Quantitative entanglement witnesses of isotropic and Werner classes via local measurements

P Silvi\textsuperscript{1}, F Taddei\textsuperscript{2}, R Fazio\textsuperscript{2} and V Giovannetti\textsuperscript{2}

\textsuperscript{1} International School for Advanced Studies, SISSA, Via Bonomea 265, I-34136 Trieste, Italy
\textsuperscript{2} NEST, Scuola Normale Superiore and Istituto Nanoscienze-CNR, Piazza dei Cavalieri 7, I-56126 Pisa, Italy

E-mail: psilvi@sissa.it

Received 18 November 2010, in final form 18 February 2011
Published 10 March 2011
Online at stacks.iop.org/JPhysA/44/145303

Abstract
Quantitative entanglement witnesses allow one to bound the entanglement present in a system by acquiring a single expectation value. In this paper, we analyze a special class of such observables which are associated with (generalized) Werner and isotropic states. Their optimal bounding functions can be easily derived by exploiting known results on twirling transformations. By focusing on an explicit local decomposition for these observables, we then show how simple classical post-processing of the measured data can tighten the entanglement bounds. Quantum optics implementations based on hyper-entanglement generation schemes are analyzed.

PACS numbers: 03.65.Ud, 03.67.Mn, 42.50.Ex

(Some figures in this article are in colour only in the electronic version)

1. Introduction

Detecting and characterizing entanglement in quantum systems is a major issue in quantum information theory, both from the theoretical and experimental point of view [1, 2]. In this context, entanglement witnesses (EWs) turn out to be useful tools that allow one to address this problem with minimal experimental effort, even in those cases where a complete tomographic reconstruction of the state is not experimentally accessible [3]. Indeed, this technique merely requires one to measure a single (non-local) observable whose expectation value yields a certifiable evidence of the presence of entanglement in the system. Unfortunately, it cannot provide a definitive answer for all system configurations (indeed in most cases, the result of the measurement will be inconclusive). Furthermore, even when successful, the result of an EW measure will not be useful in general to quantify the amount of entanglement present in the system. Still, it was recently pointed out [4–7] that the outcome of such measurements can
be used to provide at least non-trivial (lower) bounds for the latter. These observations led to the notion of quantitative EW (QEW) \[8\], which identify those operators whose expectation values can be translated into a quantification of the entanglement present in the measured system.

From a purely technical point of view, it is worth stressing that even though QEWs were originally discussed in the context of standard EWs, they can be thought as independent of the latter. Indeed an EW is characterized by well-established properties, e.g., it must assume positive values on all separable states and must have at least one negative eigenvalue \[3\]. In contrast, the definition of a QEW can be applied to any operator \(Q\) for which it is possible to associate a non-trivial bounding function (see the next section) with a given entanglement measure. Notable examples of such QEWs were already known in the literature even before the formal introduction of this notion, e.g., see \[9\]. In particular, the idea of bounding the entanglement of a state via the acquisition of a single expectation value was addressed in a series of papers \[10\] that applied previous ideas on twirling transformations \[11\] and Werner states \[12, 13\]; the same argument has been extended lately to isotropic states \[9, 13–19\]. The geometrical properties of Werner and isotropic states can be used to build explicitly a QEW capable of detecting up to a high degree of bipartite entanglement. In this paper, we review the derivation of such QEW \[5–8\] adopting a formal yet simple scheme. Subsequently, using an explicit decomposition of such operators in terms of correlated local observables \[20\], we perform a data-processing optimization which will enable further tightening of the entanglement bounds. Expressing an EW or a QEW in terms of local operation and classical communication (LOCC) measurement schemes is yet another important aspect of entanglement detection \[1, 9, 19–26\]. It deals with the problem of acquiring expectations values of a non-local observable (i.e. an observable that does not admit a separable basis of eigenvectors), when only local resources are accessible experimentally. Even though the decomposition we assume is not the most efficient one in terms of setup preparations \[1\], it has the main advantage of being easily implementable in some specific experimental configurations which exploit multi-rail quantum encoding strategies \[27–33\]. We will explicitly focus on some optical scheme based on hyper-entanglement setups \[27–29\] where bipartite entangled qudits states are generated by exploiting spatial and polarization degrees of freedom.

The paper is organized as follows. After reviewing some basic definitions around a QEW, in section 2 we focus on QEWs of the Werner and isotropic classes for which we present the optimal bounding functions. In section 3, we introduce the local decompositions for such QEWs which use 2\(d\)-independent setup preparations (\(d\) being the local dimension of a system site). We then show how the same measurement outcomes can be post-processed to strengthen the entanglement bounds. Finally, in section 4 we discuss how to implement the proposed measurement scheme in the context of optical multi-rail encodings.

2. Quantitative entanglement witnesses

Let \(Q\) be an observable acting on a bipartite system \(AB\), and let \(E_M\) be a given convex and continuous measure of bipartite entanglement, e.g. the entanglement of formation \[1, 2, 11\]. We say that \(Q\) is a QEW for the measure \(E_M\) if we can define a (non-trivial) non-negative function \(b\) for which it holds:

\[
E_M(\rho) \geq b(\text{Tr}[Q \rho]),
\]

for all \(\rho\) density matrices of \(AB\). Analogous to a standard EW \[1–3\], a QEW provides an operational method for detecting entanglement. Furthermore, it also yields a quantitative estimation of it. As pointed out in \[5, 6, 8\], a fundamental problem is to find the optimal
bounding function of a given QEW $Q$, i.e. the highest positive $b$ for which (1) holds. The latter can be formally defined as

$$b_{\text{opt}}(x) = \inf_{\rho} \{ \mathcal{E}_M(\rho) | \text{Tr}[Q\rho] = x \},$$

(2)

where the minimization is performed over the whole set of density matrices. This is a convex (generally not monotonic) function which typically is rather difficult to obtain. A partial characterization of $b_{\text{opt}}$ was proposed in [6], where it was pointed out that, due to the convexity of entanglement measures, we can write

$$b_{\text{opt}}(x) = \sup_{\lambda} \{ \lambda x - \mathcal{L}_M(\lambda Q) \},$$

(3)

where the supremum is taken over all real values of the parameter $\lambda$ and $\mathcal{L}_M$ is the Legendre transform of the measure functional $\mathcal{E}_M$, with respect to the trace scalar product, i.e.

$$\mathcal{L}_M(A) \equiv \sup_{\rho} \{ \text{Tr}[A\rho] - \mathcal{E}_M(\rho) \}. 
$$

(4)

(Here, the supremum is taken over the whole set of density matrices.)

A useful property of the set of QEWs is that it is invariant under LOCC. In particular, any (non-trivial) unitary LOCC transformation $\Gamma$ takes a given QEW to a new QEW while preserving its bounding functions including the optimal one, which is trivially again optimal. This is a simple consequence of the cyclicity of the trace which gives

$$\mathcal{E}_M(\rho) = \mathcal{E}_M(\Gamma(\rho)) \geq b(\text{Tr}[\Gamma^*(Q)\rho]),$$

(5)

for all $b$ of $Q$. (In this expression, $\Gamma^*$ stands for the adjoint super-operator of $\Gamma$.) Note that exploiting the properties of the transformations $\Gamma$, we can derive a stronger bound on $\mathcal{E}_M(\rho)$, namely [4]

$$\mathcal{E}_M(\rho) \geq \max_\Gamma b_{\text{opt}}(\text{Tr}[\Gamma^*(Q)\rho]) = b_{\text{opt}}(q_*),$$

(6)

where the maximization is now performed over the set of all LOCC unitaries, and where $q_*$ indicates the value of $\text{Tr}[\Gamma^*(Q)\rho]$ that yields the highest value of $b_{\text{opt}}$. Unfortunately, $q_*$ is a rather complicated (nonlinear) function of the state $\rho$ which is arguably hard to compute. Still, we will see in the following that, for the cases we consider here, an operational characterization of it can be obtained by focusing on a subclass of $\Gamma$.

### 2.1. Werner and isotropic QEWs

Finding the optimal bounding functions for QEWs is a hard problem [5, 6, 8]. However, there are certain highly symmetric choices of $Q$ for which $b_{\text{opt}}$ can be obtained without going through the optimizations of (3). Prototypical examples are given by the operators

$$F = \sum_{\alpha,\beta=1}^{d} |\alpha\beta\rangle \langle \beta\alpha|, \quad G = \sum_{\alpha,\beta=1}^{d} |\alpha\alpha\rangle \langle \beta\beta|,$$

(7)

where $\{|\alpha\rangle; \alpha = 1, \ldots, d\}$ is the canonical basis for the subsystem $A$ and $B$ (both assumed to be $d$ dimensional). The first is nothing but the swap operator on the composite system $AB$, while the second is a projector into the (not normalized) maximally entangled state with constant coefficients. They are related via partial transpose (i.e. $F = G^T$) and are QEWs

---

3 The QEW set is clearly invariant under the whole set of LOCC transformations. However, if the mapping $\Gamma$ is not invertible (i.e. not unitary), then the optimal bounding function of the new observable $\Gamma^*(G)$ need not coincide with the original one.
which we denote the Werner class and isotropic class, respectively, for reasons which will be clear soon. Their optimal bounding functions can be computed for several entanglement measures including entanglement of formation ($E_{\text{OF}}$), relative entropy of entanglement [13], and convex-roof extended negativity [18]. For the sake of simplicity, however in the following we focus only on the former. In this case, the optimal bounding functions for $F$ and $G$ are respectively given by

$$b_{\text{Wer}}(x) = \begin{cases} h_2 \left( \frac{1+x-\sqrt{x^2+1}}{2} \right) & \text{for } x \leq 0, \\ 0 & \text{for } x > 0, \end{cases}$$

$$b_{\text{Iso}}(x) = \begin{cases} \text{co}[h_2(\gamma) + (1-\gamma) \log(d-1)] & \text{for } x > 1, \\ 0 & \text{for } x \leq 1, \end{cases}$$

where $h_2(y) = -y \log_2 y - (1-y) \log_2 (1-y)$ is the binary Shannon entropy, $\gamma(x) = \frac{1}{d} (\sqrt{x} + \sqrt{(d-1)(d-x)})$, and ‘co’ means taking the convex-hull of the inner expression, i.e. the largest convex curve nowhere larger than the given one—note however that in (8) the argument is already convex when $d = 2$.

Proving that the operator $F$ is a QEW with the optimal bounding function (8) involves the knowledge of the properties of Werner states [12, 13] and notion of twirling transformations [11]. Werner states $\rho_W$ in an $H_d \otimes H_d$ space are those which are unvariant under the whole $U \otimes U$ group of transformations:

$$(U \otimes U) \rho_W (U \otimes U)^\dagger = \rho_W,$$

for all unitaries $U$ on $H_d$. They form a one-parameter manifold of states which can be expressed as

$$\rho_W(f) = \frac{1}{d(d^2-1)} \mathbb{1} (d-f) \mathbb{1} + (fd-1)\mathbb{F},$$

where $\mathbb{1}$ is the $d^2 \times d^2$ identity operator and where $\mathbb{F}$ is the swap operator (7). The parameter $f$ takes values between $-1$ and $1$ to ensure positivity, and it coincides with the expectation value of $\mathbb{F}$ over $\rho_W(f)$, i.e.

$$f = \text{Tr}[\mathbb{F} \rho_W(f)].$$

The entanglement of formation $E_{\text{OF}}(\rho_W(f))$ for the state $\rho_W(f)$ is a known quantity [13] which can be expressed in terms of function (8), i.e.

$$E_{\text{OF}}(\rho_W(f)) = b_{\text{Wer}}(f) = b_{\text{Wer}}(\text{Tr}[\mathbb{F} \rho_W]).$$

To demonstrate that $b_{\text{Wer}}$ is a valid bounding function for $F$ let us define the following completely-positive trace-preserving mapping, acting on a generic density matrix $\rho$ of $AB$:

$$P_\Lambda(\rho) = \int_\Lambda d\mu(U) \rho (U \otimes U)^\dagger,$$

where $\Lambda$ is the unitary group of $H_d$, and $d\mu$, its Haar measure, is uniquely defined by compactness of $\Lambda$. The transformation $P_\Lambda$ is often labeled as the twirling operation [11]. Three properties of $P_\Lambda$ are important to mention for our purposes [11, 13, 18]. First, $P_\Lambda$ is a projector, i.e. $P_\Lambda^2 = P_\Lambda$, and its range is exactly the 1D manifold of Werner states. Moreover,

4 This is a rather special case where one can easily show that the two QEWs are basically LOCC equivalent in the sense discussed in (5). Indeed in this case, $G$ can be expressed in terms of $F$ via a proper local unitary transformation plus a constant term which can be compensated via classical data post-processing.
\[ \mathcal{P}_\Lambda \] preserves the expectation value of \( \mathcal{F} \), i.e. \( \text{Tr} [\mathcal{F} \rho] = \text{Tr} [\mathcal{F} \mathcal{P}_\Lambda (\rho)] \), which, along with (11), tells us that

\[ \mathcal{P}_\Lambda (\rho) = \rho_W (\text{Tr} [\mathcal{F} \rho]) . \] (14)

Finally, since its Kraus decomposition (13) is made up of tensor product operators, \( \mathcal{P}_\Lambda \) is a (non-invertible) LOCC transformation. Therefore, it can only decrease the entanglement, thus providing the following inequality:

\[ \mathcal{E}_{\text{OF}} (\mathcal{P}_\Lambda (\rho)) \leq \mathcal{E}_{\text{OF}} (\rho) . \] (15)

Putting together (12), (14) and (15), we finally obtain [10, 11, 13, 18]

\[ \mathcal{E}_{\text{OF}} (\rho) \geq b_{\text{Wer}} (\text{Tr} [\mathcal{F} \rho]), \] (16)

for any given state \( \rho \), which is the proper definition of the bounding function for \( b_{\text{Wer}} \). To check optimality, it is sufficient to see that for any \( x \in [-1, 1] \) there is a density matrix \( \rho \) with \( \text{Tr} [\mathcal{F} \rho] = x \) for which equality holds: this happens for the Werner state \( \rho_W (x) \). Then, since for any function \( b \) higher in \( x \) than \( b_{\text{Wer}} \), condition (1) would be broken, \( b_{\text{Wer}} \) must be optimal.

The proof of the optimality of function (8) for the isotropic QEW \( \mathcal{G} \) is similar to the one we exploited in the Werner case. This time we consider isotropic states [13–17], i.e. the set of states \( \rho_I \) invariant under every transformation of the form

\[ (U \otimes U^*) \rho_I (U \otimes U^*)^\dagger = \rho_I, \] (17)

where \( \ast \) denotes the complex conjugation of the matrix entries, say, with respect to the canonical basis. Isotropic states form again a one-parameter family, defined by

\[ \rho_I (g) = \frac{1}{d(d^2 - 1)} [(d - g) \mathbb{1} + (gd - 1) \mathcal{G}], \] (18)

with \( 0 \leq g \leq d \) and \( \mathcal{G} \) as in (7). The parameter \( g \) is also the expectation value \( g = \text{Tr} [\mathcal{G} \rho_I (g)] \). On top of that, their entanglement of formation is known [15] to be given by \( \mathcal{E}_{\text{OF}} (\rho_I (g)) = b_{\text{Iso}} (g) \). Again, we exploit the properties of a twirling operation, this time of the form

\[ \mathcal{P}'_\Lambda (\rho) = \int \mathcal{P}_\Lambda (\rho) \mathcal{G} d\mu(U), \] (19)

which projects [13, 18] any given state \( \rho \) in the isotropic state \( \rho_I (\text{Tr} [\mathcal{G} \rho]) \) while preserving the expectation value of \( \mathcal{G} \) and degrading the entanglement of formation. Following these considerations, we obtain [9, 19]

\[ \mathcal{E}_{\text{OF}} (\rho) \geq b_{\text{Iso}} (\text{Tr} [\mathcal{G} \rho]), \] (20)

so that \( b_{\text{Iso}} \) is a bounding function for \( \mathcal{G} \), and it is optimal since any function pointwise higher would break the bounding condition (1) upon one isotropic state at least. (It is worth pointing out that inequality (20) was originally derived in [9] without invoking the properties of twirling.)

3. Local decomposition of a QEW

A practical problem in entanglement detection is how to acquire expectation values of non-local observables (for instance our QEWs), while we are typically allowed to perform only LOCC measurements [1, 20–26]. The standard solution is to expand the witness operator \( \mathcal{Q} \) to a sum of local observables of the form

\[ \mathcal{Q} = \sum_i \epsilon_i A_i \otimes B_i, \] (21)
with \( c_i \) being some complex coefficients. The expectation value of \( Q \) can then be reconstructed by simple data post-processing from the expectation values of the tensor product observables \( A_i \otimes B_j \), which can be acquired by performing classical correlated measures on \( A \) and \( B \). The decomposition (21) is not unique and, unless \( Q \) admits a set of separable eigenvectors\(^5\), it will involve a certain number of non-commuting operators \( A_i \otimes B_j \). In this case, several independent experimental setups need to be prepared, one for each non-commuting set of observables \( A_i \otimes B_j \) entering into (21). It is thus reasonable to try to optimize the local decomposition in a way that it corresponds to the smallest possible experimental effort. This problem attracted a lot of attention [20, 21, 24–26] but it appears that no general solution has been found yet. Instead, several specific solutions for particular classes of witnesses have been proposed.

In this section, we focus on a local decomposition of the QEWs of the Werner and the isotropic class which was first discussed in [20]. Such a solution is not the most efficient one in terms of the number of setup re-preparations. Indeed it requires \( \sim 2d \)-independent setups, while a theoretical lower bound of \( d \) was presented in [26] which, at least for \( d \), is prime and known to be achievable [25]. Yet the proposed decomposition presents at least two advantages. First, as we shall see in section 3.1, by simply reprocessing the data accumulated during the various experiment runs, one can obtain expectation values of an (exponentially large) class of QEW operators which are LOCC equivalent, in the sense of (5), to the one we started with. This leads us to the experimental acquisition of several entanglement bounds which are optimal in the sense of (2): by taking their highest value we can hence provide an estimation of the quantity \( q_* \) of (6). Secondly, as we shall see in section 4, the proposed decomposition scheme appears to be naturally suited for multi-rail encoding implementations.

To start, we introduce the following set of (local) single-site observables:

\[
\begin{align*}
P_\alpha &= |\alpha\rangle\langle\alpha|, \\
X_{\alpha\beta} &= |\alpha\rangle\langle\beta| + |\beta\rangle\langle\alpha|, \\
Y_{\alpha\beta} &= i|\alpha\rangle\langle\beta| - i|\beta\rangle\langle\alpha|, \\
\end{align*}
\]

(22)
defined for any \( \alpha, \beta \in \{0, \ldots, d\}, \alpha \neq \beta \) (as before the vectors \( |\alpha\rangle \) stands for the canonical basis of the site). Since they form a basis for the space of operators acting on \( \mathcal{H}_d \), we can express \( F \) and \( G \) as follows:

\[
\begin{align*}
F &= \sum_\alpha P_\alpha \otimes P_\alpha + \frac{1}{4} \sum_{\beta \neq \alpha} (X_{\alpha\beta} \otimes X_{\alpha\beta} + Y_{\alpha\beta} \otimes Y_{\alpha\beta}), \\
G &= \sum_\alpha P_\alpha \otimes P_\alpha + \frac{1}{4} \sum_{\beta \neq \alpha} (X_{\alpha\beta} \otimes X_{\alpha\beta} - Y_{\alpha\beta} \otimes Y_{\alpha\beta}), \\
\end{align*}
\]

(23)
which is of the form (21) with \( d^2 \) terms and the coefficients \( c_i \) being real. Therefore, the expectations values of \( F \) and \( G \) can both be acquired with at most order \( d^2 \)-independent experimental setups. Of course, some of the operators appearing in sums (23) commute with each other, so in principle it is possible to evaluate them within the same experiment. For instance any pair of operators of the set \( \Sigma = \{ P_\alpha, X_{\alpha\beta}, Y_{\alpha\beta} \} \) commute iff all the indices are different, and clearly this commutation rule extends to all the observable of the form \( A \otimes A \) with \( A \in \Sigma \). Exploiting this simple property, one can indeed reduce [20] the number of setup re-preparations needed to obtain all the values \( \langle P_\alpha \otimes P_\alpha \rangle, \langle X_{\alpha\beta} \otimes X_{\alpha\beta} \rangle \) and \( \langle X_{\alpha\beta} \otimes X_{\alpha\beta} \rangle \) from \( d^2 \) to order \( d \). (Here, we use \( \langle \cdot \cdot \cdot \rangle \) to represent the expectation value with respect to the state \( \rho \).) For the sake of clarity, we review this result by re-deriving using a different method than that presented in [20]. To do so we find it convenient to turn the problem into a graph coloring

\(^5\) This case however is not relevant as an operator of that form would not be able to detect entanglement.
Figure 1. Sketch of the minimization strategy for $d = 7$, first two steps. A red line connecting vertices $\alpha$ and $\beta$ means measuring $X_{\alpha\beta} \otimes X_{\alpha\beta}$; the green dot is the measurement of $P_{\alpha} \otimes P_{\alpha}$.

Figure 2. Sketch of the minimization strategy for $d = 6$. Turns 1, 2 and 11 are drawn. The sixth vertex has been placed in the geometrical center of the pentagon.

game. Let us assume that we have a graph made of $d$ vertices and all the possible edges among them (also known as a complete graph), e.g. see figures 1 and 2; we are also given one green brush to color the vertices with, alongside with one blue and one red brush to paint the edges of the graph. The idea is that when we paint the $\alpha$th vertex green we are actually acquiring $\langle P_{\alpha} \otimes P_{\alpha} \rangle$, while if we paint in red (resp. blue) the edge connecting the nodes $\alpha$ and $\beta$, we evaluate $\langle X_{\alpha\beta} \otimes X_{\alpha\beta} \rangle$ (resp. $\langle Y_{\alpha\beta} \otimes Y_{\alpha\beta} \rangle$). As a rule, during a single turn the player can paint everything he wants, but not

- an edge and a vertex at one of its ends,
- two consecutive edges,
- the same edge twice.

Clearly, these game rules encode the commutation relations we discussed. The aim of this game is to paint all the vertices green and all the segments in both red and blue spending the minimal number of turns. The optimal solutions for the game are presented next.

Colouring for $d$ odd. For simplicity, we draw the graph like a regular polygon with $d$ edges, including all the diagonals, see figure 1. At the first turn, the player paints in red one edge and all the diagonals that are parallel to that edge, as well as the farthest vertex, obviously green. From turn 2 to $d$, he chooses every time another edge of the polygon and repeats the process,
until all the edges have been chosen. From turn \(d + 1\) to \(2d\), he does the same thing he did in the first \(d\) turns but using the blue brush instead of the red one. That completes the game in \(2d\) turns. This coloring strategy is optimal because, due to the rules, every turn a player cannot paint more than \((d - 1)/2\) segments, and we have \(d(d - 1)\) segments to paint, so that \(2d\) is actually a lower bound for the game turns.

**Colouring for \(d\) even.** This time we draw the graph as a \((d - 1)\)-sided regular polygon, with the additional vertex in the geometric center, and consider all the edges, diagonals, and vertex–center segments, see figure 2. At the first turn the player paints in red one edge, the parallel diagonals, and the segments joining the center with the farthest vertex. Afterward, he rotates the pattern; and subsequently he does the same with the blue brush. When he is finished, he uses a turn to paint all the vertices. That completes the game in \(2d - 1\) turns. As before, \(2d - 2\) is a lower bound for painting the segments, but if that is the case, we need at least one additional turn to paint the dots. Therefore, this coloring strategy is optimal.

### 3.1. Multiple bounds from data processing

In the previous section, we presented a strategy for evaluating \(\langle F \rangle\) and \(\langle G \rangle\) according to the local decompositions (23). Now we want to show that it is possible to reprocess the same (classical) data to obtain expectation values of other QEWs. These are generated from \(F\) and \(G\) by means of local invertible transformations as discussed in section 2. Specifically, we focus on the following phase-shift operators:

\[
W(\xi) = \sum_j \xi(\alpha) |\alpha\rangle\langle \alpha|, \tag{24}
\]

where \(\xi : \{1, \ldots, d\} \rightarrow \{-1, 1\}\) is a generic two-valued function. Now, let us define

\[
\mathbb{F}(\xi) = (\mathbb{1} \otimes W(\xi))\mathbb{F}(\mathbb{1} \otimes W^\dagger(\xi)),
\]

\[
\mathbb{G}(\xi) = (\mathbb{1} \otimes W(\xi))\mathbb{G}(\mathbb{1} \otimes W^\dagger(\xi)). \tag{25}
\]

For all choices of \(\xi\), these observables are QEWs of the Werner and the isotropic class, respectively. (They are just associated with Werner and isotropic states defined on a different basis.) Furthermore, since they are LOCC equivalent to the original \(F\) and \(G\), their optimal bounding functions are still given by (8). By optimizing with respect to all \(W(\xi)\), we can write

\[
\mathcal{E}_{\text{OF}}(\rho) \geq b_{\text{W}}(f_\star(\rho)), \quad \mathcal{E}_{\text{OF}}(\rho) \geq b_{\text{Iso}}(g_\star(\rho)), \tag{26}
\]

where the maximum is taken over all possible choices of \(\xi\). These expressions can be further simplified by exploiting the monotonicity properties of \(b_{\text{W}}(x)\) and \(b_{\text{Iso}}(x)\) (the former is non-increasing while the latter is non-decreasing) to move the optimization over \(\xi\) on their arguments. This yields

\[
\mathcal{E}_{\text{OF}}(\rho) \geq b_{\text{W}}(f_\star), \quad \mathcal{E}_{\text{OF}}(\rho) \geq b_{\text{Iso}}(g_\star), \tag{27}
\]

where we defined the quantities

\[
f_\star(\rho) = \min_{\xi} \{\text{Tr}[\mathbb{F}(\xi) \rho]\}, \quad g_\star(\rho) = \max_{\xi} \{\text{Tr}[\mathbb{G}(\xi) \rho]\}. \tag{28}
\]

These are complicated (nonlinear) functions of the state \(\rho\) which are related to the quantity \(q_\star\) of (6). The main difference is that the latter was optimized with respect to the whole set
of unitary LOCC transformations \( \Gamma \), while \( f_\ast \) and \( g_\ast \) are optimized only with respect to the a proper subset of such set, namely the phase flips (25). Luckily enough, we can simplify \( f_\ast \) and \( g_\ast \) by rewriting the operators \( \mathcal{F}(\xi) \) and \( \mathcal{G}(\xi) \) in terms of operators (22). Indeed by doing so, we find that they are made of the very same \( d^2 \) operators appearing in (23), namely

\[
\mathcal{F}(\xi) = \sum_\alpha P_\alpha \otimes P_\alpha + \frac{1}{4} \sum_{\alpha \neq \beta} \xi(\alpha) \xi(\beta) (X_{\alpha\beta} \otimes X_{\alpha\beta} + Y_{\alpha\beta} \otimes Y_{\alpha\beta}),
\]

and

\[
\mathcal{G}(\xi) = \sum_\alpha P_\alpha \otimes P_\alpha + \frac{1}{4} \sum_{\alpha \neq \beta} \xi(\alpha) \xi(\beta) (X_{\alpha\beta} \otimes X_{\alpha\beta} - Y_{\alpha\beta} \otimes Y_{\alpha\beta}).
\]

This implies in particular that their expectation values can be recovered by the same set of LOCC measurements required for \( \mathcal{F}(\xi) \) and \( \mathcal{G}(\xi) \), allowing us to build \( 2^{d-1} \)-independent new Werner QEWs of the form \( \mathcal{F}(\xi) \), as well as \( 2^{d-1} \)-independent new isotropic QEWs of the form \( \mathcal{G}(\xi) \) by simple data post-processing. (The number \( 2^{d-1} \) follows from the fact that there are \( 2^d \) possible \( \xi \) functions, but \( \mathcal{F}(\xi) = \mathcal{F}(-\xi) \) and \( \mathcal{G}(\xi) = \mathcal{G}(-\xi) \).) Clearly, this is a very simple and cheap way to substantially increase the information we can gather on \( \rho \). As a matter of fact, expressions (29) and (30) are also useful to simplify the optimizations of (28).

In particular, define the \( d^2 \times d^2 \) matrix \( M^{(\ast)} \) of elements

\[
M^{(\ast)}_{\alpha,\beta} = (P_\alpha \otimes P_\beta),
\]

\[
M^{(-)}_{\alpha,\beta} = \frac{1}{2} (X_{\alpha\beta} \otimes X_{\alpha\beta} \pm (Y_{\alpha\beta} \otimes Y_{\alpha\beta})), \quad \text{for } \alpha \neq \beta
\]

where, as usual, \( (\cdots) \) stands for taking the expectation value with respect to \( \rho \). These matrices are composed of the measurement outcomes one acquires when recovering the expectation values of \( \mathcal{F} \) and \( \mathcal{G} \) via the local decomposition introduced in the previous section. In particular, we can write

\[
\text{Tr} [\mathcal{F}(\xi) \rho] = \sum_{\alpha,\beta} \xi(\alpha) M^{(\ast)}_{\alpha,\beta} \xi(\beta) \equiv \vec{\xi}^T \cdot M^{(\ast)} \cdot \vec{\xi},
\]

\[
\text{Tr} [\mathcal{G}(\xi) \rho] = \sum_{\alpha,\beta} \xi(\alpha) M^{(-)}_{\alpha,\beta} \xi(\beta) \equiv \vec{\xi}^T \cdot M^{(-)} \cdot \vec{\xi},
\]

where \( \vec{\xi} \equiv (\xi(1), \ldots, \xi(d))^T \) is the vector formed by the output entries of the binary function \( \xi \). Therefore, in order to determine values, (28) is sufficient to respectively minimize or maximize these quantities over \( \vec{\xi} \), i.e.

\[
f_\ast(\rho) \equiv \min \vec{\xi} \vec{\xi}^T \cdot M^{(\ast)} \cdot \vec{\xi},
\]

\[
g_\ast(\rho) \equiv \max \vec{\xi} \vec{\xi}^T \cdot M^{(-)} \cdot \vec{\xi}.
\]

```
\[
\begin{align*}
\frac{1}{4} \sum_{\alpha \neq \beta} |\langle X_{\alpha\beta} \otimes X_{\alpha\beta} \rangle + \langle Y_{\alpha\beta} \otimes Y_{\alpha\beta} \rangle|,
\frac{1}{4} \sum_{\alpha \neq \beta} |\langle X_{\alpha\beta} \otimes X_{\alpha\beta} \rangle - \langle Y_{\alpha\beta} \otimes Y_{\alpha\beta} \rangle|,
\end{align*}
\]

```

Unfortunately, apart from the trivial case of \( d = 2 \), providing a general analytical solution for these expressions is non-trivial and dependent on \( M^{(\ast)} \) (indeed it is formally equivalent to find a ground state of a frustrated, classical many-body spin system). On the other hand, one can easily evaluate upper bounds as follows:

\[
\begin{align*}
f_\ast(\rho) & \geq \sum_\alpha (P_\alpha \otimes P_\alpha) - \frac{1}{4} \sum_{\alpha \neq \beta} |\langle X_{\alpha\beta} \otimes X_{\alpha\beta} \rangle + \langle Y_{\alpha\beta} \otimes Y_{\alpha\beta} \rangle|,
\end{align*}
\]

\[
\begin{align*}
g_\ast(\rho) & \leq \sum_\alpha (P_\alpha \otimes P_\alpha) + \frac{1}{4} \sum_{\alpha \neq \beta} |\langle X_{\alpha\beta} \otimes X_{\alpha\beta} \rangle - \langle Y_{\alpha\beta} \otimes Y_{\alpha\beta} \rangle|,
\end{align*}
\]

even though for \( d > 3 \) they will be in general not tight and will not be useful to lower bound the entanglement of formation of the system. Luckily enough however, the optimizations (33) are treatable numerically, especially considering that for the experimental implementation we are interested in, \( d \) is usually of the order of magnitude of units (see the next section).
Figure 3. Example of data sampling for the $AB$ system in which the local degree of freedom of each site is composed of three independent optical channels (here represented as spatially separated for simplicity). The figure represents the first step of the detection strategy: applying the 50% transmissive beam splitter and measuring $\langle (N_{A1} - N_{A2})(N_{B1} - N_{B2}) \rangle$ gives us the expectation value of $X_{1,2} \otimes X_{1,2}$ (if we put a $\Phi = \pi/2$ phase shift on channel 2, we would have $Y_{1,2} \otimes Y_{1,2}$ instead); at the same time we also acquire $\langle P_3 \otimes P_3 \rangle = \langle N_{A3} N_{B3} \rangle$. For later steps, we need to permute the channels in the picture and perform those measurements as well.

4. Implementation: multi-rail encodings in quantum optics

QEW techniques for estimating lower bounds of the entanglement of a quantum state are suitable, in principle, to any experimental setting or physical apparatus. Yet, we are inclined to think that the proposed local decomposition scheme of Werner and isotropic witnesses we presented in the previous section would be particularly well-suited to multi-rail encoding experimental settings. In such implementations, a state $|\alpha\rangle$ of the canonical basis is represented by the presence of a particle/carryer in a specific transmission channel (e.g. spatial mode, polarization, time-bin or orbital angular momentum) indexed by $\alpha$. Multi-rail encodings are common in both optics [27–31] and electronics (e.g., see [32, 33]). In the following however we will focus only on the former case as it is intrinsically more flexible. For these architectures, the QEW detection strategy discussed in section 3 entails to properly mixing a pair of incoming modes via simple beam-splitter transformations and perform coincidence counting measurements at the exit port of the device, see figure 3. Accordingly the proposed LOCC measurements appear to be more suited to detect entanglement in these systems than (say) the optimal schemes of [25] which requires coherent mixing among multiple channels.

Consider for instance two separated sets $A$ and $B$ of (monochromatic) spatial optical modes. Within each of the two subsystems, a single photon is injected and allowed to flow through $d$ possible optical paths. In one of the experimental setups of [28] for instance, $d = 8$ was realized by exploiting polarization and two different spatial degrees of freedom. A generic (pure) state of the system can then be written as

$$|\Psi_m\rangle = \sum_{\alpha, \beta=1}^{d} \Phi_{\alpha\beta} \hat{a}_{A,\alpha}^\dagger \hat{a}_{B,\beta}^\dagger |0\rangle = \sum_{\alpha, \beta=1}^{d} \Phi_{\alpha\beta} |\alpha\beta\rangle,$$  \hspace{1cm} (35)
where the $\hat{a}_{j,\alpha}$ is the annihilation operator of a photon flowing through channel $\alpha$ at port/subsystem $j \in \{A, B\}$, and where $|\alpha\beta\rangle$ denotes the two-photon state $\hat{a}_{\alpha}^\dagger \hat{a}_{\beta}^\dagger |0\rangle$.

We will describe local transformations on a photonic state as a forward elastic scattering process:

$$\hat{b}_{A,\alpha} = \sum_{\gamma=1}^{d} U_{\alpha,\gamma} \hat{a}_{A,\gamma}, \quad \hat{b}_{B,\beta} = \sum_{\gamma=1}^{d} V_{\beta,\gamma} \hat{a}_{B,\gamma},$$

with $U$ and $V$ unitaries which acts on the $A$ and $B$ sets, respectively. After applying such transformation all we have to do is record the channel-selective photo-coincidence measurements

$$\langle \hat{N}_{A,\alpha} \hat{N}_{B,\beta} \rangle = \langle \Psi_{in} | \hat{b}_{A,\alpha}^\dagger \hat{b}_{A,\alpha} \hat{b}_{B,\beta}^\dagger \hat{b}_{B,\beta} | \Psi_{in} \rangle = |\langle \alpha\beta | U \otimes V | \Psi_{in} \rangle|^2.$$  

(37)

This way, we achieve simultaneously the expectation values of the $d^2$ one-dimensional projectors of the form $U^\dagger \otimes V^\dagger |\alpha\beta\rangle \langle \alpha\beta| U \otimes V$. The last step is to identify a proper set of unitaries $U$ and $V$ that lead to the determination of the quantities $\langle P_{\alpha} \otimes P_{\alpha} \rangle$, $\langle X_{\alpha\beta} \otimes X_{\alpha\beta} \rangle$, and $\langle Y_{\alpha\beta} \otimes Y_{\alpha\beta} \rangle$ which are needed for characterizing the selected QEW. For instance, assume we want to measure the expectation value of $X_{\alpha\beta} \otimes X_{\alpha\beta}$, where $\alpha$ and $\beta$ represent two distinct channels. Then, $U$ and $V$ could be implemented as Hadamard gates, acting in both the subsystems $A$ and $B$, which couples the pair of such channels. This entails the implementation of a beam-splitter transformation of transmissivity 50% which coherently mixes the two channels. After that, we can convert coincidence measurements into data via

$$\langle X_{\alpha\beta} \otimes X_{\alpha\beta} \rangle = \langle \hat{N}_{A,\alpha} - \hat{N}_{A,\alpha} \rangle \langle \hat{N}_{B,\beta} - \hat{N}_{B,\beta} \rangle.$$  

(38)

Similarly, if our goal were to obtain $\langle Y_{\alpha\beta} \otimes Y_{\alpha\beta} \rangle$ instead, the unitarities $U$ and $V$ could be chosen as Hadamard transformations, this time applied after a local phase shift (e.g. of $\pi/2$ acting upon channel $\beta$). See figure 3 for details. We can now translate into experimental operations every possible strategy for the local QEW decomposition of section 3.

As a concluding remark, we point out that the same analysis can be generalized to an electronic implementation of multi-rail encodings [32, 33]. In this case, ports $A$ and $B$ will correspond to independent leads, whose transverse modes play the role of the optical paths. Under appropriate conditions [32], the coincidence counting (37) can be related to the zero-frequency shot noise, which is a measurable quantity. In [34], for example, a procedure for constructing EWs was employed for different solid-state systems for the case of two transverse modes per lead. In such systems, however, an experimental limitation is represented by the difficulty of working with more than two modes.

5. Conclusions

We have reviewed the notion of QEWs and clarified that there exist a class of such operators for which the optimal bounding function can be easily determined without necessarily going through the complex optimization procedure of (2) and (3). For these QEWs, we then considered an explicit local decomposition that allows one to acquire the needed expectation values via a limited number of LOCC measurements. The proposed scheme is not optimal in terms of the number of independent setups one need to prepare, but has two main advantages with respect to other schemes [25]. In particular, we can easily optimize with respect to a large class of unitary LOCC providing an operational characterization of the associated nonlinear quantities, see (32). Furthermore, we have shown that it is suitable for experimental implementations based on multi-rail encodings by means of simple two-mode transformations (beam-splitter mixings) followed by correlated photo-counting detections.
Acknowledgments

We thank P Hyllus, C Macchiavello, P Mataloni, and A Osterloh, for useful comments and discussions. We are particularly grateful to P Hyllus for pointing out [26]. This work was in part supported by the FIRB IDEAS project RBID08B3FM.

References

[1] Gühne O and Tóth G 2009 Phys. Rep. 474 1
[2] Bruss D 2002 J. Math. Phys. 43 4237
[3] Horodecki M, Horodecki P and Horodecki R 1996 Phys. Lett. A 223 1
Terhal B M 2000 Phys. Lett. A 271 319
Bruss D, Cirac J I, Horodecki P, Hulpke F, Kraus B, Lewenstein M and Sanpera A 2002 J. Mod. Opt. 49 1399
Barbieri M, De Martini F, Di Nepi G, Mataloni P, D’Ariano G M and Macchiavello C 2003 Phys. Rev. Lett. 91 227901
[4] Audenaert K M R and Plenio M B 2006 New J. Phys. 8 266
[5] Gühne O, Reimpell M and Werner R F 2007 Phys. Rev. Lett. 98 110502
[6] Gühne O, Reimpell M and Werner R F 2008 Phys. Rev. A 77 052317
[7] Osterloh A and Hyllus P 2009 arXiv:0908.3467[quant-ph]
[8] Eisert J, Brandão F G S and Audenaert K M R 2007 New J. Phys. 9 46
[9] Barreiro J T, Langford N K, Peters N A and Kwiat P G 2005 Phys. Rev. Lett. 95 260501
[10] Barbieri M, Cinelli C, Mataloni P and De Martini F 2005 Phys. Rev. A 72 052110
Vallone G, Ceccarelli R, De Martini F and Mataloni P 2008 Phys. Rev. A 78 062303
Vallone G, Ceccarelli R, De Martini F and Mataloni P 2009 Phys. Rev. A 79 030301
Ceccarelli R, Vallone G, De Martini F, Mataloni P and Cabello A 2009 Phys. Rev. Lett. 103 160401
[11] Bennett C H, DiVincenzo D P, Smolin J A and Wootters W K 1999 Phys. Rev. Lett. 76 722
Bennett C H, DiVincenzo D P, Smolin J A and Wootters W K 1996 Phys. Rev. A 54 3824
[12] Werner R F 1998 Phys. Rev. A 58 826
[13] Horodecki M, Horodecki P and Horodecki R 1999 Phys. Rev. A 59 4206
Terhal B M and Vollbrecht K G H 2000 Phys. Rev. Lett. 85 2625
[14] Vidal G and Werner R F 2002 Phys. Rev. A 65 032314
[15] Rungta P and Caves C M 2003 Phys. Rev. A 67 012307
Lee S, Chi D P, Oh S D and Kim J 2003 Phys. Rev. A 68 062304
[16] Brandão F G S L 2005 Phys. Rev. A 72 022310
Datta A, Flammia S T, Shaji A and Caves C M 2007 Phys. Rev. A 75 062117
[17] Gühne O, Hyllus P, Bruss D, Ekert A, Lewenstein M, Macchiavello C and Sanpera A 2003 J. Mod. Opt. 50 1079
[18] Gühne O, Hyllus P, Bruss D, Ekert A, Lewenstein M, Macchiavello C and Sanpera A 2002 Phys. Rev. A 66 062305
Sanpera A, Tarrach R and Vidal G 1998 Phys. Rev. A 58 826
[19] Terhal B M 2002 J. Theor. Comput. Sci. 287 313
[20] Filip R 2002 Phys. Rev. A 65 062320
[21] Pittenger A O and Rubin M H 2003 Phys. Rev. A 67 012327
[22] Hyllus P 2005 PhD Thesis Universität Hannover, Germany
[23] Kwiat P G and Weinfurter H 1998 Phys. Rev. A 58 2623
[24] Barreiro J T, Langford N K, Peters N A and Kwiat P G 2005 Phys. Rev. Lett. 95 260501
Barbieri M, Cinelli C, Mataloni P and De Martini F 2005 Phys. Rev. A 72 052110
Vallone G, Ceccarelli R, De Martini F and Mataloni P 2008 Phys. Rev. A 78 062303
Vallone G, Ceccarelli R, De Martini F and Mataloni P 2009 Phys. Rev. A 79 030301
Ceccarelli R, Vallone G, De Martini F, Mataloni P and Cabello A 2009 Phys. Rev. Lett. 103 160401
[25] Gao W-B, Lu C-Y, Yao X-C, Xu P, Gühne O, Goebel A, Chen Y-A, Peng C-Z, Chen Z-B and Pan J-W 2008 arXiv:0809.4277[quant-ph]
[26] Arnaut H and Barbosa G A 2000 Phys. Rev. Lett. 85 286
Mair A, Vaziri A, Weih G and Zeilinger A 2001 Nature 412 313
Molina-Terriza G, Torres J P and Torner L 2003 Opt. Commun. 228 115
Vaziri A, Pan J W, Jennewein T, Weih G and Zeilinger A 2003 Phys. Rev. Lett. 91 227902
Franke-Arnold S, Barnett S M, Padgett M J and Allen L 2002 Phys. Rev. A 65 033823
Molina-Terriza G, Torres J P and Torner L 2007 Nature Phys. 3 305
Nagali E, Sciarrino F, De Martini F, Piccirillo B, Karimi E, Marrucci L and Santamato E 2009 arXiv:0902.0740[quant-ph]

[31] Knill E, Laflamme R and Milburn G 2001 Nature 409 46
[32] Beenakker C W J 2006 Electron-hole entanglement in the Fermi sea Quantum Computers, Algorithms and Chaos (Proc. Int. School of Physics ‘Enrico Fermi’ (Varenna, Italy, 2005) vol 162) ed G Casati, D L Shepelyansky, P Zoller and G Benenti (Amsterdam: IOS Press)
Beenakker C W J 2005 arXiv:cond-mat/0508488

[33] Beenakker C W, Di Vincenzo D P, Emary C and Kindermann M 2004 Phys. Rev. Lett. 93 020501
[34] Faoro L and Taddei F 2007 Phys. Rev. B 75 165327