Contrast in Multipath Interference and Quantum Coherence

Kai von Prillwitz\textsuperscript{1}, Łukasz Rudnicki\textsuperscript{1,2} and Florian Mintert\textsuperscript{1,3}

\textsuperscript{1}Freiburg Institute for Advanced Studies, Albert-Ludwigs University of Freiburg, Albertstrasse 19, 79104 Freiburg, Germany
\textsuperscript{2}Center for Theoretical Physics, Polish Academy of Sciences, Aleja Lotników 32/46, PL-02-668 Warsaw, Poland
\textsuperscript{3}Department of Physics, Imperial College London, London SW7 2AZ, United Kingdom

Abstract. We present various ways, in which an interference pattern can be used to rigorously describe coherence properties of a quantum state. Experimentally accessible, statistical moments of such pattern permit to characterize the extent of quantum coherence and to bound commonly employed coherence measures. In that context, we develop a smoothing technique improving the analysis of interference patterns with complicated structure. With explicit examples, we demonstrate that the assessment method works reliably even for clearly sub-optimal experimental sampling.

1. Introduction

Interference resulting from quantum coherence causes an abundance of effects that contradict our classical intuition and that challenge the ambition of the engineer that is hidden within most scientists. Most people would be inclined to negate both the interference of independent photons \cite{1} or of mesoscopic molecules \cite{2,3} if there was no clear experimental evidence for the existence of both effects. Exploiting interference on quantum mechanical systems has been an actively pursued dream for decades, and first realizations emerge \cite{4}.

Despite the fact that interference phenomena in quantum mechanical systems have been observed for more than a century, we still have surprisingly large missing pieces in our understanding of quantum coherence. We learn in our freshmen courses that the coherent superposition of at least two path alternatives (two state-vectors in a more general, abstract description) is necessary for an interference pattern to emerge, and that the achievable contrast increases with the number of states that are coherently over-imposed. While going beyond this qualitative observation, our intuition is however not able to answer the question of how many path-alternatives are needed to generate the particular interference pattern with reduced contrast. The aim of this paper is thus to provide a suitable analysis of interference patterns that brings exactly this information.
The formal definition of quantum coherence requires a set of states $|j\rangle$ that can potentially be over-imposed. In an interferometer these states would correspond to different path alternatives and the number of paths that are being taken coherently is often referred to in terms of the lateral coherence length. In molecular networks one is typically interested in the number of chromophores over which an excitation is coherently distributed; in this case, the coherent delocalization is defined in terms of the excited states of the individual chromophores [5, 6].

In general, a pure state $|\Psi\rangle$ is considered $k$-coherent in terms of a given set of basis states $|j\rangle$, if at least $k$ of the amplitudes $\langle \Psi |k \rangle$ are non-vanishing. Since decoherence processes which are unavoidably present in actual physical situations result in the deterioration of quantum coherence, the description in terms of mixed states or density matrices becomes necessary.

A mixed state $\rho$ can intuitively be understood as an average over pure states, so that averaging over incoherent states $|\Upsilon_i\rangle$ will not result in any interference phenomena. Consequently, any mixed state $\rho = \sum_i p_i |\Upsilon_i\rangle \langle \Upsilon_i|$ with $p_i \geq 0$ is considered incoherent; analogously, any mixed state that can be expressed in terms of an average over pure states with no more than $k$-coherence is not $k+1$-coherent. This motivates the commonly employed definition [7, 8, 9] that a mixed state $\rho^{(k)}$ is $k$-coherent, if any ensemble $\{|\Psi_i\rangle\}$ that satisfies $\rho^{(k)} = \sum_i p_i |\Psi_i\rangle \langle \Psi_i|$ for some set of probabilities $p_i$ contains at least one $k$-coherent state vector. The notion of $k$-coherence is thus similar to the concept of multipartite entanglement, since a mixed state is called $k$-partite entangled if any of its ensemble decompositions involves at least one $k$-partite entangled pure state [11].

Quantum coherence defined in the above manner is a resource [12, 7, 8] in the sense that there are processes whose realization is facilitated by the consumption of coherence. Accordingly tools for the classification and quantification of quantum coherence have been introduced [7, 9]. Currently, however, knowledge of the complete density matrix is required to assess these tools; on the one hand, this poses a high threshold for the analysis of coherence in laboratory experiments, and, on the other hand, the abstract nature of these tools limits the intuition that we might gain from their use. This is why we set out for a characterization of quantum coherence based on information that can directly be read off from an interference pattern only.

### 2. The interference pattern as a coherence classifier

We consider a rather general physical situation in which a superposition of different states is being established, and a certain level of decoherence results in the fact that this superposition is not perfectly coherent. A specific realization of such a situation would be a Mach-Zehnder type of interferometer, as schematically depicted in Fig. 1 where the different path-alternatives define the basis states $|j\rangle$. An incoming object impinges on a beam-splitter that creates a coherent superposition of the basis states. The phase shifters $\phi_i$ permit to generate an interference pattern that can be read off, once the object has crossed the second beam splitter.
The interference pattern is defined as the normalized probability distribution

\[ P(\rho, \phi) = \langle \Phi | \rho | \Phi \rangle = 1 + \sum_{j \neq m} \rho_{jm} \left( e^{i\phi_j - i\phi_m} \right), \quad \text{with} \quad |\Phi\rangle = \sum_{j=1}^{d} e^{-i\phi_j} |j\rangle, \quad (1) \]

to observe an object in the output mode, where \( \rho \) is the state before crossing the second beam splitter. Since on average only one out of \( d \) objects exits through the output mode, the interference pattern is given in terms of an over-normalized state with \( \langle \Phi | \Phi \rangle = d \).

The case of \( d = 2 \) path alternatives corresponds to the original Mach-Zehnder-Interferometer, and tuning a single phase shifter permits to record the interference pattern. In general, the interference pattern is obtained by tuning \( d - 1 \) phase shifters. Beyond this increase of dimensionality also the structure of the pattern can easily get more complicated with growing \( d \) since the dependence of the detection probabilities on \( \phi_i \) gets more sensitive.

The interference perspective brings a natural explanation to the notion of \( k \)-coherence (\( k \)-path coherence) in the case of mixed states. A simple definition of the pure-state coherence stems from the fact that the superposition of different paths is always coherent. For mixed states the above statement is no longer true, i.e. a delocalization over several path alternatives can also be incoherent (or in general partially coherent). In particular, if a \emph{complicated} interference pattern can be decomposed into a sum of only \( k \)-path contributions, it does not represent a scenario with \( k + 1 \)-path coherence (see an example in Fig. 2 with \( d = 3 \) and \( k = 2 \)).

2.1. Moments of the interference pattern

One can certainly obtain some information on coherence from the maximum of the interference pattern

\[ \max_{\phi} P(\rho, \phi), \quad (2) \]
as its value, when larger than \( k - 1 \), unambiguously identifies \( \rho \) to be \( k \)-coherent. In practice, however, this is not necessarily the best choice. In particular, for highly
Figure 2. Upper part: interference patterns corresponding to different types of coherence in mixed-states. Lower part: an interference pattern that results from a state distributed over three path alternatives. Reduced coherence between the three paths can sometimes be expressed as the sum of two-path interference patterns. In that case the mixed state in question will be only 2-path coherent.

coherent states, the interference pattern is a rapidly oscillating function so that optimizations will often identify only local maxima with a resulting under-estimation of coherence properties. Since, again, for highly coherent states, the optimum is given by a very narrow peak, an extremely accurate reconstruction of the interference pattern becomes necessary. This however limits the applicability of Eq. (2) to the data originating from laboratory experiments.

A much more practical alternative would be to employ the uniform statistical moments

\[ m_n = \int_0^{2\pi} \frac{d^d\phi}{(2\pi)^d} [P(\rho, \phi)]^n \equiv \int_0^{2\pi} \frac{d^d\phi}{(2\pi)^d} \langle \Phi | \rho | \Phi \rangle^n . \]  

(3)

The first moment \( m_1 = 1 \) is just the norm of the interference pattern, but the higher moments carry non-trivial information. One would expect that increasing the order of moments improves the identification of \( k \)-coherence, because taking the limit \( \lim_{n \to \infty} (m_n)^{1/n} \) is equivalent to finding the maximum \([2]\). On the other hand the required accuracy on experimental data to assess a moment grows with \( n \).

We thus strive for an approach that is based on moments of reasonably low order, but that nevertheless is strong in the identification of quantum coherence. As we will show later-on, the \textit{generalized moments}

\[ Q_n [\rho] = \int d^d\phi \ F(\phi) [P(\rho, \phi)]^n , \]  

(4)

defined in terms of a suitably chosen \( d \)-dimensional probability distribution \( F(\phi) \) characterize coherence properties rather reliably.

The best performance of the above quantity is certainly expected to take place if the maximum of \( F(\phi) \) coincides with the maximum of the interference pattern. This however calls for the search of an optimal probability distribution which might be flawed by the same issues as encountered for Eq. [2]. On the other hand, with a sufficiently wide
distribution the optimization landscape is substantially flatter than in Eq. 2 what eases the optimization a lot. Beyond that, as we will show with specific examples later-on, even a clearly sub-optimal probability distribution still yields rather strong criteria.

2.2. Lower bounds for the measure of coherence

Starting from the next section we focus on a particular choice of the probability distribution $F(\phi)$. Before we go into this direction, let us first establish a general link between the quantities $Q_n[\rho]$ and the theoretical approach to quantum coherence. The quantitative description of the coherence can be patterned after the entanglement theory, so that various proper measures (fulfilling certain monotonicity requirements) like $l_1$-norm of coherence

$$C_{l_1}[\rho] = \sum_{j \neq m} |\rho_{jm}|,$$

(5)

can be introduced [7]. The above quantity could be considered as probably the most intuitive measure, since it involves linear amplitude contributions from all off-diagonal elements.

Combining both defining Eqs. 1 and 4 we can easily obtain an estimate:

$$|Q_n[\rho]| \leq \int d^d \phi F(\phi) \left[ 1 + \sum_{j \neq m} |\rho_{jm}| \right]^n = \left[ 1 + \sum_{j \neq m} |\rho_{jm}| \right]^n. \quad (6)$$

For every $n$ and any choice of $F(\phi)$, this inequality leads to a lower bound on the $l_1$-norm of coherence:

$$|Q_n[\rho]|^{1/n} - 1 \leq C_{l_1}[\rho]. \quad (7)$$

While an experimental determination of the exact value of the measure (5) seems to be a formidable task, it can be lower bounded by any generalized moment based on the experimentally accessible interference pattern. It is thus not surprising, that $Q_n[\rho]$ contains enough information to provide a valuable description of quantum coherence as well as handy criteria allowing for identification of $k$-coherence.

2.3. Threshold values for the wrapped normal distribution

Since the interference pattern defined in (1) is given in terms of all the phases $\phi_j$ as independent variables, it is reasonable to define $F(\phi)$ in terms of independent distributions for each phase, i.e. $F(\phi) = \prod_{j=1}^d f(\phi_j)$. The $j$th angle $\phi_j$ becomes thus associated with the one dimensional distribution $f(\phi_j)$ described further in terms of its unique mean value $\mu_j$ and the standard deviation $\sigma$ which we assume to be equal for all angles.

An evaluation of Eq. 4 requires the construction of so called trigonometric moments

$$\Theta_n(\mu) \equiv \int_0^{2\pi} d\phi f(\phi) e^{i n \phi}, \quad (13)$$

defined for integer $n$. Due to the fact, that the phases $\phi_j$ are defined only in an interval of width of $2\pi$, this step can be done explicitly for most typically employed distributions like the Lorentz or Gauss distributions. The integrals
can be easily evaluated if one takes advantage of a \textit{wrapped} version of the involved distribution \cite{13}. In the case of the wrapped normal distribution, the trigonometric moments are equal to the characteristic function of the normal (unwrapped) distribution evaluated at integer arguments
\[ \Theta_n (\mu) = e^{in\mu}R_n \] with
\[ R_n = e^{-n^2\sigma^2/2} \].
With the help of the function \( \Theta_n (\mu) \) we can perform the integration that is required by Eq. \ref{eq:4} and express the generalized moments as:
\begin{equation}
Q_n (\mu, \sigma) = \sum_{i_1, i_2, \ldots, i_{2n}=1}^d \rho_{i_1i_{n+1}}\rho_{i_2i_{n+2}}\cdots\rho_{i_{2n}i_{2n}} \prod_{j=1}^d \Theta_{n_j} (\mu_j),
\end{equation}
where
\begin{equation}
n_j = \sum_{i=1}^n \delta_{j,i} - \sum_{i=n+1}^{2n} \delta_{j,i}.
\end{equation}
As argued above, in practice, the use of low order moments is desirable; we will therefore focus in the following on \( n = 1, 2, 3 \), but there is no fundamental obstacle for generalizations to higher values of \( n \).
Before one can use Eq. \ref{eq:4} or Eq. \ref{eq:8} to rigorously identify coherence properties, one needs to find the maximum that \( Q_n \) can adopt for \( k \)-coherent states. In the present case such an optimization can be done explicitly, confirming that the maximum among all \( k \)-coherent states is provided by
\begin{equation}
|W_k\rangle = \frac{1}{\sqrt{k}} \sum_{j=1}^k e^{i\varphi_j} |j\rangle,
\end{equation}
i.e. a perfectly balanced coherent superposition of \( k \) basis states.
To arrive at this conclusion, one may first realize that the generalized moments \( Q_n \) are convex functions, \textit{i.e.}
\begin{equation}
Q_n[\eta \rho_1 + (1-\eta)\rho_2] \leq \eta Q_n[\rho_1] + (1-\eta)Q_n[\rho_2],
\end{equation}
for \( 0 \leq \eta \leq 1 \) and any pair of density matrices \( \rho_1 \) and \( \rho_2 \). This is a direct consequence of the two facts that the \( n \)th power of a linear functional like \( \langle \Phi | \rho | \Phi \rangle^n \) is convex, and that the integral \( \int d^n\phi \) preserves convexity. Since states \( \rho^{(k)} \) that are at most \( k \)-coherent (for any value of \( k \)) define a convex set (\textit{i.e.} \( \eta \rho^{(k)}_1 + (1-\eta)\rho^{(k)}_2 \) is no more than \( k \)-coherent) the maximum of \( Q_n \) over \( k \)-coherent density matrices is always reached for a pure state.
The most general \( k \)-coherent pure state reads
\begin{equation}
|\Psi^{(k)}\rangle = \sum_{j=1}^k \sqrt{\lambda_j} e^{-i\varphi_j} |j\rangle, \text{ with } \lambda_j \geq 0,
\end{equation}
assuming (without loss of generality) that exactly the first \( k \) basis states are comprised with non-vanishing weight \( \lambda_j \) in the coherent superposition. In \textit{Appendix A} it is shown that the optimization of the phase factors \( e^{-i\varphi_j} \) can be performed independently of the optimization over the real amplitudes \( \lambda_j \), and that the maximum is obtained if the \( \varphi_j \) coincide with the expectation values \( \mu_j \) of \( f(\phi_j) \). Also the remaining
optimization over the $\lambda_j$ can be performed very generally. As shown in Appendix A, the quantity to be optimized is always a Schur-concave function which is maximized for $\lambda_1 = \ldots = \lambda_k = 1/k$.

That is, the maximum of $Q_n$ that can be adopted for $k$-coherent states reads

$$Q_n^{(k)}(\sigma) = k^{1-n} \sum_{l=0}^{n^2} v_l^{(n,k)} e^{-l\sigma^2},$$

with the coefficients $v_l^{(n,k)}$ given by ($K_m = k - m$):

$$v_0^{(1,k)} = 1, \quad v_1^{(1,k)} = K_1,$$
$$v_0^{(2,k)} = 2k - 1, \quad v_1^{(2,k)} = 4, \quad v_2^{(2,k)} = K_1 K_2 K_3,$$
$$v_3^{(2,k)} = 2K_1 K_2, \quad v_4^{(2,k)} = K_1,$$
$$v_0^{(3,k)} = 4 - 9k + 6k^2, \quad v_1^{(3,k)} = 3K_1 (11 + 3k(2k - 5)),$$
$$v_2^{(3,k)} = 9K_1 K_2^2 K_3, \quad v_3^{(3,k)} = K_1 K_2^2 (45 + kK_10),$${
$$v_4^{(3,k)} = 3K_1 (55 + 2kK_9) - 52), \quad v_5^{(3,k)} = 9K_1 K_2 K_3,$$
$$v_6^{(3,k)} = 2K_1 K_2 K_3, \quad v_7^{(3,k)} = 6K_1 K_2, \quad v_8^{(3,k)} = 0, \quad v_9^{(3,k)} = K_1.$$  

Any excess of the value $Q_n^{(k)}(\sigma)$ is an unambiguous identification of coherence properties beyond $k$-coherence.

Note that for $n = 1$ we can simplify

$$Q_1(\mu, \sigma) = 1 + e^{-\sigma^2} \sum_{i_1 \neq i_2} \rho_{i_1 i_2} \prod_{j=1}^d e^{i n_j \mu_j},$$

with $n_j \in \{-1, 0, 1\}$. We observe that the $\sigma$-dependent contribution factors out, so that the true benefits related to the long tail of the distribution come with $n > 1$.

3. Numerical results

Having established the rigorous properties of the generalized moments $Q_n$, it remains to identify the range of optimal values of $\sigma$. While narrow distributions require an accurate identification of the maximum in the interference pattern, they yield strong criteria if this maximum is found. On the other hand wider distributions give robust but potentially weaker criteria.

To address this questions, let us consider an ensemble of density matrices, and check how many states are detected to be $k$-coherent. Ideally, one would like to identify the ratio of detected and all $k$-coherent states; but since the latter is unknown, we will focus on the percentage of detected $k$-coherent states (subsequently called the detection ratio $R$) within a given ensemble, which typically is substantially smaller than unity.

Let us assume that the centers of the distributions $\mu$ do not fit perfectly with the maximum of the interference pattern, but are shifted by some vector $\delta$. Since one can
always redefine the origin of the phase variables $\phi_i$, we can restrict the following analysis to states that give rise to interference patterns with the global maximum at $\phi_i = 0$ (for $i = 1, ..., d$).

The above assumptions imply that the generalized moments are defined in terms of the probability distribution with $\mu_i = \delta_i$ so that the $\delta_i$ define the centers of the underlying one-dimensional distributions. If all the $\delta_i$ vanish, the distribution is centered around the maximum of the interference pattern, and since the entire interference pattern is invariant under a global phase shift, this holds also if all the $\delta_i$ coincide. We therefore define the invariant deviation vector $\tilde{\delta}_i = \delta_i - \bar{\delta}$, where the term $\bar{\delta} = \sum_{i=1}^{d} \delta_i / d$ removes the above-mentioned ambiguity.

Since the ability to identify coherence is expected to depend on the specific choice of the $\delta_i$, we average the detection ratio with a centered Gaussian distribution, associated to the components of the invariant deviation vector. All the $\tilde{\delta}_i$ are treated as independent random variables

$$\tilde{\delta}_i \sim \mathcal{N} \left(0, \sigma_G^2\right), \quad \langle R \rangle \quad \text{(22)}$$

classified by a single width $\sigma_G$. The averaged detection ratio $\langle R \rangle$ depends thus on the parameter $\sigma_G$ which in a natural way quantifies the shift between the mean value $\mu$ and the maximum of the interference pattern. The particular choice $\sigma_G = 0$ refers to the case when the maximum of the pattern is always found, while positive widths cause random, but statistically controlled shifts. Note that in order to obtain the desired width in (22), the primary, non-invariant parameters $\delta_i$ need to be distributed according to $\mathcal{N} \left(0, \frac{d}{d-1} \sigma_G^2\right)$. In other words, the global phase invariance provides the decrease of the width by the factor $\sqrt{(d-1)/d}$, so that in the limit of large $d$ the invariance in question does not play any role.

Let us start with a simple ensemble

$$\rho_a = a \left| \Psi_W \right\rangle \left\langle \Psi_W \right| + \frac{1-a}{d} \mathbf{1}, \quad \text{(23)}$$
with \( |\Psi_W \rangle = \sum_{i=1}^{d} |i\rangle / \sqrt{d} \), parametrized by the real parameter \( a \) that satisfies \( 1 \geq a \geq 0 \). All \( d \) basis states are populated with equal weight independently of \( a \); for \( a = 1 \), \( \rho_a \) describes a perfectly coherent superposition, and \( a = 0 \) corresponds to the situation without any phase coherence. The averaged detection ratio for the third moment and \( d = k = 7 \) is shown in Fig. 3 (left plot) as function of \( \sigma \) and \( \sigma_G \). As expected, \( \langle R \rangle \) decays monotonically with \( \sigma_G \) for fixed \( \sigma \) parameter. A more interesting feature can be captured while looking at the dependence on \( \sigma \) with fixed \( \sigma_G \). There is a broad range of \( \sigma_G \), for which \( \langle R \rangle \) is (up to some optimal point) an increasing function of \( \sigma \). Fig. 3 (right plot) explicitly shows the above effect (in comparison with the first and second moments) present in the exemplary intersection taken at \( \sigma_G = 0.4 \). The largest value of \( \langle R \rangle \) is in this case not achieved for \( \sigma = 0 \), but for a finite width \( \sigma \approx 0.9 \). This observation provides a perfect illustration of the role played by the width parameter \( \sigma \). When the maximum of the interference pattern cannot be identified faithfully, the long tail of the distribution still allows one to capture the coherence properties of the state.

In order to test the universal validity of the above observations we performed a similar computation for a general ensemble of states \( \rho = U \Lambda U^\dagger \), where \( \Lambda \) is a diagonal matrix describing the spectrum of \( \rho \) while \( U \) is a unitary transformation. The matrix \( U \) is drawn from the Circular Unitary Ensemble (CUE \[14, 15\]), while \( \Lambda \) contains squared absolute values of components from a single column of a unitary matrix, which was also generated with the help of CUE.

| \( k \) | \( \sigma_G \) | \( \langle R \rangle_\text{ref} \) | \( \langle R \rangle_2^{\max} / \langle R \rangle_\text{ref}; n = 2 \) | \( \langle R \rangle_3^{\max} / \langle R \rangle_\text{ref}; n = 3 \) | \( \sigma_{\text{max}} \) |
|---|---|---|---|---|---|
| \( \geq 4 \) | 0.3 | 0.45 | 1.00 | 1.01 | 0.3 |
| 0.8 | 0.11 | 1.10 | 1.21 | 0.7 |
| \( \geq 5 \) | 0.3 | 0.10 | 1.03 | 1.05 | 0.5 |
| 0.8 | 0.007 | 1.21 | 1.48 | 0.8 |
| \( \geq 6 \) | 0.3 | 0.002 | 1.13 | 1.21 | 0.6 |

Table 1. Maximal detection ratios for the first three moments (\( \langle R \rangle_\text{ref} \) corresponds to the first moment and by construction is independent of \( \sigma \)) and several values of \( k \) and \( \sigma_G \). By \( \sigma_{\text{max}} \) we denote the value of \( \sigma \) for which the detection ratios in the case \( n = 2 \) and \( n = 3 \) are maximal.

For several values of \( k \) and \( \sigma_G \) we found (see Table 1) the reference detection ratio \( \langle R \rangle_\text{ref} \) obtained with the help of \( Q_1 \), and the maximal (achieved for particular \( \sigma_{\text{max}} \)) detection ratios \( \langle R \rangle_2^{\max} \) and \( \langle R \rangle_3^{\max} \) corresponding to \( Q_2 \) and \( Q_3 \). The maximal detection ratios are optimal with respect to the parameter \( \sigma \) scanned with a step size equal to 0.1. In all cases we can observe the desired increase of the detection ratio with the number \( n \) of the moment used. We can also see that the optimal scenario is typically provided by a finite width of the distribution \( F(\phi) \), so that the long tail of this distribution indeed plays a significant role.
4. Analysis for mixed states

In Section 2 we showed that the coherence classifiers based on \( Q_n \) are convex with respect to the density matrix, what further implies that their maxima are provided by the pure states \( (10) \). The efficiency of the present criteria can thus be improved if additional information on the purity \( \mathcal{P} = \text{Tr}\rho^2 \) is available. Our aim is to study the purity-dependent threshold values \( Q_n^{(k)} (\sigma, \mathcal{P}) \) and compare them with the upper bounds given by the formula \( (13) \).

In our analysis we focus on the first non-trivial case \( n = 2 \). First of all, we find that the global maximum (corresponding to the case \( k = d \)) for the second moment is provided by the state (see Appendix B)

\[
\rho (\mathcal{P}) = \frac{1}{d} \left( 1 - \sqrt{\frac{d\mathcal{P} - 1}{d - 1}} \right) \mathbb{I}_d + \sqrt{\frac{d\mathcal{P} - 1}{d - 1}} |\Psi_d\rangle \langle \Psi_d|.
\] (24)

Moreover, if \( \mathcal{P} \leq \mathcal{P}_k \) with

\[
\mathcal{P}_k = \frac{k^2 - 2k + d}{d(d - 1)},
\] (25)

the state (24) is at most \( k \)-coherent, so that it provides the value of \( Q_2^{(k)} (\sigma, \mathcal{P}) \) in this range of purity.

**Figure 4.** The threshold values \( Q_2^{(k)} = Q_2^{(k)} (1, \mathcal{P}) \) found numerically for \( k = 2, 3, 4, 5 \) and \( d = 5 \). Black line denotes the analytical result obtained for the state (24).

In order to capture the behavior of the threshold values for \( \mathcal{P}_k < \mathcal{P} < 1 \) we have numerically evaluated their dependence on purity, in the case \( d = 5 \) and \( \sigma = 1 \). The black line in Fig. 4 corresponds to \( Q_2^{(5)} (1, \mathcal{P}) \) and perfectly covers the range \( \mathcal{P} \leq \mathcal{P}_k \) in all cases \( k = 2, 3, 4, 5 \). The most important result is however brought by the fact that the threshold values for \( \mathcal{P} \geq \mathcal{P}_k \) are almost purity-independent. This observation strongly suggests that the coherence classifiers, while being maximal for pure states, can be reliably applied also for general mixed states. When the state is highly mixed,
it very likely is the state with a weak coherence or with no coherence at all. As can
however be seen in Fig. 4, even for low values of purity the threshold values \(^{(13)}\) are
still below the black borderline provided by the maximal state \(^{(24)}\). There are thus
highly-mixed states, such that their coherence can be confirmed with the help of the
pure-states threshold values \(^{(13)}\). The robustness of the studied criteria is especially
visible for small values of \(k\) and is expected to hold in the case when \(k \ll d\). Such
scenario, when the coherence is distributed between a small number of sites (small \(k\)
being a part of the large system (large \(d\)), naturally appears in the context of quantum
biology.

5. Conclusion

Since the quantum coherence is a prerequisite for the interference, information carried by
the interference pattern shall allow for accurate description of the coherence properties
of the quantum state. As we have seen, the interference pattern indeed permits to
draw rigorous conclusions on coherence. From a practical point of view, the freedom in
choice of sampling \([encoded in the number of the statistical moment and the distribution
\(F(\phi)\)]\ as well as the possibility to include additional information (like purity) make this
approach flexible, so that it can be tailored for the specific properties of a system under
investigation.

Here, we have been considering the case of \(d\) independently adjustable phases, but the underlying framework works independently of such assumptions. Limitations
on experimentally variable quantities may be compensated through suitably chosen
distributions with variable widths that reflect the realistically achievable sampling.
Moreover, we developed the general framework which can utilize a wrapped version of an
arbitrary probability distribution. While the normal distribution seems to be the most
natural first choice, more sophisticated distributions can even better support a particular
experimental realization. Beyond the conceptual connection between the directly
observable interference pattern and the underlying, abstract coherence properties, the
present approach thus provides a versatile method to characterize coherence properties
in a wide range of systems.

Acknowledgments

We thank Alexander Stibor and Björn Witt for interesting and helpful conversations.
Financial support by the European Research Council under the project Odycquent is
gratefully acknowledged. The article processing charge was funded by the German
Research Foundation (DFG) and the Albert Ludwigs University Freiburg in the funding
programme Open Access Publishing.
Appendix A. Derivation of threshold values

We start the derivation by inserting $\rho^{(k)} = |\Psi^{(k)}\rangle \langle \Psi^{(k)}|$ with $|\Psi^{(k)}\rangle$ given by (12) into the expression (8):

$$Q_n (\mu, \sigma) = \sum_{i_1, i_2, \ldots, i_{2n}}^{d} \sqrt{\lambda_{i_1} \lambda_{i_2} \ldots \lambda_{i_{2n}}} \prod_{i=1}^{d} R_{n_i} \cos \left[ \sum_{j=1}^{d} n_j (\mu_j - \varphi_j) \right]$$

$$\leq \sum_{i_1, i_2, \ldots, i_{2n}}^{d} \sqrt{\lambda_{i_1} \lambda_{i_2} \ldots \lambda_{i_{2n}}} \prod_{j=1}^{d} R_{n_j}$$

$$\equiv g_n (\lambda) \quad \text{(A.1)}$$

An estimate $\cos (\cdot) \leq 1$ applied in the second line implies that the maximum of $Q_n (\mu, \sigma)$ with respect to $\mu$ is achieved when the peak position of the probability distribution coincides with the maximum of the interference pattern.

In the next step we employ the concept of Schur-concavity [16, 17]. For any two vectors $\lambda$ and $\lambda'$ such that $\lambda'$ is majorized by $\lambda$ ($\lambda' \prec \lambda$) and any Schur-concave function $g (\lambda)$ one gets $g (\lambda') \geq g (\lambda)$. In the case of pure, $k$-coherent states all vectors $\lambda$ majorize the uniform vector $\lambda_k = \left( \frac{1}{k}, \frac{1}{k}, \ldots, \frac{1}{k}, 0, \ldots, 0 \right)$. To finish the proof we thus only need to show that the function $g_n (\lambda)$ defined in (A.1) is Schur-concave for $n = 2, 3$. To this end it is however sufficient to show that $g_n (\lambda)$ satisfies the well known Schur condition:

$$S [g] = (\lambda_i - \lambda_j) \left( \frac{\partial g}{\partial \lambda_i} - \frac{\partial g}{\partial \lambda_j} \right) \leq 0 \quad \forall i, j = 1, 2, \ldots, d.$$ \hspace{1cm} \text{(A.2)}

Since $g_n (\lambda)$ is symmetric (under permutation of its arguments) we shall only show this inequality for one pair of indices.

To proceed further we need an explicit form of both functions. Define

$$W^B_A = \prod_{i=1}^{A} \lambda_i \prod_{n=1}^{B} \sqrt{\lambda_n}, \quad G_{AB} = \sum_{(1, d, \neq)} W^B_A \quad \text{(A.3)}$$

where the sum $\sum_{(1, d, \neq)}$ is taken over all indices $i_1, \ldots, i_A, j_1, \ldots, j_B$ which are assumed to be pairwise different and each index runs from 1 to $d$. We have:

$$g_2 (\lambda) = 1 + (1 + R_2^2) G_{20} + 2 R_1^2 G_{02} + 2 R_1^2 (1 + R_2) G_{12} + R_1^4 G_{04}, \quad \text{(A.4)}$$

$$g_3 (\lambda) = 1 + 3 R_1^2 G_{02} + 3 \left( 1 + R_2^2 \right) G_{20} + 6 R_1^2 (1 + R_2) G_{12} + 3 R_1^4 G_{04}$$

$$+ 2 \left( 1 + 3 R_2^2 \right) G_{30} + 3 R_1^2 \left( 3 + 4 R_2 + 3 R_2^2 \right) G_{22} + 6 R_1^4 (1 + R_2) G_{14}$$

$$+ R_1^6 G_{06} + 6 R_1 (2 R_1 + R_1 R_2 + R_2 R_3) \sum_{\neq} \sqrt{\lambda_i} \lambda_j \sqrt{\lambda_m} \quad \text{(A.5)}$$

$$+ 2 R_1^3 (3 R_1 + R_3) \sum_{\neq} \sqrt{\lambda_i} \sqrt{\lambda_j \lambda_m \lambda_l} + (3 R_1^2 + R_3^2) \sum_{\neq} \sqrt{\lambda_i} \lambda_j^3.$$
The symbol $\sum_{\neq}$ represents the sum from 1 to $d$ with all indices being pairwise different.

Since all parameters $R_n$ are non-negative we can treat each term in (A.4-A.5) separately. We find
\[
\frac{\partial G_{AB}}{\partial \lambda_1} = A \sum_{(2,d,\neq)} W^{B}_{A-1} + \frac{B}{2\sqrt{\lambda_1}} \sum_{(2,d,\neq)} W^{B-1}_{A},
\]
so that:
\[
S [G_{AB}] = - \sum_{(3,d,\neq)} \left\{ (\lambda_1 - \lambda_2)^2 \left[ A (A - 1) W^{B-2}_{A} + \frac{B (B - 1)}{2\sqrt{\lambda_1} \lambda_2} W^{B-2}_{A} \right] 
+ B (\lambda_1 - \lambda_2) \left( \sqrt{\lambda_1} - \sqrt{\lambda_2} \right) \left[ AW^{B-1}_{A-1} + \frac{1}{2\sqrt{\lambda_1} \lambda_2} W^{B-1}_{A} \right]
+ \frac{AB}{2\sqrt{\lambda_2}} (\lambda_1 - \lambda_2) \left( \lambda_1^{3/2} - \lambda_2^{3/2} \right) W^{B-1}_{A-1} \right\} \leq 0. \tag{A.7}
\]
g$_2$ ($\lambda$) is thus Schur-convex as a linear combination of Schur-concave functions $G_{AB}$. Note that this result is valid independently of the probability distribution defining the coefficients $R_n$. The function $g_3$ ($\lambda$) involves additional terms which are not of the $G_{AB}$ form and are not Schur-concave. By applying the same procedure as in (A.7) it is however possible to show that the combinations:
\[
2R_1^3 (3R_1 + R_3) \sum_{\neq} \sqrt{\lambda_i} \sqrt{\lambda_j \lambda_m \lambda_l} + 6R_1^4 (1 + R_2) G_{14}, \tag{A.8}
\]
\[
6R_1 (2R_1 + R_1 R_2 + R_2 R_3) \sum_{\neq} \sqrt{\lambda_i} \sqrt{\lambda_j \lambda_m} + (3R_1^2 + R_3^2) \sum_{\neq} \sqrt{\lambda_i \lambda_j}^3
+ 3R_1^2 (3 + 4R_2 + 3R_2^2) G_{22},
\]
are Schur-concave if a single condition $R_1 \geq R_3$ is satisfied. For the wrapped normal distribution we get $R_1 = e^{-\sigma^2/2} \geq e^{-9\sigma^2/2} = R_3$, so that $g_3$ ($\lambda$) is Schur-convex as well.

Appendix B. Derivation of the global maximum

Since the generalized moments (8) are real and non-negative we have the following, simple estimate:
\[
Q_2 (\mu, \sigma) = |Q_2 (\mu, \sigma)| \leq \sum_{i_1,i_2,i_3,i_4=1}^d |\rho_{i_1 i_3}| |\rho_{i_2 i_4}| \prod_{j=1}^d R_{n_j}. \tag{B.1}
\]
Note that the right hand side does not depend on $\mu$, so the same upper bound applies to the quantity optimized with respect to the center of the distribution $F(\phi)$.
\[
Q_2 (\mu, \sigma) \leq 1 + (1 + R_2^2) \sum_{\neq} |\rho_{ij}|^2 + 2R_1^2 \sum_{\neq} |\rho_{ij}|
+ 2R_1^2 (1 + R_2) \sum_{\neq} |\rho_{im}| |\rho_{ji}| + R_1^4 \sum_{\neq} |\rho_{ij}| |\rho_{kl}|. \tag{B.2}
\]
The above bound saturates if \( \mathbb{R} \ni \rho_{ij} \geq 0 \) for all \( i \neq j \) so that the density matrix has only real and non-negative entries.

Using the arithmetic-geometric mean inequality we get the following estimates:

\[
\sum_{\neq} |\rho_{im}| |\rho_{ji}| \leq (d - 2) \sum_{\neq} |\rho_{ij}|^2, \quad (B.3)
\]

\[
\sum_{\neq} |\rho_{ij}| |\rho_{ml}| \leq (d - 3) (d - 2) \sum_{\neq} |\rho_{ij}|^2. \quad (B.4)
\]

Since we assume that the purity \( \mathcal{P} = \text{Tr} \rho^2 \) is fixed we get

\[
\sum_{\neq} |\rho_{ij}|^2 = \mathcal{P} - \sum_{i=1}^d \rho_{ii}^2, \quad (B.5)
\]

what also implies that

\[
\sum_{\neq} |\rho_{ij}| \leq \sqrt{d (d - 1) \left( \mathcal{P} - \sum_{i=1}^d \rho_{ii}^2 \right)}. \quad (B.6)
\]

The maximum of (B.5), as well as maxima of (B.3, B.4) and (B.6) are provided by the uniform choice

\[
\forall i \rho_{ii} = \frac{1}{d}. \quad (B.7)
\]

This observation immediately leads to a state independent maximum of \( Q_2 \). The above maximum might be saturated only when the inequalities used in (B.3, B.4) and (B.6) saturate too, i.e. when

\[
\forall_{i \neq j} |\rho_{ij}| = \frac{1}{d} \sqrt{\frac{d \mathcal{P} - 1}{(d - 1)}}, \quad (B.8)
\]

The last conclusion proves Eq. (24), showing that the maximum is always global.

The next step is to determine when the state (24) does not happen to be \( k+1 \)-coherent. For example, if \( \mathcal{P} = 1 \), then the global maximum is attained by the \( d \)-coherent pure state. We start with the following observation. To realize any \( k \)-coherent state it is sufficient to consider the form:

\[
\rho^{(k)} = \sum_{m=1}^{D_k} \sum_{i,j \in I_m(k)} \Xi^{(m)}_{ij} |i⟩⟨j|, \quad D_k = \binom{d}{k}, \quad (B.9)
\]

with \( I_m(k) \) being for each \( m = 1, \ldots, D_k \) a unique set of \( k \) different indices taken from \( \{1, \ldots, d\} \). By \( \Xi^{(m)}_{ij} \) we denote arbitrary (not normalized) \( k \times k \) non-negative definite matrices.

We shall now construct \( \rho^{(k)} \) with all diagonal elements equal to \( \kappa \equiv 1/d \) and all off-diagonal elements equal to \( 0 \leq \beta \leq 1/d \). The maximal value of \( \beta \) is provided by the case when for every \( m \), all diagonal elements of \( \Xi^{(m)}_{ij} \) are equal to some \( h > 0 \) and
all corresponding off-diagonal elements are given by some $b \leq h$. From combinatorial considerations we find that:

$$\frac{1}{d} \equiv \kappa = \binom{d-1}{k-1} h, \quad \beta = \binom{d-2}{k-2} b.$$  \hfill (B.10)

We thus obtain

$$\text{Tr} \left( \rho^{(k)} \right)^2 = \frac{1}{d} + d (d-1) \beta^2 = \frac{1}{d} + d (d-1) \left[ \binom{d-2}{k-2} \right]^2 b^2 \leq \frac{1}{d} + d (d-1) \left[ \binom{d-2}{k-2} \right]^2 h^2 = \mathcal{P}_k.$$  \hfill (B.11)

In that way we have recovered the range $[25]$.

References

[1] Hong C K, Ou Z Y and Mandel L 1987 Phys. Rev. Lett. 59 2044
[2] Nairz O, Arndt M and Zeilinger A 2003 Am. J. Phys. 71 319
[3] Facchi P 2004 J. Mod. Opt. 51 1049
[4] Eibenberger S, Cheng X, Cotter J P and Arndt M 2014 Phys. Rev. Lett. 112 250402
[5] Fassoli F, Dinshaw R, Arpin P C and Scholes G D 2014 J. R. Soc. Interface 11 20130901
[6] Scholak T, de Melo F, Wellens T, Mintert F and Buchleitner A 2011 Phys. Rev. E 83 021912
[7] Cramer M, Baumgratz T and Plenio MB 2014 arXiv:1311.0275
[8] Girolami D 2014 [arXiv:1403.2446]
[9] Levi F and Mintert F 2014 New J. Phys. 16 033007
[10] Smyth C and Scholes G D 2014 [arXiv:1401.7412]
[11] Horodecki R et al 2009 Rev. Mod. Phys. 81 865
[12] Vedral V Arndt M and Juffmann T 2009 HFSP Journal 3 386
[13] Jammalamadaka S R and SenGupta A 2001 Topics in circular statistics (World Scientific)
[14] Życzkowski K and Kuś M 1994 J. Phys. A: Math. Gen. 27 4235
[15] Życzkowski K and Sommers H-J 2001 J. Phys. A: Math. Gen. 34 7111
[16] Wang S, Zhang T and Xi B 2011 Schur convexity for a class of symmetric functions. In Information Computing and Applications (Springer Berlin Heidelberg)
[17] Marshall A W and Olkin I 1979 Inequalities: Theory of Majorization and Its Applications (Academic Press)