THE CGLMP BELL INEQUALITIES AND QUANTUM THEORY

B. J. Dalton$^{1,2}$

$^1$Centre for Quantum and Optical Science, Swinburne University of Technology, Melbourne, Victoria 3122, Australia
$^2$School of Physics and Astronomy, University of Glasgow, Glasgow G12 8QQ, United Kingdom

*Correspondence: bdalton@swin.edu.au

Abstract: Quantum non-locality tests have been of interest since the original EPR paper. The present paper discusses whether the CGLMP (Bell) inequalities obtained by Collins et al are possible tests for showing that quantum theory is not underpinned by hidden variable theory (HVT). It is contended that CGLMP inequalities only require that HVT applies without requiring it to be local, so hidden variable theories of a more general class could also be ruled out. Restricting the HVT to be local does not result in a weaker CGLMP inequality. The HVT involved is based on a probability for outcomes of simultaneous measurements of pairs of observables corresponding to non-commuting quantum operators, which is allowed in classical theory. Although the CGLMP inequalities involve probabilities for measurements that are compatible with the Heisenberg uncertainty principle and for which both HVT and quantum expressions exist, there is no unambiguous quantum measurement process linked to the probabilities in the CGLMP inequalities. Quantum measurements corresponding to the different classical measurements that give the same CGLMP probability are found to yield different CGLMP probabilities. However, violation of a CGLMP inequality based on any one of the possible quantum measurement sequences is sufficient to show that the Collins et al HVT does not predict the same results as quantum theory, such as for a state considered in their paper (though for observables lacking a physical interpretation). In spite of the problems of comparing the HVT inequalities with quantum expressions, it is concluded that the CGLMP inequalities are suitable for ruling out even non-local hidden variable theories - a significance not previously recognised.

Keywords: Hidden variable theory, Quantum non-locality, Bell inequalities, Copenhagen quantum interpretation, Localism
1 Introduction

The concept of hidden variable theory was introduced in papers by Einstein, Schrodinger, Bell and Werner ([1], [2], [3], [4], [5]). Einstein suggested that quantum theory, though correct was incomplete - in that the probabilistic measurement outcomes predicted in quantum theory could be just the statistical outcome of an underlying deterministic theory, where the possible measured outcomes for all observables always have specific values, and measurement merely reveals what these values are. Hence observable quantities (such as position and momentum) could be regarded as elements of reality irrespective of whether an actual measurement has taken place. The EPR paradox is based on this assumption and involved an entangled state for two well-separated and no longer interacting distinguishable particles, which had well-defined values for the position difference and the momentum sum. For this state, measuring the position (or the momentum) for the first particle would instantly affect the outcome for measuring the position (or the momentum) of the second particle (a feature we now refer to as steering). Einstein regarded this as being in conflict with causality. The paradox is that by measuring the position for the first particle, the position for the second particle is then known without doing a measurement. So by then measuring the momentum for the second particle a joint precise measurement of both the position and momentum for the second particle would have occurred - apparently contradicting the Heisenberg uncertainty principle. The Schrodinger cat paradox [3] is another example, but now involving a macroscopic sub-system (the cat) in an entangled state with a microscopic sub-system (the two state radioactive atom). From the Einstein concept of reality the cat must be either alive or dead even before the box is opened to see what is the case. However, in the Copenhagen interpretation of quantum theory (see [6] for a discussion), the values for observables do not have a presence in reality until measurement takes place. Hence, from the Copenhagen viewpoint the cat is neither dead nor alive until the box is opened - which is a paradox in the Einstein concept of reality but not from the Copenhagen viewpoint. Bohm [7] described a similar paradox to EPR, but now involving a system consisting of two spin 1/2 particles in a singlet state, and where the observables were spin components with quantised measured outcomes rather than the continuous outcomes that applied to EPR.

Einstein believed that an underlying realist theory could be found, based on what are now referred to as hidden variables - which would specify the real or underlying state of the system. However, it was not until 1965 before a quantitative general form for local hidden variable theory was proposed by Bell [4]. This was relevant for the EPR paradox and could be tested in experiments. In its simplest form, the key idea is that hidden variables are specified probabilistically when the state for the composite system is prepared, and these would determine the actual values for all the sub-system observables even after the
sub-systems have separated - and even if the observables were *incompatible* with simultaneous precise measurements (such as two different spin components). In the EPR experiment they would specify *both* the position and momentum for each distinguishable particle. More elaborate versions of local hidden variable theory only require the hidden variables to determine the probabilities of measurement outcomes for each sub-system observable, with the overall expressions for the joint sub-system measurement outcomes being obtained in accordance with classical probability theory (see [8], [9], [10], [11], [12] for a description). States where the joint probability can be described via local hidden variable theory are referred to as *Bell local*. Quantum states for composite systems that could be described by local hidden variable theory were such that certain inequalities would apply involving the mean values of products for the results of measuring pairs of observables for both sub-systems - the *Bell inequalities* [4], [13]. States for which a local hidden variable theory does not apply (and hence violate Bell inequalities) are the *Bell non-local* states. Based on the entangled singlet state of two spin $1/2$ particles Clauser et al [14] proposed an experiment that could demonstrate a violation of a Bell inequality. This would show that local hidden variable theory could not account for experiments that can be explained by quantum theory. Subsequent experimental work violating Bell inequalities confirmed that there are some quantum states for which a local hidden variable theory does not apply and where quantum theory was needed to explain the results (see Brunner et al [10] for a recent review). The existence of some quantum states (such as the two qubit *Bell states* [15]) for which the Bell inequalities are not obeyed and which was confirmed experimentally is itself sufficient to show that Einstein’s hope that an underlying reality represented by a local hidden variable theory could always underpin quantum theory cannot be realised.

As indicated above, Bell inequalities are usually based on local versions of hidden variable theory, with separate hidden variable measurement probabilities for each sub-system. Such a version of HVT is clearly inspired by the EPR paradox which involves separated sub-systems, and finding quantum states that are not Bell local has been a key activity of researchers - presumably motivated to show that Einstein’s idea of underpinning quantum theory with a classical realist theory could not be accomplished. To some, an ongoing program of trying to replace quantum theory by an underlying classical theory might now be seen as an exercise in nostalgia. However, a key aim of science is to find the simplest fundamental theory that explains physical effects, so if we are to confirm that quantum theory is currently that theory, it is important to rule out all versions of hidden variable theory, and not just those which are Bell local. In *non-local* forms of hidden variable theory, the hidden variables would determine the probabilities of measurement outcomes for observables in *both* sub-systems, but unlike the local case these probabilities would no longer factorise into separate probabilities for each sub-system - even for sub-systems that are well separated. Consequently, measurement outcomes for the separate sub-systems would no longer be independent.
It is here that a class of Bell inequalities introduced by Collins et al [16] are of particular interest. Although the authors claim to be dealing with a local hidden variable theory, it is contended in the present paper that their results actually apply for non-local hidden variable theory. The essential reason for this view is the absence in the derivation of CGLMP inequalities of needing to factorise the probability for measurement outcomes into separate probabilities for the observables of each sub-system - a key requirement for locality. Hence, the CGLMP inequalities are of particular significance in that if a quantum state could be shown to violate a CGLMP inequality, it follows that hidden variable theories of a more general class could also be ruled out. It also follows that any local HVT must also not be valid, as local HVT are particular cases of non-local HVT. Thus, violation of the CGLMP inequalities could be more powerful than violation of local HVT inequalities in ruling out hidden variable theories in general. This gives the CGLMP inequalities a significance not previously recognised.

The Collins et al [16] formalism is based on a HVT in which the fundamental probability $C(j, k, l, m)$ (for which no quantum expression exists) is for the outcomes of measuring four observables (two from each sub-system), but where the observables for the same sub-system would be incompatible according to quantum theory. The validity of doing this in a classical HVT is not being challenged, but it does have consequences. It results in there being no unambiguous quantum measurement process for treating the probabilities involved in the CGLMP inequalities. However, it is contended here that in spite of there being no unambiguous quantum measurement process, comparisons between the HVT and quantum predictions are still possible - and these are sufficient to show that the Collins et al [16] HVT does not predict the same results as quantum theory.

The basis for our contention may be summarised as follows: The overall aim is of course to see whether or not a HVT can predict the same results for experiment as quantum theory. Hence, in order to compare the predictions of the Collins et al [16] version of HVT with those from quantum theory, the two sets of predictions must be applied to the outcomes for the same measurement processes. Secondly, in order to avoid an immediate conflict with the Heisenberg uncertainty principle, then irrespective of the fundamental probabilities in the HVT being allowed in classical theory, the actual measurement processes involved in determining the quantities in the CGLMP inequalities must avoid the simultaneous measurement of observables that are incompatible according to quantum theory - otherwise there would be no quantum theory expressions available to determine the relevant probabilities. The measurement processes considered by Collins et al [16] do avoid conflict with the Heisenberg uncertainty principle by just involving steps where only one observable from each sub-system at a time is being measured. In compliance with this requirement, the CGLMP inequalities involve probabilities for the outcomes of such pairs of observables with the outcomes for the other pair of observable being left unrecorded. However, there are a number of differing measurement processes that are equivalent.
in classical HVT and which yield the same probability for the outcomes just described. These differ in the *order* in which measurements on the two pairs of sub-system observables is made, and on whether or not measurements are actually made on the pair of observables whose outcomes are left unrecorded. Though these differing measurement processes yield the same final outcome probability in classical physics, the same is not the case in quantum physics. Fundamentally this is because quantum measurements change the state whereas classical measurements do not. As we will see, the quantum theory expressions differ if the pairs of observables are measured in a different order, and not measuring a pair of observables yields a different outcome probability for the other pair of observables, than if the first pair are measured and their outcomes disregarded. Hence there are a number of different quantum theory expressions that correspond to the probabilities occurring the CGLMP inequalities, a feature that has not previously been recognised. However, as each of these quantum measurement processes is equivalent to a classical measurement process from which the probabilities in the CGLMP inequalities can be obtained, then a violation of a CGLMP inequality based on *any* one of the quantum measurement processes is *sufficient* to show that the Collins et al [16] HVT does not predict the same results as quantum theory. The most convenient quantum measurement process is the one where pairs of observables whose results are to be left unrecorded, are *never measured* at all. Based on this expression, Collins et al [16] find a quantum state that displays a violation of the CGLMP inequality $I \leq 3$. So far no experiments confirming this have been reported.

An issue still remaining is that the observables involved for the CGLMP violation presented in [16] have no obvious physical interpretation. As Bose-Einstein condensates in cold atomic gases are now available based on double-well potentials supporting localised modes, cases are available where there are two localised modes per well associated with different hyperfine states. With the modes for each well defining two sub-systems, the spin components could provide two different observables for each sub-system. It would be of particular interest to see if there is a violation of the CGLMP inequalities for such a system, since if the number of bosonic atoms is large, an experimental situation for confirming general non-locality in a macroscopic system may be available.

To avoid confusion, the general approach in this paper should be clarified. There are two main issues. The first concerns the question of "counterfactuality". In hidden variable theory it is meaningful to discuss measurements or introduce theoretical quantities that violate the Heisenberg uncertainty principle - after all, HVT is a classical theory. The introduction of the probability $C(j,k,l,m)$ for joint measurements of all four observables $A_1, A_2, B_1, B_2$ leading to outcomes listed as $j,k,l,m$ is an example of such a theoretical quantity (the pair for each sub-system are incompatible) and in HVT it is regarded as being measurable. The probabilities such as $P(j,k,l,m|A_1, A_2, B_1, B_2, \lambda)$ associated with the non-deterministic version of non-local HVT is another such theoretical quantity, it being regarded as measureable when the hidden variables $\lambda$ are known. Such quantities would presumably be labelled as "counterfactual".
However, this is not to say that such quantities need to be measured since HVT also provides theoretical expressions for quantities that can be (and are) measured experimentally without violating the Heisenberg uncertainty principle, and for which quantum theory expressions for the measured quantities can also be provided. The probabilities such as \( P(A_1 = j, B_1 = l) \) and \( P(A_1 = B_1) \) are examples of quantities that can be (and are) measured, and where measurement processes can equally well be described in HVT and in quantum theory. Though there is no unique classical HVT measurement process that would determine these probabilities, there is a corresponding quantum measurement process for each, neither of which involve measurements that violate the Heisenberg uncertainty principle. Here we analyse the quantum measurement process in accord with the Copenhagen interpretation. Within this framework, predictions based on the HVT such as the CGLMP inequalities can be compared with the predictions from quantum theory regarding the inequalities, with the outcome being finally decided by experiments in which the quantities appearing in the CGLMP inequalities are actually measured. Such experiments will involve a series of measurements for the same state preparation process - quantities such as \( P(A_1 = B_1) \) and \( P(B_1 = A_2 + 1) \) are measured separately. This issue is fully discussed in this paper, since it is important to verify that the CGLMP inequalities could be tested. The second issue concerns the question of "contextuality". The original version of hidden variable theory [4] considered two well separated sub-systems, and for each choice of hidden variable introduced separate measurement probabilities for the outcomes of measurements (at the same time) of an observable for each sub-system. The overall probability for a joint measurement for both observables was then obtained in accord with classical probability theory (Kolmogorov). This has been referred to as a local (non-contextual) HVT, and the present paper follows the same criterion of the factorisability of the probability for measurement outcomes for different sub-systems for each choice of hidden variable, for designating a HVT as either local or non-local (contextual). Both local and non-local versions of HVT are fully discussed in this paper, since the issue of the CGLMP inequalities being based on a non-local HVT is important.

In Section 2 the basic features of the CGLMP formalism will be reviewed - including the relationship between the fundamental probability introduced and probabilities that appear in the CGLMP inequalities. The classical measurement processes associated with such inequalities is identified. In Section 3 the relationship between the CGLMP formalism and hidden variable theory (both non-local and local) is described, and one of key CGLMP inequalities is derived. Applying the locality constraint to this CGLMP inequality is then carried out and the method of Lagrange undetermined multipliers used to investigate whether a lower upper bound occurs for the CGLMP inequality. In Section 4 the issue of replicating the classical measurement processes associated with the CGLMP inequalities with measurement processes for which there is a quantum theory formalism is treated - including a comparison with the standard local HVT situation where the fundamental probability introduced only involves
one observable for each sub-system. The quantum theory expressions used to
determine the quantities in the CGLMP inequalities are identified, and a quan-
tum state for which an inequality violation occurs is referred to. The Section 5
summarises the results. Details are set out in Appendices 8 and 9.

In this paper the same symbols will be used for the measurement outcomes,
but classical HVT observables will generally be distinguished from quantum
observables by the absence of the operator symbol. Quantum theory probability
expressions will have a subscript $Q$.

2 The CGLMP Formalism

In this section the fundamental probabilities $C(j, k, l, m)$ for joint measurements
of two sub-system observables for each of the two sub-systems introduced by
Collins et al [16] are seen as describing measurement outcomes possible in clas-
sical physics, though not in quantum physics. They represent a deterministic
form of hidden variable theory (HVT), and the issue of whether the HVT is local
or non-local will be examined. It can be seen that all of the probabilities for joint
measurements of one sub-system observable for each of the two sub-systems $A,$
$B$ of a bipartite system introduced by Collins et al [16] are also recognisable as
standard probabilities in classical physics. All can be validly expressed in terms
of the fundamental joint measurement probabilities $C(j, k, l, m)$.

2.1 Probabilities Introduced by Collins et al

In the standard notation we would express the probability $C(j, k, l, m)$ in Collins
et al [16] that the measurement of observables $A_1, A_2, B_1, B_2$ results in outcomes
listed as $j, k, l, m$ as

$$C(j, k, l, m) \equiv P(j, k, l, m|A_1, A_2, B_1, B_2) \quad (1)$$

Here the observables for sub-system $A$ are $A_1, A_2$ and those for sub-system
$B$ are $B_1, B_2$, and all four observables have the same number $d$ of different
outcomes listed as $j, k, l, m = 0, 1, ..., d-1$. As stated in [16] this formulation is a
deterministic version of hidden variable theory - as was the original treatment by
Bell [4] (see Eq. (14) therein), where the hidden variables determine the actual
outcomes for measurements. As stated in Collins et al [16], their treatment can
also be presented in a more general non-deterministic version of HVT where the
hidden variables merely determine the probabilities for measurement outcomes
- such as presented in recent work in Refs. [8], [10]. Both approaches lead to
the same CGLMP inequalities, and the choice between them is not relevant to
the other issues raised in this paper.

The fundamental Collins et al [16] probability $C(j, k, l, m)$ is based on the
simultaneous measurement of two observables for each sub-system - which is
allowed in a classical theory such as hidden variable theory. Due to the Heisenberg uncertainty principle, such probabilities do not occur in quantum theory unless the quantum operators for each sub-system commute - so in general there is no quantum expression for $P(j, k, l, m|A_1, A_2, B_1, B_2)$. Thus in the Collins et al \[16\] approach, the fundamental probability in the classical theory which is intended to underly quantum theory does not itself have a quantum counterpart. Such an approach is perfectly valid, but it will mean that only probabilities such as $P(j, l|A_1, B_1)$ of outcomes $j, l$ for measurements of observables $A_1, B_1$ (which are for different sub-systems) could have a quantum counterpart. In the Collins et al \[16\] classical theory such probabilities are derivable from the $C(j, k, l, m)$ and can be interpreted in terms of classical measurements, as we point out in the next two paragraphs. Furthermore, there are a number of measurement processes that would be equivalent in classical physics and lead to the same probability such as $P(j, l|A_1, B_1)$. However, as we will see in Section 4 the quantum expressions for such probabilities $P(j, l|A_1, B_1)$ obtained by applying these different but equivalent measurement processes are all different, so the question arises as to which of these quantum descriptions should be used to evaluate the probabilities that appear in the CGLMP inequalities? Should all be used to test the inequalities or is any one of them enough?

The four different joint probabilities for measurement outcomes for one observable from each of the two sub-systems discussed in Collins et al \[16\] are

$$P(A_1 = j, B_1 = l) = P(j, l|A_1, B_1) = \sum_{k,m} P(j, k, l, m|A_1, A_2, B_1, B_2) = \sum_{k,m} C(j, k, l, m)$$

$$P(A_2 = k, B_2 = m) = P(k, m|A_2, B_2) = \sum_{j,l} P(j, k, l, m|A_1, A_2, B_1, B_2) = \sum_{j,l} C(j, k, l, m)$$

(2)

Here $P(A_1 = j, B_1 = l) = P(j, l|A_1, B_1)$ is the probability for outcomes $j, l$ for measurement of observables $A_1, B_1$ irrespective of the outcomes for measurement of observables $A_2, B_2$. Such joint probabilities for one observable for each sub-system are obviously allowed in the CGLMP classical hidden variable theory, given that simultaneous measurements all four observables $A_1, A_2, B_1, B_2$ are allowed. They also can be described in quantum theory, as will be seen in Section 4. One possible classical measurement process is for all specific outcomes $j, l$ for measurement of observables $A_1, B_1$ and the outcomes $k, m$ for measurements on $A_2, B_2$ to be recorded. The probability for the outcomes $j, l$ for measurements of $A_1, B_1$ \textit{irrespective} of the outcomes $k, m$ for measurements of $A_2, B_2$ is then obtained by dividing the number of results with the same $j, l$ by the total number of results. Note that as classical measurements can be made without disturbing the system, the order in which the pairs of measurements for observables $A_1, B_1$ and $A_2, B_2$ occur is irrelevant, so doing the measurements in a different order would be another measurement process that is classically
equivalent for determining \(P(A_1 = j, B_1 = l)\). Similar remarks apply to the other three joint probabilities \(P(A_1 = j, B_2 = m)\), \(P(A_2 = k, B_1 = l)\) and \(P(A_2 = k, B_2 = m)\).

However, there is further way to measure probabilities such as \(P(A_1 = j, B_1 = l)\). All specific outcomes \(j, l\) for measurement of observables \(A_1, B_1\) could be recorded but no measurements would be made on \(A_2, B_2\), so again the outcomes \(k, m\) would be unrecorded. The probability for the outcomes \(j, l\) for measurements of \(A_1, B_1\) is obtained by dividing the number of results with the same \(j, l\) by the total number of results. Exactly the same expression as in Eq. (2) would apply for this different classical measurement process as for the ones described in the previous paragraph. Similar remarks apply to the other three joint probabilities \(P(A_1 = j, B_2 = m)\), \(P(A_2 = k, B_1 = l)\) and \(P(A_2 = k, B_2 = m)\).

Collins et al also introduce probabilities for when the outcomes for observables of the two sub-systems are either the same or differ by a fixed amount. Thus in terms of both the standard notation and in terms of the \(C(j, k, l, m)\) we have for example

\[
P(A_1 = B_1) = \sum_j P(A_1 = j, B_1 = j) = \sum_j P(j, j|A_1, B_1) = \sum_{j,k,m} P(j, k, j, m|A_1, A_2, B_1, B_2) = \sum_{j,k,m} C(j, k, j, m) \quad (3)
\]

and

\[
P(B_1 = A_2 + 1) = \sum_k P(A_2 = k, B_1 = k + 1 \mod d) = \sum_{j,k,m} C(j, k, k + 1 \mod d, m) \quad (4)
\]

Here \(P(A_1 = B_1) = \sum_j P(j, j|A_1, B_1)\) is the probability that the outcomes for measurements of \(A_1\) and \(B_1\) are the same, irrespective of what the outcome \(j\) is and irrespective of what the outcomes are for measurements of \(A_2\) and \(B_2\). Such probabilities for one observable for each sub-system are allowed in classical hidden variable theory, since simultaneous measurements all four observables \(A_1, A_2, B_1, B_2\) are allowed, and then all the specific outcomes \(j, l\) for measurement of observables \(A_1, B_1\) which are the same and irrespective of the outcomes \(k, m\) for measurements on \(A_2, B_2\) can be recorded. Again, the same expressions would apply if the outcomes for the observables \(A_2, B_2\) were just left unmeasured or if the order in which the pairs \(A_1, B_1\) and \(A_2, B_2\) were measured was reversed. Similar considerations apply to \(P(B_1 = A_2 + 1)\) except here the
outcomes for measurements of $A_2$ and $B_1$ are $k$ and $k + 1 \pmod{d}$, irrespective of what the outcome $k$ is.

As we will see, the CGLMP inequalities are based on the four HVT probabilities $P(A_1 = j, B_1 = l)$, ..., $P(A_2 = k, B_2 = m)$ - or to be more specific $P(A_1 = B_1)$, $P(B_1 = A_2 + 1)$, $P(A_2 = B_2)$ and $P(B_2 = A_1)$. The key point is that the classical measurement process envisaged by Collins et al [16] on which the CGLMP inequalities are based could involve measuring the outcomes for all four observables $A_1, A_2, B_1, B_2$ and then combining the results for which two outcomes (such as $j, l$ for observables $A_1, B_1$) are present irrespective of the outcomes (such as $k, m$ for $A_2, B_2$) for the other two observables, to determine probabilities such as $P(A_1 = j, B_1 = l) = P(j, l \mid A_1, B_1)$. The CGLMP inequalities could also be based on measurements of compatible observables (such as the pairs $A_1, B_1$ and $A_2, B_2$) and in either order. Alternatively, a different classical measurement process would be one in which the outcomes for two observables (such as $k, m$ for $A_2, B_2$) are never measured at all. In classical physics the probability for outcome $j, l$ for measuring a pair of observables $A_1, B_1$ and not recording the outcome $k, m$ after also measuring the other pair of observables $A_2, B_2$ would be the same as when measuring the other pair of observables never occurred. Similar remarks apply for determining the probabilities such as $P(A_1 = B_1)$, where only results for all the same outcomes (such as $j = l$ for observables $A_1, B_1$) are combined.

3 Local Hidden Variable Theory and CGLMP Inequalities

In this section the issue of whether the Collins et al [16] probabilities are consistent with hidden variable theory is considered, and if so whether that hidden variable theory is a local one. The paper by Collins et al [16] clearly states that their formalism is a local theory, but it is contended here that locality has not been invoked. Consequently, the CGLMP inequalities apply for a more general HVT, not just one restricted by the locality requirement. We begin by first reviewing the standard approach to hidden variable theory (both local and non-local) based on considering just one observable at a time for both sub-systems. The Collins et al approach involving two observables for both sub-system is then discussed. The derivation of a key CGLMP inequality then follows, with the proof not requiring the fundamental probabilities $C(j, k, l, m)$ to be determined in accord with local HVT. This is the basis of our contention that the CGLMP approach is not restricted to a local HVT. Furthermore - as shown below, adding the requirement that the $C(j, k, l, m)$ are determined in accord with local HVT does not lead to an inequality of the form $I \leq I_L$, where $I_L$ is less than 3, as might be hoped for if $C(j, k, l, m)$ is restricted to be based on local HVT. The presentation set out below is in terms of the more general non-deterministic HVT, but the deterministic version can be obtained by replacing the hidden variable probability distribution $P(\lambda)$ by a delta function.
3.1 Standard Approach: One Observable per Sub-System

In contrast to Collins et al \cite{16}, the usual discussions on Bell non-locality are framed in terms of basic probabilities for measurement outcomes of just one observable for each sub-system, which in standard notation are of the form $P(\alpha, \beta|\Omega_A, \Omega_B)$. The joint probabilities $P(A_1 = j, B_1 = l) \equiv P(j, l|A_1, B_1)$ etc (which are expressed by Collins et al \cite{16} in terms of the $C(j, k, l, m)$ as set out in the previous section) would be particular examples of the $P(\alpha, \beta|\Omega_A, \Omega_B)$. Such probabilities have a direct counterpart in quantum theory $P_Q(\alpha, \beta|\Omega_A, \Omega_B) \equiv \text{Tr} \left( \hat{\Pi}_\alpha^A \otimes \hat{\Pi}_\beta^B \right) \hat{\rho}$, and both expressions correspond to the single measurement process in which measurements on $\Omega_A, \Omega_B$ lead to outcomes $\alpha, \beta$.

In a general non-local hidden variable theory based around single measurements for each sub-system we would have

$$P(j, l|A, B) = \sum_\lambda P(\lambda) P(j, l|A, B, \lambda)$$  \hspace{1cm} (5)

where $P(\lambda)$ is the probability distribution for the hidden variables $\lambda$ and $P(j, l|A, B, \lambda)$ is the probability that measurement of sub-system observables $A, B$ leads to outcomes $j, l$ if the hidden variables are $\lambda$. For simplicity the hidden variables are assumed to be discrete - the generalisation to continuous hidden variables is trivial.

However, for the hidden variable theory to be local requires the $P(j, l|A, B, \lambda)$ to factorise into separate probabilities $P(j|A, \lambda)$ and $P(l|B, \lambda)$ for each sub-system. This criterion for locality has been set out in numerous papers (see for example, \cite{5}, \cite{9}, \cite{10}, \cite{11}, \cite{12}). Hence

$$P(j, l|A, B) = \sum_\lambda P(\lambda) P(j|A, \lambda) P(l|B, \lambda)$$  \hspace{1cm} (6)

Note that this expression shows that the measurement events where $A$ leads to outcome $j$ and $B$ leads to outcome $l$ are classically correlated.

From the local hidden variable theory expression we can then demonstrate the no signaling conditions

$$P(j|A) = \sum_l P(j, l|A, B) = \sum_\lambda P(\lambda) P(j|A, \lambda)$$

$$P(l|B) = \sum_j P(j, l|A, B) = \sum_\lambda P(\lambda) P(l|B, \lambda)$$  \hspace{1cm} (7)

Here $P(j|A) = \sum_l P(j, l|A, B)$ is the probability that measurement of $A$ leads to outcome $j$ irrespective of what the outcome is for measurement of $B$, and the result that this is given by $\sum_\lambda P(\lambda) P(j|A, \lambda)$ shows that this probability for the measurement outcome for $A$ would be the same irrespective of what observable $B$ was chosen for the other sub-system. Thus, if $B$ were to be
replaced by $B^\#$ and hence $P(l|B, \lambda)$ by $P(l^\#|B^\#, \lambda)$ we still have $P(j|A) = \sum_\lambda P(\lambda)P(j|A, \lambda)$ so the measurement outcomes for A for one sub-system do not even depend on the choice of observable B for the other sub-system. So whatever measurement is carried out for B for one sub-system has no effect on the outcome for a measurement on observable A for the other, a result that would be expected never to be violated if the two sub-systems were well-separated. Similar considerations apply for $\sum_{j,l} P(j, l|A, B)$.

### 3.2 Collins et al Approach: Two Observables per Sub-System

For the probabilities introduced by Collins et al [16] involving measurements on two observables per sub-system were described via hidden variable theory, and where the basic probability $C(j, k, l, m)$ gives the probability that measurement of observables $A_1, A_2, B_1, B_2$ results in outcomes listed as $j, k, l, m$, a general approach where hidden variables $\lambda$ are introduced we would result in the following expression for the basic probability

$$C(j, k, l, m) \equiv P(j, k, l, m|A_1, A_2, B_1, B_2) = \sum_\lambda P(\lambda)P(j, k, l, m|A_1, A_2, B_1, B_2, \lambda) \quad (8)$$

for the situation of a non-local HVT. Here $P(j, k, l, m|A_1, A_2, B_1, B_2, \lambda)$ is the probability for the outcomes $j, k, l, m$ for measurement of $A_1, A_2, B_1, B_2$ to occur when the hidden variables are $\lambda$. This corresponds to a non-deterministic hidden variable theory, and it is suggested in Collins et al [16] that the $C(j, k, l, m)$ are compatible with such an expression [17].

However, if these probabilities introduced by Collins et al [16] were described via a hidden variable theory that is local, then the simplest form of the locality condition would be to write $C(j, k, l, m)$ as the product of probabilities $A(j, k)$ and $B(l, m)$

$$C(j, k, l, m) = A(j, k) \times B(l, m) \quad (9)$$

where $A(j, k)$ gives the probability that measurement of observables $A_1, A_2$ results in outcomes listed as $j, k$, and $B(l, m)$ gives the probability that measurement of observables $B_1, B_2$ results in outcomes listed as $l, m$. However, a more general approach where hidden variables are involved would be to express the basic probability in the form

$$C(j, k, l, m) \equiv P(j, k, l, m|A_1, A_2, B_1, B_2) = \sum_\lambda P(\lambda)P(j, k|A_1, A_2, \lambda)P(l, m|B_1, B_2, \lambda) \quad (10)$$

where $P(j, k|A_1, A_2, \lambda)$ is the probability for the outcomes $j, k$ to occur for measurement on one sub-system of $A_1, A_2$ when the hidden variables are $\lambda$ and...
\(P(l, m|B_1, B_2, \lambda)\) is the probability for the outcomes \(l, m\) to occur for measurement on the other sub-system of \(B_1, B_2\) when the hidden variables are \(\lambda\). No such expression is found in the paper by Collins et al \[16\] though it is implied by their statement that they are considering a local hidden variable theory that such an expression is intended to apply for their LHVT results. We will therefore assume that if Collins et al \[16\] did involve a LHVT then this would have as its true underlying probabilities the two sub-system probabilities \(P(j, k|A_1, A_2, \lambda)\) and \(P(l, m|B_1, B_2, \lambda)\), along with the hidden variable distribution function \(P(\lambda)\). These would be the minimal basis for claiming that Collins et al \[16\] are treating a LHVT.

The derivation of the CGLMP inequalities in Ref. \[16\] is based on Eqs. (1) and (2) therein. However, whilst Eq. (2) is consistent with non-local expressions for the fundamental probability \(C(j, k, l, m)\) such as Eq. (8), there is no reason why it would follow from either of the two local expressions in Eqs. (9) or (10). The condition in Eq. (2) implies that all the measurement outcomes for \(A_1, A_2, B_1, B_2\) are non-locally correlated, even though the pairs of observables \(A_1, A_2\) and \(B_1, B_2\) apply to different sub-systems - whereas locality would require a lack of correlation between the outcomes for \(A_1, A_2\) and those for \(B_1, B_2\). The only correlation demonstrated in the standard local hidden variable theory expression for the overall joint probability \(C(j, k, l, m)\) is the classical correlation due to the hidden variables determining the separate local probabilities for the measurement outcomes for the pairs of sub-system observables \(A_1, A_2\) and \(B_1, B_2\). As we will see below, the LHVT expression leads to the no-signalling feature, where the outcomes for \(A_1, A_2\) do not depend on the choice of observables \(B_1, B_2\) for the other sub-system (and vice-versa). This feature is characteristic of localism, and it is hard to see how the non-local HVT expression would result in this feature (though it is theoretically possible \[10\]).

The basic LHVT probabilities must obey the constraints

\[
\begin{align*}
C_A &= \sum_{j,k} P(j, k|A_1, A_2, \lambda) - 1 = 0 \\
C_B &= \sum_{l,m} P(l, m|B_1, B_2, \lambda) - 1 = 0 \\
C_P &= \sum_{\lambda} P(\lambda) - 1 = 0
\end{align*}
\] (11)

Thus, the joint probabilities for one observable for each sub-system would be given by expressions such as

\[
P(A_1 = j, B_1 = l) \equiv P(j, l|A_1, B_1) = \sum_{k, m} \sum_{\lambda} P(\lambda)P(j, k|A_1, A_2, \lambda)P(l, m|B_1, B_2, \lambda) = \sum_{\lambda} P(\lambda)P(j|A_1, \lambda)P(l|B_1, \lambda)
\] (12)
where
\[ P(j|A_1, \lambda) = \sum_k P(j, k|A_1, A_2, \lambda) \quad P(l|B_1, \lambda) = \sum_m P(l, m|B_1, B_2, \lambda) \quad (13) \]

and \( P(j|A_1, \lambda) \) is the probability that measurement of sub-system observable \( A_1 \) results in outcome \( j \) if the hidden variables are \( \lambda \), with \( P(l|B_1, \lambda) \) having an analogous interpretation. Similar results apply for \( P(A_1 = j, B_2 = m) \), \( P(A_2 = k, B_1 = l) \), and \( P(k|A_2, \lambda) \) and \( P(m|B_2, \lambda) \) defined similarly to (13). From (11) we find the constraints
\[ \sum_j P(j|A_1, \lambda) = 1 \quad \sum_l P(l|B_1, \lambda) = 1 \quad \text{all } \lambda \quad (14) \]

3.3 No Signalling Result

These results are sufficient to demonstrate that the no signalling conditions apply. Thus for the cases where two observables are measured in one of the sub-systems
\[ \sum_{j,k} P(j, k|A_1, A_2, B_1, B_2) = \sum_{\lambda} P(\lambda)P(j, k|A_1, A_2, \lambda) \]
\[ \sum_{j,k} P(j, k|A_1, A_2, B_1, B_2) = \sum_{\lambda} P(\lambda)P(l, m|B_1, B_2, \lambda) \quad (15) \]

showing that the outcome for measurements in one of the sub-systems are not affected if the measurement outcomes in the other sub-system are not recorded, and more importantly that it would not matter if the other sub-system observables were replaced by a different pair, such as when \( B_1, B_2 \to B_1^#, B_2^# \) with outcomes \( l^#, m^# \). Thus the no signalling condition follows from local hidden variable theory based on the Collins et al [16] basic probabilities \( P(j, k|A_1, A_2, \lambda)P(l, m|B_1, B_2, \lambda) \).

However, if only a general hidden variable theory applied based only on the \( C(j, k, l, m) \) we would not necessarily obtain the no-signalling condition. For example
\[ \sum_{l,m} P(j, k, l, m|A_1, A_2, B_1, B_2) = \sum_{l,m} C(j, k, l, m) \]
\[ = \sum_{\lambda} P(\lambda) \sum_{l,m} P(j, k, l, m|A_1, A_2, B_1, B_2, \lambda) \]
\[ (16) \]

and the right side could still depend on the choice of \( B_1, B_2 \) even though all their outcomes are summed over. Thus \( \sum_{l,m} P(j, k, l, m|A_1, A_2, B_1, B_2, \lambda) \) might be different to \( \sum_{l^#, m^#} P(j, k, l^#, m^#|A_1, A_2, B_1^#, B_2^#, \lambda) \), where the observables \( B_1, B_2 \) have been replaced by \( B_1^#, B_2^# \) with outcomes \( l^#, m^# \).
Thus the Collins et al \cite{16} basic probabilities $C(j, k, l, m)$ only definitely results in the no signalling condition being satisfied if the hidden variable theory is local. This is consistent with statements in Brunner et al \cite{10} that locality implies no signaling, but no signaling does not imply locality.

3.4 The Basic Collins et al Inequality

In this section we will derive the Collins et al \cite{16} inequality $I \leq 3$.

The quantity $I$ is defined by

$$I = P(A_1 = B_1) + P(B_1 = A_2 + 1) + P(A_2 = B_2) + P(B_2 = A_1) \quad (17)$$

so without invoking a local hidden variable theory we have

$$I = \sum_{j,k,m} C(j, k, j, m) + \sum_{j,k,m} C(j, k, k + 1 (\text{mod } d), m) + \sum_{j,k,l} C(j, k, l, k) + \sum_{j,k,l} C(j, k, l, j)$$

$$= \sum_{j,k,l,m} C(j, k, l, m) \left[ \delta_{l,j} + \delta_{l,k+1 (\text{mod } d)} + \delta_{m,k} + \delta_{m,j} \right] \quad (18)$$

Now the quantity in the brackets $[]$ is never negative and could only have possible values of $0, 1, 2, 3, 4$ in view of the Kronecker delta only having values of $0$ or $1$. The value $4$ is impossible since this would require $j = k = m = l$ for the first, third and fourth Kronecker $\delta$ to equal $1$ but requires $l = k + 1 (\text{mod } d)$ for the second Kronecker $\delta$ to also equal $1$. Hence $[] \leq 3$ and thus

$$I \leq \sum_{j,k,l,m} C(j, k, l, m) [3] \leq 3$$

Although Collins et al \cite{16} state that this applies for local hidden variable theory, no use of any local hidden variable theory constraint seems to have been made. This inequality is valid irrespective of whether $C(j, k, j, m)$ is given by the general hidden variable theory expression \cite{18} or by the local hidden variable theory expression \cite{10}. It is also valid when the hidden variable probability $P(\lambda)$ is replaced by a delta function, which turns the non-deterministic HVT into a deterministic HVT. Collins et al \cite{16} state that the inequality $I \leq 4$ applies to general (non-local) hidden variable theory. This is of course trivially true, since $I \leq 3$ already applies to general (non-local) hidden variable theory. However, if the LHVT expression is applied, it may be possible to show that $I$ is less than a value smaller than $3$. We consider this possibility by adding the LHVT requirement to the Collins et al formalism. Finally, whether the CGLMP inequalities turn out to be useful in finding quantum states for which measurement outcomes cannot be interpreted via HVT (local or otherwise) depends on identifying measurement processes that replicates those in the Collins et al \cite{16} approach, but whose outcomes can be uniquely treated using quantum theory. This will be considered in section 4.
3.5 Local HVT Constraints and CGLMP Inequalities

Suppose that local hidden variable theory requirements are introduced. The question is - What is the maximum value for $I$ when the $C(j, k, l, m)$ are given by the LHVT expression (10) in terms of all the possible choices of $P(j, k|A_1, A_2, \lambda)$ and $P(l, m|B_1, B_2, \lambda)$? We would then have

$$I = \sum_{j, k, l, m} C(j, k, l, m) \left[ \delta_{l,j} + \delta_{l,k+1(\text{mod } d)} + \delta_{m,k} + \delta_{m,j} \right]$$

which only involves the one observable probabilities $P(j|A_1, \lambda), P(j|A_2, \lambda), P(j|B_1, \lambda)$ and $P(j|B_2, \lambda)$ introduced in Eq. (13). We note that this is a bilinear function of these probabilities.

Suppose we consider the basic quantities in the LHVT $P(j, k|A_1, A_2, \lambda), P(l, m|B_1, B_2, \lambda)$ and $P(\lambda)$ as variational functions. These will be subject to the constraints in Eq. (11). Writing

$$I = \sum_{\lambda} P(\lambda) \sum_j P(j|A_1, \lambda)P(j|B_1, \lambda)$$

$$+ \sum_{\lambda} P(\lambda) \sum_j P(j|A_2, \lambda)P(j+1(\text{mod } d)|B_1, \lambda)$$

$$+ \sum_{\lambda} P(\lambda) \sum_j P(j|A_2, \lambda)P(j|B_2, \lambda)$$

$$+ \sum_{\lambda} P(\lambda) \sum_j P(j|A_1, \lambda)P(j|B_2, \lambda)$$

(20)

Using these variational functions we can then find any local maximum (or minimum) value for $I$ as determined from (18) by using the method of Lagrange
undetermined multipliers. With \( \delta I = \sum_{\lambda} \delta P(\lambda) F(\lambda) + \sum_{\lambda} P(\lambda) \delta F(\lambda) \) we have

\[
\delta I - \mu_p \delta C_P - \mu_A \delta C_A - \mu_B \delta C_B = \sum_{\lambda} \delta P(\lambda)(F(\lambda) - \mu_p)
\]

\[
+ \sum_{\lambda} P(\lambda) \left\{ \sum_{j,k,l,m} \delta P(j,k|A_1,A_2,\lambda) P(l,m|B_1,B_2,\lambda) \Delta (j,k,l,m) \right\}
\]

\[
+ \sum_{\lambda} P(\lambda) \left\{ \sum_{j,k,l,m} P(j,k|A_1,A_2,\lambda) \delta P(l,m|B_1,B_2,\lambda) \Delta (j,k,l,m) \right\}
\]

\[
- \sum_{\lambda} P(\lambda) \left\{ \mu_A \sum_{j,k} \delta P(j,k|A_1,A_2,\lambda) + \mu_B \sum_{l,m} \delta P(l,m|B_1,B_2,\lambda) \right\}
\]

\[
= 0 \quad (22)
\]

Collecting terms gives

\[
\delta I - \mu_p \delta C_P - \mu_A \delta C_A - \mu_B \delta C_B = \sum_{\lambda} \delta P(\lambda)(F(\lambda) - \mu_p)
\]

\[
+ \sum_{\lambda} P(\lambda) \sum_{j,k} \delta P(j,k|A_1,A_2,\lambda) \left( \sum_{l,m} P(l,m|B_1,B_2,\lambda) \Delta (j,k,l,m) - \mu_A \right)
\]

\[
+ \sum_{\lambda} P(\lambda) \sum_{l,m} \delta P(l,m|B_1,B_2,\lambda) \left( \sum_{j,k} P(j,k|A_1,A_2,\lambda) \Delta (j,k,l,m) - \mu_B \right)
\]

\[
= 0 \quad (23)
\]

so that the maxima (or minima) are determined from

\[
F(\lambda) - \mu_p = 0 \quad \text{all } \lambda
\]

\[
\sum_{l,m} P(l,m|B_1,B_2,\lambda) \Delta (j,k,l,m) - \mu_A = 0 \quad \text{all } j,k, \lambda
\]

\[
\sum_{j,k} P(j,k|A_1,A_2,\lambda) \Delta (j,k,l,m) - \mu_B = 0 \quad \text{all } l,m, \lambda \quad (24)
\]

After some algebra and introducing the one observable measurement probabilities from (13) we have, after changing some of the indices and rearranging terms the following equations. Details are set out in Appendix 9.
to show that for LHVT we always have $I^3$ we must conclude that only the inequality $I$ applies to hidden variable theories, both non-local and local. However, finding a quantum state in which

$$I$$

is a saddle point. Hence no useful result arises from this attempt to find a maximum value for $LHVT$ expression (10) for $C$. But also that Bell non-locality occurs.

Consideration of the second order changes to $\mu_I$ to be a maximum (or minimum) only depend on their sums in the form of the one observable probabilities $P(j|A_1, \lambda)$, $P(j|A_2, \lambda)$, $P(j|B_1, \lambda)$ and $P(j|B_2, \lambda)$. This is to be expected in view of (20).

The last two equations in (25) can be separated into separate equations for $P(j|A_1, \lambda)$, $P(j|A_2, \lambda)$, $P(j|B_1, \lambda)$ and $P(j|B_2, \lambda)$. These are given by

$$P(j|B_1, \lambda) + P(k + 1 (mod d)|B_1, \lambda) = 0 \quad \text{all } j, k, \lambda$$

$$P(k|B_2, \lambda) + P(j - 1 (mod d)|B_2, \lambda) = 0 \quad \text{all } j, k, \lambda$$

$$P(j|A_1, \lambda) + P(k + 1 (mod d)|A_1, \lambda) = 0 \quad \text{all } j, k, \lambda$$

$$P(k|A_2, \lambda) + P(j - 1 (mod d)|A_2, \lambda) = 0 \quad \text{all } j, k, \lambda$$

so the equations for $P(k|B_1, \lambda)$ and $P(j|A_1, \lambda)$ are of the same form, as are those for $P(k|B_2, \lambda)$ and $P(j|A_2, \lambda)$. Details are given in Appendix 9.

These equations have been solved resulting in the following outcomes: All the single observable probabilities are equal, $P(j|B_1, \lambda) = P(j|A_1, \lambda) = P(j|B_2, \lambda) = P(j|A_2, \lambda) = 1/d$, with the resulting value for $I_{\text{max,min}} = 4/d$. Hence there is a local maximum or minimum for $I$ which never exceeds 2 and decreases as the number of measurement outcomes $d$ for each of the four observables increases. Consideration of the second order changes to $I$ around this value shows that in these cases $I_{\text{max,min}} = 4/d$ is neither a maximum or a minimum, but is in fact a saddle point. Hence no useful result arises from this attempt to find a maximum value for $I$ that applies only for LHVT by explicitly exploiting the LHVT expression (10) for $C_{j,k,l,m}$. The details are set out in Appendix 9.

Unless there is some other way of exploiting the LHVT expression for $C_{j,k,l,m}$ to show that for LHVT we always have $I$ smaller than a number that is less than 3, we must conclude that only the inequality $I \leq 3$ applies to hidden variable theories, both non-local and local. However, finding a quantum state in which $I_Q > 3$ would show not only that non-local hidden variable theory is not valid, but also that Bell non-locality occurs.
4 Possible Quantum Theory Measurement Processes in CLGMP Formalism

We now examine three possible quantum measurement processes that replicate classical measurement processes in the Collins et al [16] approach for the specific case of $P(A_1 = j, B_1 = l)$. These differ by the order in which measurements on the recorded observables $A_1, B_1$ and the unrecorded observables $A_2, B_2$ occur, and on whether the unrecorded observables $A_2, B_2$ are measured at all. Treatment of the other probabilities in Eq.(2) would be similar. Following the Copenhagen interpretation of quantum theory, in each case the density operator changes during the process via quantum projector operators that correspond to the quantum measurement that has taken place. The new density operator must of course still satisfy the condition $\text{Tr} \hat{\rho} = 1$. The density operator also changes when outcomes are left unrecorded. On the other hand, it does not change if no measurement is made. It is found that for each of these three equivalent classical measurement processes there is a measurement process in accord with quantum theory that replicates the measurement process that underpins the CGLMP inequalities. However, different quantum theory expressions apply in each case. Although the three measurement processes are different, it is concluded that for showing that the Collins et al HVT [16] does not predict the same results as quantum theory, it is sufficient to demonstrate a CGLMP violation for any one of the three (or more) quantum expressions that could be considered. As we will see, the measurement process in which the pair of observables with unrecorded outcomes are not measured at all leads to quantum expressions for the probabilities in the CGLMP inequalities that enabled Collins et al [16] to identify a quantum state that violates an inequality.

This ambiguity regarding the quantum measurement process is fundamentally due to the approach of basing these inequalities on a HVT probability $C(j, k, l, m)$ for the outcome of a measurement that is disallowed in quantum theory. In contrast, there is no such issue involved for Bell inequalities based on HVT of the standard type, where measurements of one observable for both sub-systems are allowed in quantum theory. To make this clear we first consider the standard type.

4.1 Quantum Theory Measurements: One Observable per Sub-System

In the case where the fundamental HVT probability $P(j, l|A, B, \lambda)$ involves measurements for a single observable for each sub-system there is no ambiguity in relating the classical HVT measurement for $P(j, l|A, B)$ to the quantum measurement process. In the classical HVT the order of measuring $A, B$ is irrelevant. In the quantum case if $\hat{A}$ is measured first with outcome $j$, then the probability of this outcome is given by $Tr(\hat{\Pi}_j^A \otimes \hat{1}^B)\hat{\rho}$ and the quantum state changes to
\( \hat{\rho}^\# \), where (see Sect. 8.3.1 in Ref. [6])

\[
\hat{\rho}^\# = (\hat{\Pi}_j^A \otimes \hat{\Pi}_l^B)\hat{\rho}(\hat{\Pi}_j^A \otimes \hat{\Pi}_l^B)/Tr(\hat{\Pi}_j^A \otimes \hat{\Pi}_l^B)\hat{\rho}. \tag{27}
\]

If \( \hat{B} \) is measured second with outcome \( l \), then the conditional probability for this outcome given the previous outcome \( j \) for measuring \( \hat{A} \) would be

\[
Tr(\hat{\Pi}_j^A \otimes \hat{\Pi}_l^B)\hat{\rho}^\# = Tr((\hat{\Pi}_j^A \otimes \hat{\Pi}_l^B)\hat{\rho}(\hat{\Pi}_j^A \otimes \hat{\Pi}_l^B)/Tr(\hat{\Pi}_j^A \otimes \hat{\Pi}_l^B)\hat{\rho})
= Tr((\hat{\Pi}_j^A \otimes \hat{\Pi}_l^B)\hat{\rho})/Tr(\hat{\Pi}_j^A \otimes \hat{\Pi}_l^B)\hat{\rho} \tag{28}
\]

The overall probability for the measurement of \( \hat{A} \) with outcome \( j \) and \( \hat{B} \) with outcome \( l \) is obtained by multiplying the conditional probability with the probability for first measuring \( \hat{A} \) with outcome \( j \), and equals the usual quantum expression for the probability of a joint measurement

\[
P_Q(j, l|A, B) = Tr(\hat{\Pi}_j^A \otimes \hat{\Pi}_l^B)\hat{\rho} \tag{29}
\]

After the second measurement the new quantum state will be \( \hat{\rho}^{##} \), where

\[
\hat{\rho}^{##} = (\hat{\Pi}_j^A \otimes \hat{\Pi}_l^B)\hat{\rho}(\hat{\Pi}_j^A \otimes \hat{\Pi}_l^B)/Tr(\hat{\Pi}_j^A \otimes \hat{\Pi}_l^B)\hat{\rho}
= (\hat{\Pi}_j^A \otimes \hat{\Pi}_l^B)\hat{\rho}(\hat{\Pi}_j^A \otimes \hat{\Pi}_l^B)/Tr(\hat{\Pi}_j^A \otimes \hat{\Pi}_l^B)\hat{\rho} \tag{30}
\]

after some operator algebra and using \( \hat{\Pi}^2 = \hat{\Pi} \). The expressions for the joint measurement probability \( P(j, l|A, B) \) and the final quantum state \( \hat{\rho}^{##} \) are those expected from quantum theory. A key point is that the same results are obtained if \( \hat{B} \) with outcome \( l \) is measured first and \( \hat{A} \) with outcome \( j \) is measured second. Hence a HVT expression for \( P(j, l|A, B) \) based on classical measurements of \( A, B \) taken in either order is linked to a unique quantum theory expression describing the same measurement process.

Since the Bell inequalities involve joint probabilities such as \( P(j, l|A, B) \) or joint mean values derived from these such as

\[
\langle A \otimes B \rangle = \sum_{j,l} (j \times l) P(j, l|A, B) \tag{31}
\]

then the HVT expression based on \( \hat{\rho}^{##} \) and the quantum theory expression based on \( (29) \) can be compared in terms of the same measurement process. The mean value expressions are

\[
\langle A \otimes B \rangle_{HVT} = \sum_{\lambda} P(\lambda) \sum_{j,l} (j \times l) P(j, l|A, B, \lambda) = \sum_{\lambda} P(\lambda) \langle (A \otimes B)\lambda \rangle_{HVT}
\]

\[
\langle A \otimes B \rangle_Q = Tr \sum_{j,l} (j \times l) (\hat{\Pi}_j^A \otimes \hat{\Pi}_l^B)\hat{\rho} = Tr(\hat{A} \otimes \hat{B})\hat{\rho} \tag{32}
\]

im an obvious notation.
4.2 Quantum Theory Measurements: Two Observables per Sub-System

We now examine three possible quantum measurement processes that attempt to replicate classical measurement processes in the Collins et al [16] approach for the specific case of $P(A_1 = j, B_1 = l)$.

The first possibility would be if the measurements on $A_2, B_2$ which led to outcomes $k, m$ were performed first. In this case the original density operator $\hat{\rho}$ changes to $\hat{\rho}^\#$ where (see Sect 8.3.1 in Ref [16])

$$\hat{\rho}^\# = \left(\hat{\Pi}^{A_2}_k \otimes \hat{\Pi}^{B_2}_m\right) \hat{\rho} \left(\hat{\Pi}^{A_2}_k \otimes \hat{\Pi}^{B_2}_m\right) / Tr \left(\hat{\Pi}^{A_2}_k \otimes \hat{\Pi}^{B_2}_m\right) \hat{\rho} \quad (33)$$

Here $\hat{\Pi}^{A_2}_k$ etc are the usual quantum projectors, with $\sum_k \hat{\Pi}^{A_2}_k = \hat{1}_A$, etc. If the results for all the outcomes $k, m$ are then left unrecorded the density operator changes again to $\hat{\rho}^\##$ where

$$\hat{\rho}^\## = \sum_{k,m} \hat{\rho}^\# \times Tr \left(\hat{\Pi}^{A_2}_k \otimes \hat{\Pi}^{B_2}_m\right) \hat{\rho} = \sum_{k,m} \left(\hat{\Pi}^{A_2}_k \otimes \hat{\Pi}^{B_2}_m\right) \hat{\rho} \left(\hat{\Pi}^{A_2}_k \otimes \hat{\Pi}^{B_2}_m\right) \hat{\rho} \quad (34)$$

This of course differs from the original density operator $\hat{\rho}$. So it is not as if $A_2, B_2$ had never been measured at all. The quantum probability for measurements on $A_2, B_2$ which led to all outcomes $k, m$ is obviously $P_Q(l, m|A_2, B_2) = \sum_{k,m} Tr \left(\hat{\Pi}^{A_2}_k \otimes \hat{\Pi}^{B_2}_m\right) \hat{\rho} = 1$, where the notation $(k, m)$ indicates outcome events for all $k, m$.

The probability that subsequent measurement of observables $A_1, B_1$ resulting in outcomes $j, l$ will be given by the conditional probability

$$Tr \left(\hat{\Pi}^{A_1}_j \otimes \hat{\Pi}^{B_1}_l\right) \hat{\rho}^\## = \sum_{k,m} Tr \left(\hat{\Pi}^{A_1}_j \otimes \hat{\Pi}^{B_1}_l\right) \left(\hat{\Pi}^{A_2}_k \otimes \hat{\Pi}^{B_2}_m\right) \hat{\rho} \left(\hat{\Pi}^{A_2}_k \otimes \hat{\Pi}^{B_2}_m\right) \hat{\rho} \quad (35)$$

so that the overall quantum probability for the event where measurement of $A_1, B_1$ results in outcomes $j, l$ after measurement of $A_2, B_2$ results in all outcomes $k, m$ is obtained by multiplying this conditional probability by $P_Q(l, m|A_2, B_2) = 1$, and is given by

$$P_Q(j, l, (k, m)|A_1, B_1, (A_2, B_2)) = \sum_{k,m} Tr \left(\hat{\Pi}^{A_2}_k \otimes \hat{\Pi}^{B_2}_m\right) \hat{\rho} \left(\hat{\Pi}^{A_2}_k \otimes \hat{\Pi}^{B_2}_m\right) \hat{\rho} \quad (36)$$

The notation $P_Q(j, l, (k, m)|A_1, B_1, (A_2, B_2))$ indicates that the $A_2, B_2$ measurements were carried out first and the results left unrecorded. The subsequent measurement of observables $A_1, B_1$ leading to outcomes $j, l$ results in the further
change of the density operator from $\hat{\rho}^{\#\#}$ to $\hat{\rho}^{\#\#\#}$ where

$$
\hat{\rho}^{\#\#\#} = \sum_{k,m} \left( \hat{\Pi}_j^A \hat{\Pi}_k^A \otimes \hat{\Pi}_l^B \hat{\Pi}_m^B \right) \hat{\rho} \left( \hat{\Pi}_k^A \hat{\Pi}_j^A \otimes \hat{\Pi}_l^B \hat{\Pi}_m^B \right) /
$$

$$
\left\{ \sum_{k,m} Tr \left( \hat{\Pi}_j^A \hat{\Pi}_k^A \hat{\Pi}_l^B \hat{\Pi}_m^B / \hat{\Pi}_l^B \hat{\Pi}_m^B \right) \right\}
$$

(37)

We next consider a second possibility for the quantum measurements, namely what happens in quantum theory if observables $A_1, B_1$ resulting in outcomes $j, l$ are measured first, followed by measurement of $A_2, B_2$ leading to outcomes $k, m$, which are then left unrecorded. The overall probability for this process is

$$
P_Q(j, l, (k, m)|A_1, B_1, (A_2, B_2)_2) = \sum_{k,m} Tr \left( \hat{\Pi}_j^A \hat{\Pi}_k^A \hat{\Pi}_l^B \hat{\Pi}_m^B \hat{\Pi}_l^B \hat{\Pi}_m^B \right) \hat{\rho}
$$

(38)

and the final density operator is

$$
\hat{\rho}^{k,k;kk} = \sum_{k,m} \left( \hat{\Pi}_j^A \hat{\Pi}_k^A \otimes \hat{\Pi}_l^B \hat{\Pi}_m^B \right) \hat{\rho} \left( \hat{\Pi}_j^A \hat{\Pi}_k^A \otimes \hat{\Pi}_l^B \hat{\Pi}_m^B / \hat{\Pi}_l^B \hat{\Pi}_m^B \right)
$$

(39)

The notation $P_Q(j, l, (k, m)|A_1, B_1, (A_2, B_2)_2)$ indicates that the $A_2, B_2$ measurements were carried out second and the results left unrecorded. The details are set out in Appendix S.

In the general case where the two operators of each sub-system do not commute, the results for $P_Q(j, l, (k, m)|A_1, B_1, (A_2, B_2)_2)$ and $\hat{\rho}^{\#\#\#}$ or $P_Q(j, l, (k, m)|A_1, B_1, (A_2, B_2)_2)$ and $\hat{\rho}^{k,k;kk}$ for these two measurement processes are not the same. There are even further possibilities that could have been considered, such as involving measuring observables $A_1, B_2$ resulting in outcomes $j, m$ are measured first with the outcome $m$ left unrecorded, followed by measurement of $A_2, B_1$ leading to outcomes $k, l$ with the outcome $k$ then left unrecorded. This confirms that the two classically equivalent measurement processes that equally determine the probability $P(A_1 = j, B_1 = l)$ can each be described via quantum theory, but the two quantum theory predictions are different.

We now consider a third possibility. In the general case where the two operators of each sub-system do not commute, the results for $P_Q(j, l, (k, m)|A_1, B_1, (A_2, B_2)_2)$ and $\hat{\rho}^{\#\#\#}$ or $P_Q(j, l, (k, m)|A_1, B_1, (A_2, B_2)_2)$ and $\hat{\rho}^{k,k;kk}$ are not the same as if measurements on $A_2, B_2$ had never taken place at all. The probability $P_Q(j, l|A_1, B_1)$ for measurement of observables $A_1, B_1$ alone that results in outcomes $j, l$ would be

$$
P_Q(j, l|A_1, B_1) = Tr \left( \hat{\Pi}_j^A \otimes \hat{\Pi}_l^B \right) \hat{\rho}
$$

(40)

and quantum density operator following this measurement just of observables $A_1, B_1$ alone would be

$$
\hat{\rho}_{j,l} = \left( \hat{\Pi}_j^A \otimes \hat{\Pi}_l^B \right) \hat{\rho} \left( \hat{\Pi}_j^A \otimes \hat{\Pi}_l^B \right) / Tr \left( \hat{\Pi}_l^B \right) \hat{\rho}
$$

(41)
So not only does the measurement probability $P_Q(j, l | A_1, B_1)$ differ from $P_Q(j, l, (k, m) | A_1, B_1, (A_2, B_2)_1)$ or $P_Q(j, l, (k, m) | A_1, B_1, (A_2, B_2)_2)$, but the final quantum states $\hat{\rho}_{j,l}$ differs from $\hat{\rho}^{###}$ or $\hat{\rho}^{kkk}$ are also different. Thus the classical probability $P(j, l | A_1, B_1) = \sum_{k,m} P(j, k, l, m | A_1, A_2, B_1, B_2)$ is linked to a third quantum probability for a possible measurement process that replicates one of the equivalent classical measurement processes on which $P(A_1 = j, B_1 = l)$ is based.

Similar considerations apply to the other probabilities $P(A_1 = j, B_2 = m)$, $P(A_2 = k, B_1 = l)$ and $P(A_2 = k, B_2 = m)$.

The above analysis confirms the situation that expressions in the classical hidden variable theory of Collins et al [16] for probabilities that occur in the CGLMP inequities are replicated by a number of different quantum theory expressions depending on which of the classically equivalent measurement processes associated with the HVT probability is considered. In classical measurement theory for the particular case of $P(A_1 = j, B_1 = l) = P(j, l | A_1, B_1)$ it should not make any difference if $A_1, B_1$ were measured first resulting in outcomes $j, l$ followed by measurements of $A_2, B_2$ resulting in all outcomes $k, m$ from that when the measurements of $A_2, B_2$ were carried out first and the outcomes $k, m$ left unrecorded, followed by measurements of $A_1, B_1$ leading to outcomes $j, l$. However, the quantum measurement theory treatment of the two different sequences give different results, both in terms of the overall quantum probabilities for the process and the final quantum state that is created. Similarly, if the unrecorded observables $A_2, B_2$ are never measured at all, a third quantum expression is involved. So which one is to be chosen to give the quantum theory analogue of the Collins et al [16] quantity $P(A_1 = j, B_1 = l)$?

Fortunately, since the overall aim is to demonstrate that the Collins et al HVT does not predict the same results as quantum theory, we may choose any one of the quantum expressions provided that a quantum state can be found for which a CGLMP inequality is violated. After all, the three different measurement processes described above are all treatable via quantum theory, and all three are equivalent classically for replicating the HVT quantities such as $P(A_1 = j, B_1 = l)$. It is sufficient to show that a CGLMP inequality is violated for one measurement process and for one quantum state to demonstrate that the Collins et al HVT [16] cannot predict the same results as quantum theory. From the point of view of simplicity in the quantum calculations, the most suitable measurement process to choose would be the one where unrecorded observables are never measured at all.

Collins et al [16] make the comparison of the CGLMP hidden variable theory predictions with those from quantum theory by choosing the measurement processes to be those where the unrecorded pair of observables are just not measured at all - the third (and simplest) possibility discussed above. It can easily be confirmed from the quantum theory probability expression set out in Eq. (14) therein for $P_{QM}(A_a = k, B_b = l)$ that this is the approach that has been adopted. Hence Collins et al [16] use the following quantum theory expression
for $P_Q(A_1 = j, B_1 = l)$

$$P(A_1 = j, B_1 = l) = \sum_{k,m} C(j,k,l,m) \equiv Tr \left( \hat{\Pi}_j^{A_1} \otimes \hat{\Pi}_l^{B_1} \right) \hat{\rho} = P_Q(j,l|A_1,B_1)$$  \hspace{1cm} (42)

For the final probabilities $P(A_1 = B_1)$, $P(B_1 = A_2 + 1)$ etc that appear in the CGLMP inequalities, similar considerations apply and expressions such as

$$P(A_1 = B_1) = \sum_{j,k,m} C(j,k,j,m) \equiv Tr \sum_j \left( \hat{\Pi}_j^{A_1} \otimes \hat{\Pi}_j^{B_1} \right) \hat{\rho} = \sum_j P_Q(j,j|A_1,B_1)$$

$$P(B_1 = A_2 + 1) = \sum_{j,k,m} C(j,k,k+1(mod \, d),m)$$

$$\equiv Tr \sum_k \left( \hat{\Pi}_k^{A_2} \otimes \hat{\Pi}_{k+1(mod \, d)}^{B_1} \right) \hat{\rho} = \sum_k P_Q(k,k+1(mod \, d)|A_2,B_1)$$ \hspace{1cm} (43)

have been assumed in Collins et al [16].

Thus, in spite of the CGLMP Bell inequalities being based on expressions such as $P(A_1 = B_1)$, $P(B_1 = A_2 + 1)$ which have several possible equivalents in quantum theory, conclusions that certain quantum states and related observables lead to violations of Bell locality can still be made.

### 4.3 Quantum State Violating Collins Inequality

An example of a quantum state that violates the inequality $I \leq 3$ is considered by Collins et al [16], based on quantum expressions in Eq. (43). Two particles with the same spin $s$ are considered, for which the spin eigenstates are $|s,m\rangle$, where $m = -s, .., +s$. For the (unnormalized) state $\sum_{m=-s}^s |s,m\rangle_A |s,m\rangle_B$ the quantum expression for $I$ is found to be greater than 3 for all $d = 2s + 1$, corresponding to a Bell inequality violation in a macroscopic system if $s$ is large. However, this violation involved introducing physical quantities $A_1, A_2, B_1, B_2$ as Hermitian operators defined by their eigenvalues and eigenvectors (see Eq. (13) in Ref [16]), the latter being linear combinations of the $|s,m\rangle_{A(B)}$. However, as the operators turn out to be off-diagonal in these basis states, it is not obvious what physical observable they correspond to. So not only is it unclear what quantum measurement sequences should be used to confirm the violation of the inequality $I \leq 3$, but also the physical significance of the four sub-system observables themselves is obscure. No experimental tests of the Bell inequalities have been carried out. However, as pointed out in Section 1 it may now be possible to test the CGLMP inequalities in Bose-Einstein condensates in cold atomic gases based on double-well potentials, with two localised modes per well associated with different hyperfine states and with spin components as the sub-system observables.
Note that if a quantum state is shown to violate the inequality $I \leq 3$, then it follows that not only is the non-local HVT of Collins et al. shown not to be valid but also any local HVT must also not be valid. As local HVT are particular cases of non-local HVT, violation of the CGLMP inequalities is more powerful than violation of local HVT inequalities, as this would rule out a more general class of hidden variable theories.

5 Conclusion

The significance of the CGLMP (Bell) inequalities as possible tests for showing that quantum theory is not underpinned even by a general non-local hidden variable theory has been discussed. The question of whether Collins et al actually used a local form of hidden variable theory has been examined, and it is concluded that the CGLMP inequalities only require the assumption that a general hidden variable theory applies, without the necessity of it being local. Such a theory is more general than the standard local theory. The CGLMP (Bell) inequalities are based on a form of hidden variable theory (HVT) that allows for simultaneous measurements of pairs of observables that correspond to non-commuting quantum operators. However, although this is allowed in a classical probability theory, it does lead to the CGLMP Bell inequalities being based on expressions for which a number of different quantum theory expressions apply, corresponding to different measurement processes that would have been equivalent in classical physics. Fundamentally, this is because quantum measurements change the quantum state whereas classical measurements leave the state unchanged, and for observables whose outcomes are unrecorded whether measurements of these observables are made and discarded, differs in quantum theory from the case where the measurements are not made at all. However, conclusions that certain quantum states and related observables lead to violations of HVT can still be made based on any one of the possible quantum theory expressions that replicates an equivalent classical measurement processes that could determine the probabilities in the CGLMP inequalities. The most convenient quantum measurement process is the one where pairs of observables whose results are to be left unrecorded are never measured at all. Based on this expression Collins et al have identified a quantum state that violates the CGLMP inequality $I \leq 3$. However, the observables found by Collins et al to be associated with the CGLMP inequality $I \leq 3$ violation have no obvious physical interpretation. We also considered what further inequality may apply if the CGLMP hidden variable theory is then restricted to being local, via maximizing $I$ using the method of Lagrange undetermined multipliers. However the result of this procedure only yielded a stationary value $I = 4/d$ for the CGLMP inequality, which was a saddle point rather than a maximum, so at present $I \leq 3$ applies for both local and non-local HVT states. Leaving aside the issue of interpreting the observables, it is concluded that the CGLMP inequalities have been shown to rule out hidden variable theories, both non-local and local. The CGLMP inequalities therefore rule out a wider class of hidden variable theories.
than those that are local, giving the CGLMP an additional importance not previously recognised. We also point out that CGLMP tests might be carried out in systems such as Bose-Einstein condensates of atomic gases with two hyperfine components in a double potential well, as the spin components for each well could be suitable observables.

6 Acknowledgements

The author is grateful for helpful discussions with Professors S M Barnett and B M Garraway, and acknowledges financial support from University of Glasgow and Imperial College, London during the conduct of this research. The author is also grateful to two referees for raising issues that have resulted in a somewhat different conclusion about CGLMP inequalities.

7 Conflicts of Interest

The author declares no conflict of interest.

References

[1] A. Einstein, B. Podolsky and N. Rosen, Phys. Rev. 47, (1935) 777.
[2] E. Schrödinger, Proc. Camb. Phil. Soc. 31, (1935) 553; Proc. Camb. Phil. Soc. 32, (1936) 446.
[3] E. Schrödinger, Naturwissenschaften 23, (1935) 807.
[4] J. S. Bell, Physics 1, (1964) 195.
[5] R. F. Werner, Phys. Rev. A 40, (1989) 4277.
[6] C. J. Isham, 1995 *Lectures on Quantum Theory* (Imperial College Press: London (1995)).
[7] D. Bohm, Quantum Theory, Prentice-Hill: New York, USA, 1951.
[8] S. J. Jones, H. M. Wiseman and A. C. Doherty, Phys. Rev. A 76, (2007) 052116.
[9] M. D. Reid, P. D. Drummond, W. P. Bowen, E. G. Cavalcanti, P. K. Lam, H. A. Bachor, U. L. Andersen and G. Leuchs, Rev. Mod. Phys. 81 (2009) 1727.
[10] N. Brunner, D. Cavalcanti, S. Pironio, V. Scarani and S. Wehner, Rev. Mod. Phys. 86, (2014) 419.
[11] S. Jevtic, M. J. W. Hall, M. R. Anderson, M. Zwierz and H. M. Wiseman, J. Opt. Soc. Amer. B 32, (2015) A40.

[12] B. J. Dalton, J. Goold, B. M. Garraway and M. D. Reid, Phys. Scr. 92, (2017) 023004.

[13] J. S. Bell, Foundations of Quantum Mechanics, Ed. B. D’Espagnat, Academic, New York, USA, 1971.

[14] J. F. Clauser, M. A. Horne, A. Shimony and R. A. Holt, Phys. Rev. Letts. 23, (1969) 880.

[15] S. M. Barnett, 2009 Quantum Information (Oxford University Press: Oxford (2009)).

[16] D. Collins, N. Gisin, N. Linden, S. Massar and S. Popescu, Phys. Rev. Letts. 88, (2002) 040404.

[17] I. C. Percival, Phys. Letts. A, 244, (1998) 495.
8 Appendix - Quantum Measurement Processes

Replicating \( P(A_1, B_1) \)

Here we consider the second possibility for what happens in quantum theory in a measurement process that replicates the classical measurement process on which \( P(A_1 = j, B_1 = l) \) is based when observables \( A_1, B_1 \) resulting in outcomes \( j, l \) are measured first, followed by measurement of \( A_2, B_2 \) leading to outcomes \( l, m \), which are then left unrecorded.

After the first measurement the original density operator \( \hat{\rho} \) changes to \( \hat{\rho}^k \) where (see Sect 8.3.1 in Ref [6])

\[
\hat{\rho}^k = \left( \hat{\Pi}_j^{A_1} \otimes \hat{\Pi}_l^{B_1} \right) \hat{\rho} \left( \hat{\Pi}_j^{A_1} \otimes \hat{\Pi}_l^{B_1} \right) / \text{Tr} \left( \hat{\Pi}_j^{A_1} \otimes \hat{\Pi}_l^{B_1} \right) \hat{\rho} \tag{44}
\]

The probability for measurement of observables \( A_1, B_1 \) resulting in outcomes \( j, l \) is given by

\[
P_Q(j, l|A_1, B_1) = \text{Tr} \left( \hat{\Pi}_j^{A_1} \otimes \hat{\Pi}_l^{B_1} \right) \hat{\rho} \tag{45}
\]

After the subsequent measurement of \( A_2, B_2 \) leading to outcomes \( k, m \) the density operator becomes

\[
\hat{\rho}^{kk} = \left( \hat{\Pi}_k^{A_2} \hat{\Pi}_l^{A_1} \otimes \hat{\Pi}_m^{B_2} \hat{\Pi}_l^{B_1} \right) \hat{\rho} \left( \hat{\Pi}_k^{A_2} \hat{\Pi}_l^{A_1} \otimes \hat{\Pi}_m^{B_2} \hat{\Pi}_l^{B_1} \right) / \text{Tr} \left( \hat{\Pi}_k^{A_2} \hat{\Pi}_l^{A_1} \otimes \hat{\Pi}_m^{B_2} \hat{\Pi}_l^{B_1} \right) \hat{\rho} \tag{46}
\]

and this outcome occurs with a conditional probability \( \text{Tr} \left( \hat{\Pi}_k^{A_2} \otimes \hat{\Pi}_m^{B_2} \right) \hat{\rho}^{kk} \).

The conditional probability for all outcomes \( k, m \) for measurements of \( A_2, B_2 \) following measurements of \( A_1, B_1 \) resulting in outcomes \( j, l \) will then be

\[
\sum_{k,m} \text{Tr} \left( \hat{\Pi}_k^{A_2} \otimes \hat{\Pi}_m^{B_2} \right) \hat{\rho}^{kk} = \sum_{k,m} \text{Tr} \left( \hat{\Pi}_k^{A_2} \otimes \hat{\Pi}_m^{B_2} \right) \left( \hat{\Pi}_j^{A_1} \otimes \hat{\Pi}_l^{B_1} \right) \hat{\rho} \left( \hat{\Pi}_j^{A_1} \otimes \hat{\Pi}_l^{B_1} \right) / \text{Tr} \left( \hat{\Pi}_j^{A_1} \otimes \hat{\Pi}_l^{B_1} \right) \hat{\rho} = \sum_{k,m} \text{Tr} \left( \hat{\Pi}_j^{A_1} \hat{\Pi}_k^{A_2} \hat{\Pi}_l^{A_1} \otimes \hat{\Pi}_m^{B_2} \hat{\Pi}_l^{B_1} \right) \hat{\rho} / \text{Tr} \left( \hat{\Pi}_j^{A_1} \otimes \hat{\Pi}_l^{B_1} \right) \hat{\rho} \tag{47}
\]

The overall quantum probability for the event where measurement of \( A_1, B_1 \) results in outcomes \( j, l \) before measurement of \( A_2, B_2 \) results in all outcomes \( k, m \) is obtained by multiplying this conditional probability by \( P_Q(j, l|A_1, B_1) \) and is given by

\[
P_Q(j, l, (k, m)|A_1, B_1, (A_2, B_2)_2) = \sum_{k,m} \text{Tr} \left( \hat{\Pi}_j^{A_1} \hat{\Pi}_k^{A_2} \hat{\Pi}_l^{A_1} \otimes \hat{\Pi}_m^{B_2} \hat{\Pi}_l^{B_1} \right) \hat{\rho} \tag{48}
\]

The notation \( P_Q(j, l, (k, m)|A_1, B_1, (A_2, B_2)_2) \) indicates that the \( A_2, B_2 \) measurements were carried out second and the results left unrecorded. Note the different expressions for \( P_Q(j, l, (k, m)|A_1, B_1, (A_2, B_2)_2) \) and \( P_Q(j, l, (k, m)|A_1, B_1, (A_2, B_2)_1) \).
which are both different to $P_Q(j, l|A_1, B_1)$, the probability for just measuring $A_1, B_1$ alone. The density operator and after the $A_2, B_2$ measurements were carried out and the outcomes $l, m$ are left unrecorded changes from $\hat{\rho}^{k,k}$ to $\hat{\rho}^{k,k,k}$ where

$$\hat{\rho}^{k,k,k} = \sum_{k,m} \left( \hat{\Pi}^A_{k} \hat{\Pi}^A_{j} \otimes \hat{\Pi}^B_{m} \hat{\Pi}^B_{l} \right) \hat{\rho} \left( \hat{\Pi}^A_{j} \hat{\Pi}^A_{k} \otimes \hat{\Pi}^B_{l} \hat{\Pi}^B_{m} \right) / \left\{ \text{Tr} \left( \hat{\Pi}^A_{j} \otimes \hat{\Pi}^B_{l} \right) \hat{\rho} \right\}$$

(49)
9 Appendix - Maximizing I via Method of Lagrange Multipliers

9.1 Deriving the Eqs. (25)

Substituting for $\Delta (j, k, l, m)$ from (21) and carrying out the sums gives

$$F(\lambda) - \mu_p = 0$$

$$\sum_{j,k,l,m} P(j,k|A_1,A_2,\lambda)P(l,m|B_1,B_2,\lambda) \left( \delta_{i,j} + \delta_{i,k+1(\mod d)} + \delta_{m,k} + \delta_{m,j} \right) - \mu_p$$

$$= 0 \quad \text{all } \lambda$$  \hspace{1cm} (50)

and

$$\sum_{l,m} P(l,m|B_1,B_2,\lambda) \left( \delta_{i,j} + \delta_{i,k+1(\mod d)} + \delta_{m,k} + \delta_{m,j} \right) - \mu_A$$

$$= 0$$

$$\sum_{m} P(j,m|B_1,B_2,\lambda) + \sum_{m} P(k+1(\mod d), m|B_1,B_2,\lambda)$$

$$+ \sum_{l} P(l,k|B_1,B_2,\lambda) + \sum_{l} P(l,j|B_1,B_2,\lambda) - \mu_A$$

$$= 0 \quad \text{all } j, k, \lambda$$  \hspace{1cm} (51)

and

$$\sum_{j,k} P(j,k|A_1,A_2,\lambda) \Delta (j,k,l,m) - \mu_B$$

$$= 0$$

$$\sum_{k} P(l,k|A_1,A_2,\lambda) + \sum_{j} P(j,l-1(\mod d)|A_1,A_2,\lambda)$$

$$+ \sum_{j} P(j,m|A_1,A_2,\lambda) + \sum_{k} P(m,k|A_1,A_2,\lambda) - \mu_B$$

$$= 0 \quad \text{all } l, m, \lambda$$  \hspace{1cm} (52)
So introducing the one observable measurement probabilities from (23) we have, after changing some of the indices and rearranging terms

\[
\sum_j P(j|A_1, \lambda)P(j|B_1, \lambda) + \sum_j P(j|A_2, \lambda)P(j+1 \mod d|B_1, \lambda) \\
+ \sum_j P(j|A_2, \lambda)P(j|B_2, \lambda) + \sum_j P(j|A_1, \lambda)P(j|B_2, \lambda) - \mu_P = 0 \quad \text{all } \lambda \\
P(j|B_1, \lambda) + P(j+1 \mod d|B_1, \lambda) + P(k|B_2, \lambda) + P(j|B_2, \lambda) - \mu_A = 0 \quad \text{all } j, k, \lambda \\
P(l|A_1, \lambda) + P(\overline{m}+1 \mod d|A_1, \lambda) + P(m|A_2, \lambda) + P(l|A_2, \lambda) - \mu_B = 0 \quad \text{all } l, m, \lambda 
\]

We note that the conditions on the basic quantities in the LHVT \( P(j, k|A_1, A_2, \lambda), P(l, m|B_1, B_2, \lambda) \) for \( I \) to be a maximum (or minimum) only depend on their sums in the form of the one observable probabilities \( P(j|A_1, \lambda), P(j|A_2, \lambda), P(j|B_1, \lambda) \) and \( P(j|B_2, \lambda) \). This is to be expected in view of (20).

### 9.2 Obtaining the Maxima, Minima

If in the last two equations of (25) the indices are interchanged via \( j \leftrightarrow k \) and \( l \leftrightarrow m \) we have

\[
P(k|B_1, \lambda) + P(j+1 \mod d|B_1, \lambda) + P(j|B_2, \lambda) + P(k|B_2, \lambda) - \mu_A = 0 \quad \text{all } j, k, \lambda \\
P(m|A_1, \lambda) + P(\overline{m}+1 \mod d|A_1, \lambda) + P(l|A_2, \lambda) + P(l|A_2, \lambda) - \mu_B = 0 \quad \text{all } l, m, \lambda 
\]

Subtracting gives

\[
-P(k|B_1, \lambda) - P(j+1 \mod d|B_1, \lambda) = 0 \quad \text{all } j, k, \lambda \\
-P(m|A_1, \lambda) - P(\overline{m}+1 \mod d|A_1, \lambda) = 0 \quad \text{all } l, m, \lambda 
\]

which are equations just involving \( P(k|B_1, \lambda) \) and \( P(l|A_1, \lambda) \) on their own.

Similarly if in the same two equations the indices are interchanged via \( j \leftrightarrow k+1 \mod d \) and \( l \leftrightarrow m+1 \mod d \) we have

\[
P(k+1 \mod d|B_1, \lambda) + P(j|B_1, \lambda) + P(j-1 \mod d|B_2, \lambda) + P(k+1 \mod d|B_2, \lambda) - \mu_A = 0 \quad \text{all } j, k, \lambda \\
P(\overline{m}+1 \mod d|A_1, \lambda) + P(l|A_1, \lambda) + P(\overline{l}-1 \mod d|A_2, \lambda) + P(\overline{m}+1 \mod d|A_2, \lambda) - \mu_B = 0 \quad \text{all } l, m, \lambda 
\]

31
Subtracting gives

\[
-P(j - 1 \mod d|B_2, \lambda) - P(k + 1 \mod d|B_2, \lambda) = P(k|B_2, \lambda) + P(j|B_2, \lambda)
\]

\[
P(l - 1 \mod d|A_2, \lambda) - P(m + 1 \mod d|A_2, \lambda) = P(m|A_2, \lambda) + P(l|A_2, \lambda)
\]

which are equations just involving \(P(k|B_2, \lambda)\) and \(P(l|A_2, \lambda)\) on their own.

Thus we have in terms of indices \(j, k, \lambda\)

\[
P(j|B_1, \lambda) + P(k + 1 \mod d|B_1, \lambda) - P(k|B_1, \lambda) - P(j - 1 \mod d|B_1, \lambda) = 0 \quad \text{all } j, k, \lambda
\]

\[
P(k|B_2, \lambda) + P(j|B_2, \lambda) - P(j - 1 \mod d|B_2, \lambda) - P(k + 1 \mod d|B_2, \lambda) = 0 \quad \text{all } j, k, \lambda
\]

\[
P(j|A_1, \lambda) + P(k + 1 \mod d|A_1, \lambda) - P(k|A_1, \lambda) - P(j - 1 \mod d|A_1, \lambda) = 0 \quad \text{all } j, k, \lambda
\]

\[
P(k|A_2, \lambda) + P(j|A_2, \lambda) - P(j - 1 \mod d|A_2, \lambda) - P(k + 1 \mod d|A_2, \lambda) = 0 \quad \text{all } j, k, \lambda
\]

so the equations for \(P(k|B_1, \lambda)\) and \(P(j|A_1, \lambda)\) are of the same form, as are those for \(P(k|B_2, \lambda)\) and \(P(k|A_2, \lambda)\).

The solution to the equations for \(P(k|B_1, \lambda)\) requires \(P(k + 1 \mod d|B_1, \lambda) - P(k|B_1, \lambda) = \alpha_{B1}, P(j|B_1, \lambda) - P(j - 1 \mod d|B_1, \lambda) = -\alpha_{B1}\) for all \(j, k, \lambda\).

We then have for the particular \(k\)

\[
P(1|B_1, \lambda) - P(0|B_1, \lambda) = \alpha_{B1} \quad P(1|B_1, \lambda) - P(0|B_1, \lambda) = 0
\]

\[
P(2|B_1, \lambda) - P(1|B_1, \lambda) = \alpha_{B1} \quad P(2|B_1, \lambda) - P(1|B_1, \lambda) = 2\alpha_{B1}
\]

\[
\vdots
\]

\[
P(d - 1|B_1, \lambda) - P(d - 2|B_1, \lambda) = \alpha_{B1} \quad P(d - 1|B_1, \lambda) - P(d - 2|B_1, \lambda) = d\alpha_{B1}
\]

\[
P(0|B_1, \lambda) - P(d - 1|B_1, \lambda) = \alpha_{B1} \quad P(0|B_1, \lambda) - P(d - 1|B_1, \lambda) = d\alpha_{B1}
\]

so that \(\alpha_{B1} = 0\) and hence all the \(P(j|B_1, \lambda)\) are equal and from the constraint \(1\) we must have \(P(j|B_1, \lambda) = P(0|B_1, \lambda) = 1/d\). Similar considerations show that \(P(j|A_1, \lambda) = P(j|B_2, \lambda) = P(j|A_2, \lambda) = 1/d\).

Overall we then find by substituting in \(20\) that

\[
I_{\text{max,min}} = 4/d
\]

so there is a local maximum or minimum for \(I\) which never exceeds 2 and decreases as the number of measurement outcomes \(d\) for each of the four observables becomes larger.

The question then arises: Is this a local maximum or minimum? To study this we consider an expansion of the expression for \(I\) given in Eq. \(20\) around the values \(P(j|B_1, \lambda) = P(j|A_1, \lambda) = P(j|B_2, \lambda) = P(j|A_2, \lambda) = 1/d\), and just retain the second order terms. The first order correction to \(I_{\text{max,min}}\) is \(\delta I_{(\text{first})} = 0\). The second order correction to
\( I_{\text{max, min}} = 4/d \) can be written in matrix form as

\[
\delta I_{(\text{sec. ord})} = \frac{1}{2} \sum_{\lambda} P(\lambda) \left[ \begin{array}{cccc}
\delta A_1' & \delta A_2' & \delta B_1' & \delta B_2'
\end{array} \right] \left[ \begin{array}{cccc}
0 & 0 & E & E \\
0 & 0 & D & E \\
E & D^t & 0 & 0 \\
E & E & 0 & 0
\end{array} \right] \left[ \begin{array}{c}
\delta A_1 \\
\delta A_2 \\
\delta B_1 \\
\delta B_2
\end{array} \right]
\]

where the notation for the sub-vectors is \( \delta A_t \equiv \{ P(0|A_1), P(1|A_1), \ldots, P(d-2|A_1), P(d-1|A_1), \} \),
\( \delta A'_2 \equiv \{ P(0|A_2), P(1|A_2), \ldots, P(d-2|A_2), P(d-1|A_2), \} \) with analogous expressions for \( \delta B_t \) and \( \delta B'_2 \). The transpose is indicated by \( t \). The \( d \times d \) sub-matrices of the \( 4d \times 4d \) matrix \( M \) shown in (61) are the unit sub-matrix \( E \) and \( D \), which is an off-diagonal matrix of the form

\[
D = \begin{bmatrix}
0 & 1 & 0 & \ldots & 0 & 0 \\
0 & 0 & 1 & \ldots & 0 & 0 \\
\vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\
0 & 0 & 0 & \ldots & 0 & 1 \\
1 & 0 & 0 & \ldots & 0 & 0 \\
(0) & (1) & \ldots & (d-2) & (d-1)
\end{bmatrix}
\]

The expression for \( \delta I_{(\text{sec. ord})} \) is a quadratic form involving the \( 4d \) real variables listed in \( \delta A_1', \delta A_2', \delta B_1' \) and \( \delta B_2' \). We note that \( M \) is a real, symmetric matrix, so its eigenvalues are all real and it would have \( 4d \) orthonormal column eigenvectors, all of which can be chosen to be real. If we choose a small real displacement of the single observable probabilities given by \( \left[ \begin{array}{c}
\delta A_1' \\
\delta A_2' \\
\delta B_1' \\
\delta B_2'
\end{array} \right] = \mu \left[ X^t \right] \), where \( \left[ X^t \right] \) is a normalised row eigenvector of \( M \) with eigenvalue \( \xi \), then in this case \( \delta I_{(\text{sec. ord})} = \frac{1}{2} \xi \mu^2 \). By considering all the eigenvalues, all possible displacements are included. Clearly, whether \( \delta I_{(\text{sec. ord})} \) is always negative (corresponding to \( I_{\text{max, min}} = 4/d \) being a maximum) requires all the eigenvalues to be negative. If all are positive then as \( \delta I_{(\text{sec. ord})} \) is positive, this corresponds to a minimum. A mixture of positive and negative (and zero) eigenvalues indicates a saddle point.

Evaluation of the eigenvalues have been carried out for the cases \( d = 2 \) and \( d = 3 \) using Mathematica. The outcome in both cases is that the eigenvalues are real and occur in pairs \(+a, -a\), including \( a = 0 \). This shows that in these cases \( I_{\text{max, min}} = 4/d \) is neither a maximum or a minimum, but is in fact a saddle point. Hence no useful result arises from this attempt to find a maximum value for \( I \) by explicitly exploiting the LHVT expression for \( C_{j,k,l,m} \).