{r,r'} \langle s \rangle_B O_{AB} \mid r' \rangle_A \langle r' \rangle_A \langle s \rangle_B. \]  

(86)

If one exchanges the order of the two (number) factors and also exchanges the two muto indices \(r'\) and \(r''\) in each term on the RHS of \((86)\), then the RHS of \((85)\) and \((86)\) are seen to be equal. Hence, so are the LHS's. Rule \((84)\) is proved analogously.

Second Rule: Getting out of the partial trace

Under the assumptions of the first rule, the following relations are always valid:

\[ \tr_B (O_A O_{AB}) = O_A \tr_B O_{AB}. \]  

(87)

\[ \tr_B (O_{AB} O_A) = (\tr_B O_{AB}) O_A. \]  

(88)

\[ \tr_A (O_B O_{AB}) = O_B \tr_A O_{AB}. \]  

(89)

\[ \tr_A (O_{AB} O_B) = (\tr_B O_{AB}) O_B. \]  

(90)

An operator that acts non-trivially only in the tensor-factor space that is opposite to the one over which the partial trace is taken behaves analogously as a constant under a full trace: it can be taken outside the partial trace. But one must observe the order (important for operators, not for numbers).

Proof. Let \(|r\rangle_A : \forall r\rangle\) and \(|s\rangle_B : \forall s\rangle\) be any complete orthonormal bases in the factor spaces. Then

\[ \langle r \rangle_A \tr_B (O_A O_{AB}) \mid r' \rangle_A = \sum_{r',r'',s'} \langle s \rangle_B (O_A \otimes I_B) \mid r'' \rangle_A \mid s' \rangle_B \times \langle r'' \rangle_A \langle s' \rangle_B O_{AB} \mid r' \rangle_A \mid s \rangle_B = \sum_{r',r''} \langle r \rangle_A \langle r' \rangle_A \langle r'' \rangle_A \langle s \rangle_B O_{AB} \mid r' \rangle_A \mid s \rangle_B. \]  

(91)

On the other hand,

\[ \langle r \rangle_A O_A \tr_B O_{AB} \mid r' \rangle_A = \sum_{r',r''} \langle s \rangle_B \langle r \rangle_A \langle r' \rangle_A \langle r'' \rangle_A \langle s \rangle_B O_{AB} \mid r' \rangle_A \mid s \rangle_B. \]  

(92)

The RHS of \((91)\) and \((92)\) are seen to be equal. Hence, so are the LHS’s. Relations \((88)\), \((89)\) and \((90)\) are proved analogously.

8.3 Equivalence of the partial scalar product and a certain partial trace

The auxiliary relations that follow stand in certain analogies with the known basic relation

\[ \tr (\Psi_{AB} \langle \Psi \mid_{AB} O_{AB} ) = \langle \Psi \mid_{AB} O_{AB} \mid \Psi_{AB} \rangle. \]  

(93)
Lemma 3.

\[ \text{tr}_B (|\phi_B\rangle\langle\phi_B| |\Psi_{AB}\rangle |\Psi_{AB}\rangle = \frac{\langle\phi_B| |\Psi_{AB}\rangle}{||\phi_B||} \frac{\langle\Psi_{AB}| |\phi_B\rangle}{||\Psi_{AB}||} . \]  

(94)

Proof. Utilizing definition \[51\], taking into account that \(\langle m|_A I_A |\tilde{m}\rangle_A = \delta_{m,\tilde{m}}\), and eventually making use of \[2\], one obtains

\[ \langle m|_A \text{tr}_B (|\phi_B\rangle\langle\phi_B| |\Psi_{AB}\rangle |\Psi_{AB}\rangle |\phi_B\rangle = \sum_n \langle m|_A \langle n| |\phi_B\rangle |\phi_B\rangle \langle\phi_B| |\Psi_{AB}\rangle |\Psi_{AB}\rangle |\phi_B\rangle \langle n| |\phi_B\rangle , \]  

\[ = \sum_n \langle m|_A \langle n| |\phi_B\rangle |\phi_B\rangle \langle\phi_B| |\Psi_{AB}\rangle |\Psi_{AB}\rangle |\phi_B\rangle \langle n| |\phi_B\rangle , \]  

\[ = \sum_n \langle m|_A \langle n| |\phi_B\rangle |\phi_B\rangle \langle\phi_B| |\Psi_{AB}\rangle |\Psi_{AB}\rangle |\phi_B\rangle \langle n| |\phi_B\rangle , \]  

\[ = \langle m|_A \Psi_{AB} \Psi_{AB} \phi_B \phi_B \rangle . \]

\[ \square \]

Lemma 4.

\[ \text{tr} (|\phi_B\rangle\langle\phi_B| |\Psi_{AB}\rangle |\Psi_{AB}\rangle = ||\phi_B||^2. \]  

(95)

Proof. According to Lemma \[3\]

\[ \text{tr} (|\phi_B\rangle\langle\phi_B| |\Psi_{AB}\rangle |\Psi_{AB}\rangle = \text{tr} (|\phi_B\rangle |\Psi_{AB}\rangle |\Psi_{AB}\rangle |\phi_B\rangle \)  

\[ = ||\phi_B||^2 |||\Psi_{AB}||^2 . \]  

Finally, the two Lemmata \[2\] and \[4\] imply the claim.

Proposition 6. The following bridge relation is valid between partial trace and partial scalar product:

\[ \frac{\text{tr}_B (|\phi_B\rangle\langle\phi_B| |\Psi_{AB}\rangle |\Psi_{AB}\rangle)}{\text{tr} (|\phi_B\rangle\langle\phi_B| |\Psi_{AB}\rangle |\Psi_{AB}\rangle)} = \frac{\langle\phi_B| |\Psi_{AB}\rangle}{||\phi_B||} \frac{\langle\Psi_{AB}| |\phi_B\rangle}{||\Psi_{AB}||} . \]

References


