Approximate joint measurement of qubit observables through an Arthur–Kelly model

Rajarshi Pal and Sibasish Ghosh

The Institute of Mathematical Sciences, CIT Campus, Taramani, Chennai-600 110, India

E-mail: rajarshi@imsc.res.in and sibasish@imsc.res.in

Received 7 June 2011, in final form 4 October 2011
Published 10 November 2011
Online at stacks.iop.org/JPhysA/44/485303

Abstract
We consider a joint measurement of two and three unsharp qubit observables through an Arthur–Kelly-type joint measurement model for qubits. We investigate the effect of the initial state of the detectors on the unsharpness of the measurement as well as the post-measurement state of the system. Particular emphasis is given on a physical understanding of the POVM to PVM transition in the model and entanglement between the system and the detectors. Two approaches for characterizing the unsharpness of the measurement and the resulting measurement uncertainty relations are considered. The corresponding measures of unsharpness are connected for the case where both the measurements are equally unsharp. The connection between the POVM elements and symmetries of the underlying Hamiltonian of the measurement interaction is made explicit and used to perform the joint measurement in arbitrary directions. Finally, in the case of three observables, we derive a necessary condition for the approximate joint measurement and use it to show the relative freedom available when the observables are non-orthogonal.

PACS numbers: 03.65.Ta, 03.67.−a, 02.30.Tb, 02.40.Ft

(Some figures may appear in colour only in the online journal)

1. Introduction

Quantum mechanics does not allow the joint measurement of non-commuting observables. The uncertainty principle, usually described as a lower bound on the product of standard deviations of the outcome statistics, does not really capture this complementary feature as it (uncertainty principle) is a statement about the quantum state of the system and does not relate to an actual situation involving apparatus which can attempt a joint measurement.

As intuition suggests, one has to allow for some degree of imprecision in order to make room for the notion of joint measurement of non-commuting observables. In quantum mechanics of a single system, observables correspond to self-adjoint operators with the
outcome statistics given by the Born rule, i.e. if \( P_i \) is the projector onto the eigenspace corresponding to the eigenvalue \( \lambda_i \), then in a measurement of the observable, the probability of obtaining the \( i \)th outcome when the system is in the state \( \rho \) is \( \text{Tr}(\rho P_i) \). However, this is not adequate. For example, suppose a system is allowed to interact with another system (ancilla) for a while and then the measurement is performed on the combined system. To describe the probabilities of obtaining various outcomes for such a measurement for different states of the original system before the interaction, one has to replace the projectors above by positive operators \( E_i \) (acting on the states of the original system) with \( \sum E_i = I \) for normalization. Such measurements are called POVMs (positive operator-valued measures). The most general measurements possible in quantum mechanics can be described through POVMs. It turns out that the idea of imprecise or unsharp measurements can be appropriately investigated if one considers instead of projective measurements POVMs. The observables corresponding to projective measurements form a subclass of those described by POVMs and are called sharp. Quantum theory of measurement is quite a developed subject and the interested reader is referred to [1] and [2].

The first model for the approximate joint measurement of position and momentum was given by Arthur and Kelly in (1965) by generalizing von Neumann’s model for measurement [3]. In von Neumann’s model for measurement, the position observable of the object is measured by coupling it to the momentum \( P_p \) of a probe system via the interaction evolution \( U = e^{-i\lambda \hat{Q} \otimes \hat{P}_p} \), and using the position \( \hat{Q}_p \) of the probe as the readout observable (because of the complete symmetry between position and momentum, one could alternatively couple through the position of the probe and use the probe momentum as readout; this is what we do here for the models we consider). The idea of Arthur–Kelly was to couple two such probe systems, respectively, to the position and momentum of the system and then perform measurement on the commuting meter observables of the probe to gain information about the position and momenta of the system. This is in fact an unsharp joint measurement and inspired the later development of the formalism. EPR-Bell argument for unsharp realities was given in [4]; the bounds on the joint measurement of qubit observables were derived starting from the condition of operational locality in [5]; the formalism for qubit observables and relevant inequalities for the two-observable case were developed in [6–9]. The connection of the approximate joint measurement with Bell inequalities was investigated in [10]; joint measurements by quantum cloning have been investigated for example in [11] and [12]. It has also been used to estimate the expectation values of qubit observables in [13]; the joint unsharp measurement of position and momentum was treated in [14] and [15]. Such measurements have also been proposed for monitoring a qubit in the context of continuous weak linear measurements, for example, by using a quantum point contact to probe a double-quantum-dot charge qubit [16–18].

In this paper, we consider an Arthur–Kelly-like model for the approximate joint measurement separately, of two and three non-commuting qubit observables. The probe systems we employ are continuous as typical in experiments like Stern–Gerlach, deviations in position or momenta of particles carrying spin are measured to obtain information about the system spin state. We consider the characterization of unsharpness or quality of approximation by two ways existing in the literature: first, by considering closeness of the marginal probability distribution of the probability distribution of the joint measurement to that of the sharp observable being approximately measured; second, by considering suitably defined closeness of the observables themselves with the meter observables in the Heisenberg picture. We show that for a symmetric joint measurement where the two marginal probability distributions are equally close to the corresponding sharp probability distributions, the two measures of closeness are proportional. The error-disturbance relationship [21] does not seem to hold for the measure based on the Heisenberg picture for our choice of the pointer observable. Numerical
analysis is performed for the two-observable case to show the validity of measurement uncertainty relations, transition from POVM to projection-valued measurement (PVM) and also the effect of the joint measurement on the system. The effect of the pre-measurement on the system turns out to be that of an asymmetric depolarizing channel and this forms the basis for a physical understanding of the POVM to PVM transition. Entanglement between the system and the detectors is also investigated. We also prove a lemma showing the connection between the joint measurement and the symmetries of the underlying Hamiltonian of the measurement interaction together with that of the initial detector states. This is then used to perform the approximate joint measurement in arbitrary directions. Moving on to the case of the three-observable joint unsharp measurement, we prove a simple necessary condition that is sufficient for the case of three orthogonal observables. The case of three orthogonal observables also generalizes the validity of the known necessary–sufficient condition for this case. This necessary condition is derived from certain geometrical considerations based on the so-called Fermat–Toricelli point. We show that it yields the known two-observable bounds in the limit that the measurement of one of the observables is pure guessing of the value of the observable. Finally, an extension of the Arthur–Kelly-like model to the three-observable case is studied.

The paper is organized as follows. In section 2, we introduce the joint unsharp spin measurement and the original Arthur–Kelly model given for the joint measurement of position and momentum. Measures characterizing the quality of approximations to the sharp observables being approximated are introduced in section 3. The corresponding measurement uncertainty relations are discussed. In section 4, we introduce the Arthur–Kelly-like model for qubits that we have considered and derive the final state of the system and meters after the measurement interaction. The approximate joint measurement of $\sigma_x$ and $\sigma_y$ is considered in section 5. Section 6 deals with the effect of the initial detector states on the joint measurement and post-measurement state of the system. In section 7, we try to develop a physical understanding of the POVM to PVM transition seen in the model based on the results obtained in the previous section. We also try to see how the entanglement between the system and the detectors behaves as the measurements become sharper. In section 8, we explore through a lemma the connection between the symmetries of the underlying Hamiltonian of the measurement interaction and initial detector states on the joint measurement. This is used in section 9 to perform the joint measurement of spin in arbitrary directions. The corresponding POVM elements are also calculated and matched with the orthogonal case in section 5. In section 10, we compute the spin direction fidelities, which were introduced as a measure of quality of approximation in section 3, for the model considered. It is compared with another measure based on the distance between outcome probabilities of sharp observables and their unsharp approximations. The approximate joint measurement of three-qubit observables is considered in section 11. A necessary condition on the parameters of the marginal POVM elements is derived by geometric considerations involving the Fermat–Toricelli point. The sufficiency of the condition is explored and the restrictions placed by it investigated in the context of the joint measurement through an extension of the model considered in section 4 to three detectors.

2. An introduction to joint unsharp spin measurements and the Arthur–Kelly model

Suppose we want to measure the spin of a spin-$\frac{1}{2}$ particle along a direction given by the unit vector $\hat{n}$. If the system is described by the density matrix $\rho$, then probabilities for obtaining outcomes $\pm$ are given in the POVM formalism, respectively, by $\text{Tr}(\rho A_+) \text{ and } \text{Tr}(\rho A_-)$. The POVM elements $A_+$ and $A_-$ are called effects. One says that the observable $A$ is characterized
by the specification of the map \( \omega_i \mapsto A_i \) where \( \omega \) belongs to the set of outcomes with \( i = +, - \) (e.g., here \( \omega_+ = +, \omega_- = - \)). If \( A_+ = |\hat{n}, +\rangle\langle \hat{n}, +| \), i.e., a projector on to the up state of \( \hat{\sigma}_z \), then the measurement is called a sharp measurement of \( \hat{\sigma}_z \) and the observable characterized by \( A_+ \) a sharp observable; else the measurement and the corresponding observable are called unsharp. One way in which such a measurement can be made is as follows. Suppose some one wants to perform a standard (sharp) measurement of \( \hat{\sigma}_z \) but because of some error in his setup, there is a finite probability of registering the outcome ‘-’ when the system is actually in the state \( |+\rangle \) and vice versa. This will be an unsharp measurement with the probabilities of error given by \( \text{Tr}(A_-|+\rangle\langle +|) \) and \( \text{Tr}(A_+|-\rangle\langle -|) \). Such a situation can arise if for example in a Stern–Gerlach setup the beam passing through the inhomogeneous magnetic field is poorly collimated initially (see [22] and also section 7).

2.1. Joint measurability

Two observables are said to be jointly measurable if there is a measurement scheme that allows the determination of values of both the observables. This means that the POVM describing the measurement scheme contains the POVM elements of the two observables as marginals. In this way, it is ensured that there is a joint probability distribution corresponding to the pairs of different values of the two observables for each state. The two observables \( \hat{\Upsilon}_1 \) and \( \hat{\Upsilon}_2 \) are jointly measurable if there is an observable \( G : \omega_{ij} \rightarrow G_{ij}, i, j = \pm \), such that

\[
\begin{align*}
\hat{\Upsilon}_1^+ &= G_{++} + G_{+-} \\
\hat{\Upsilon}_2^+ &= G_{++} + G_{-+} \\
\hat{\Upsilon}_1^- &= G_{--} + G_{-+} \\
\hat{\Upsilon}_2^- &= G_{-+} + G_{-\cdot}
\end{align*}
\]

with the POVM elements \( \hat{\Upsilon}_{ij}^\pm \) satisfying \( \hat{\Upsilon}_1^+ + \hat{\Upsilon}_1^- = I \) and \( \hat{\Upsilon}_2^+ + \hat{\Upsilon}_2^- = I \). The outcomes \( \omega_{ij} \) of \( G \) can be taken to be the pairs \( (\omega_1, \omega_2) \), where \( \omega_1, \omega_2 \in \{+, -\} \). The map implies that the probability of registering outcome, say \((+, +)\), is given by \( \text{Tr}(\rho G_{++}) \) and so on. The extension to more than two observables is done in the obvious way.

The two observables \( \hat{\Upsilon}_1 \) and \( \hat{\Upsilon}_2 \) are said to commute if \( \hat{\Upsilon}_1^i \hat{\Upsilon}_2^j = \hat{\Upsilon}_2^j \hat{\Upsilon}_1^i \forall i, j = \pm \). The possibility of joint measurability is guaranteed by the following theorem which we state without proof (for the proof, see [23]).

**Theorem 1.** A pair of sharp observables is jointly measurable iff they commute. Commutativity of unsharp observables is sufficient but not necessary for joint measurability.

2.2. Arthur–Kelly model

As mentioned earlier, Arthur–Kelly gave a model for the joint measurement of position and momentum [3]. In this model, a quantum object is coupled with two probe systems which are then individually measured to obtain information about the object’s position and momentum. The coupling to probes is based on von-Neumann’s model of measurement. They showed that this constitutes a simultaneous measurement of position and momentum in the sense that the output statistics reproduce the expectation values of the object’s position and momentum.

The position \( \hat{Q} \) and momentum \( \hat{P} \) of the object are coupled with the position \( \hat{Q}_1 \) and the momentum \( \hat{P}_2 \) of the two probe systems, respectively, which serve as the readout observables.
Neglecting the free evolutions of the three systems (assuming that the measurement interaction dominates during the short time in which it acts), the combined time evolution is described by

$$ U = \exp^{i\lambda \hat{Q}_1 \hat{P}_1 t + \kappa \hat{P}_2 \hat{Q}_2 t} .$$  

(2)

The coupling constants $\lambda$ and $\kappa$ represent the interaction strengths and can be absorbed into a rescaling of the pointer observables $\hat{Q}_1$ and $\hat{P}_2$. If $|\psi\rangle$ is an arbitrary input state of the object and $|\Psi_1\rangle$ and $|\Psi_2\rangle$ are fixed initial states of the probes (well behaved and zero expectations for each of the probe’s position and momentum), the probabilities for the values of $\alpha, \beta$ to the one defined in equation (1), through

$$ |\langle \psi | G_j^T (X \times Y) |\psi\rangle| := (\langle \psi | \otimes |\Psi_1\rangle \otimes |\Psi_2\rangle) U^j (\hat{Q}_1 (\lambda X) \otimes \hat{P}_2 (\kappa Y)) U (|\psi\rangle \otimes |\Psi_1\rangle \otimes |\Psi_2\rangle) .$$  

(3)

The marginals are given by $G_j^T (X) := G_j^T (X \times \mathcal{R})$ and $G_j^T (Y) := G_j^T (\mathcal{R} \times Y)$. The cost of the joint measurement is an increase in the variance of the marginals and is given by

$$ \Delta(G_j^T, \psi) \Delta(G_j^T, \psi) \geq \hbar ,$$  

(4)

where $\Delta(G_j^T, \psi) = \sqrt{\langle \psi | G_j^{T^2} |\psi\rangle - \langle \psi | G_j^T |\psi\rangle^2}$ for $j = 1, 2$.

3. Quality of the joint measurement and measurement uncertainty relations

The notion of approximate joint measurement naturally demands a measure of proximity to the sharp observables being approximately measured. The restriction on the measures corresponding to non-commuting observables leads one to measurement uncertainty relations. In this paper, we use two such approaches to characterize proximity.

3.1. Closeness of probabilities

Proximity between two observables can be characterized by the distance between the corresponding probability distributions for all states. Thus, the distance between the two observables $\hat{A}$ and $\hat{B}$ can be defined as

$$ D(\hat{A}, \hat{B}) := \max_j \sup_{\rho} \|\text{tr}[TA_j] - \text{tr}[TB_j]\| ,$$  

(5)

where $A_j$ (or $B_j$) corresponds to the POVM element (or effect) of the observable $A$ (or $B$) associated with the measurement outcome $j$. $T$ is the density matrix of the system [6].

We will be considering single-qubit observables characterized by the two effects $\hat{Y}^{\alpha, \alpha', \sigma}$ and $\hat{Y}^{\beta, \beta', \sigma'}$, respectively, to the two outcomes $+ \text{ and } -$ with

$$ \hat{Y}^{\alpha, \alpha', \sigma} = \frac{1}{2} (\alpha \hat{I} + \alpha' \hat{I} \sigma) ,$$  

(6)

with $(\alpha, \alpha') \in \mathbb{R}^4$.

Now for the observables $\hat{A}$ and $\hat{B}$ with the respective set of effects $\{A_+, A_-, \{B_+, B_-\}$, we have $\|\text{tr}[TA_+] - \text{tr}[TB_+]\| = |\text{tr}[TA_-] - \text{tr}[TB_-]|$.

For the two single-qubit observables $\hat{Y}^{\alpha, \alpha'}$ and $\hat{Y}^{\beta, \beta', \sigma'}$, taking $T$ to be the `$+$/` eigenstate or `$-$/` eigenstate of $\hat{\sigma} \cdot \frac{\alpha \hat{I} - \beta \hat{I}}{||\alpha \hat{I} - \beta \hat{I}||}$ according to whether $(\alpha - \beta) > 0$ or $(\alpha - \beta) < 0$, respectively, one has

$$ D(\hat{Y}^{\alpha, \alpha'}, \hat{Y}^{\beta, \beta'} ) = \frac{1}{2} ||\alpha \hat{I} - \alpha' \hat{I}|| + \frac{1}{2} |\alpha - \beta| .$$  

(7)

This shows that the distance of a certain unsharp observable $\hat{Y}^{\alpha, \alpha'}$ from any sharp observable $\hat{Y}^{\alpha, \alpha'}$ is minimum when $\alpha' = 0$. 

5
3.1.1. Unbiased observables. Observables of the form $\hat{\mathbf{a}}_{1,2}$ are called unbiased. This is because, for such an observable, both outcomes are equally likely for a maximally mixed state. Also, the expectation value of the unsharp measurement when the system is in the state $T$ is given by $\langle \hat{\mathbf{a}}_1 \rangle_u := 1 \cdot \text{tr}(T \frac{1}{2} (I + \hat{\mathbf{a}}_1 \cdot \hat{\mathbf{a}}_1)) - 1 \cdot \text{tr}(T \frac{1}{2} (I - \hat{\mathbf{a}}_1 \cdot \hat{\mathbf{a}}_1)) = ||\hat{a}_1|| ||\mathbf{a}_1||$. Again as $D(\hat{T}_{1,\mathbf{a}}, \hat{T}_{1,\mathbf{a}}) = \frac{1}{2} (1 - ||\hat{a}_1||)$, $||\hat{a}_1||$ itself serves as a measure of proximity. We will often approximate sharp observables with unbiased unsharp ones.

3.1.2. Measurement uncertainties. We will choose jointly measurable observable pairs $(\hat{T}_{\alpha,\mathbf{a}}, \hat{T}_{\beta,\mathbf{a}})$ to approximate the sharp pair $(\hat{T}_{1,\mathbf{a}}, \hat{T}_{1,\mathbf{a}})$. The necessary and sufficient conditions on $\alpha, \beta, \mathbf{a}_1, \mathbf{a}_2$ so that the first pair is jointly measurable in the sense of (1) are formulated in [7–9]. When the observables are unbiased, the conditions simplify to

$$||\hat{a}_1 + \hat{a}_2|| + ||\hat{a}_1 - \hat{a}_2|| \leq 2.$$  

(8)

It was further shown in [6] that

$$D(\hat{T}_{\alpha,\mathbf{a}}, \hat{T}_{1,\mathbf{a}}) + D(\hat{T}_{\beta,\mathbf{a}}, \hat{T}_{1,\mathbf{a}}) \geq 2D_0,$$

(9)

with $D_0 = \frac{1}{\sqrt{2}} \text{tr}(\cos(\theta/2) + \sin(\theta/2) - 1)$, $\theta$ being the angle between $\mathbf{a}_1$ and $\mathbf{a}_2$. The conditions for attainment of the lower bound $2D_0$ were also spelt out. The approximate observables should be unbiased, i.e. $\alpha = \beta = 1$ in this later case. The other conditions (see [6]) imply that for optimality, $\mathbf{a}_1$ and $\mathbf{a}_2$ lie along $\hat{n}$ and $\hat{m}$, respectively, only when the latter are orthogonal.

3.2. Closeness of observables

A completely different approach was taken in [19] and [20] in the context of the original Arthur–Kelly model to give a formulation of the complementary nature of the approximate joint measurement process. If $\hat{\mu}_x$ and $\hat{\mu}_p$ denote the pointer observables used to measure the system position and momentum, respectively, then the retrodictive error operators are defined as

$$\hat{\epsilon}_x = \hat{\mu}_x - \hat{x}, \quad \hat{\epsilon}_p = \hat{\mu}_p - \hat{p},$$  

(10)

the predictive error operators as

$$\hat{\epsilon}_x = \hat{\mu}_x - \hat{x}, \quad \hat{\epsilon}_p = \hat{\mu}_p - \hat{p},$$  

(11)

and the disturbance operators as

$$\hat{\delta}_x = \hat{x} - \hat{x}, \quad \hat{\delta}_p = \hat{p} - \hat{p},$$  

(12)

where the operators $\hat{\delta}_f \in \{\hat{\mu}_x, \hat{\mu}_p, \hat{x}, \hat{p}\}$ appearing on the right, stand for the final Heisenberg picture operator $\hat{O}$ after the measurement interaction $U$, i.e. $\hat{O} = U^\dagger \hat{O}U$ where $\hat{O}$ is the Heisenberg picture operator at the moment the interaction starts. Various errors were then defined by taking the square root of expectation of the square of the operators defined above and taking supremum over the system states. For example, the maximal error of retrodiction was defined as

$$\Delta_{e_x} = \sup_{\psi} \langle \psi | \hat{H}_m | \psi \rangle (\psi \otimes \Psi_1 \otimes \Psi_2 | \epsilon_x^2 | \psi \otimes \Psi_1 \otimes \Psi_2 ) \frac{1}{2},$$  

(13)

and similarly for $\Delta_{e_p}, \Delta_{e_f, x}, \Delta_{e_f, p}, \Delta_{d_x}$ and $\Delta_{d_p}$ (see [19, 20] for details). The measurement uncertainty or ‘error principle’ was shown to hold in the form of

$$\Delta_{e_x} \Delta_{e_p} \geq \hbar \frac{1}{2},$$  

(14)

$$\Delta_{e_f, x} \Delta_{e_f, p} \geq \hbar \frac{1}{2}, \quad \Delta_{e_x} \Delta_{d_p} \geq \hbar \frac{1}{2},$$  

(15)
\[ \Delta_{ef}, \Delta_{dp} \geq \frac{\hbar}{2}, \]

and extensions of the above uncertainties are obtained by interchanging \( x \) and \( p \). One of the important features of this approach is the difference between the error of retrodiction and that of prediction. It was shown in [19, 20] that these are not the same as long as there is a finite disturbance.

### 3.3. Qubit observables

In a similar spirit, for the case of the approximate joint measurement of qubit observables through an Arthur–Kelly-like process, fidelities were defined in the Heisenberg picture that would provide a notion of the direction of the spin [21] of the system. As in the previous case, the distinction was made between errors of retrodiction and prediction. In this paper, we consider only the type-1 measurements considered by the author of [21].

We next consider the fidelities as defined by the authors. The retrodictive fidelity is defined as

\[ \eta_s = \inf_{|\chi\rangle \in H_{\text{sys}}} \langle \psi | \otimes \langle \chi | \frac{1}{2} (\hat{n}_f \cdot \hat{S}_i + \hat{S}_i \cdot \hat{n}_f) | \psi \rangle \otimes | \chi \rangle, \]

where \( \hat{S}_i = \hat{S} \) and \( \hat{n}_i = \hat{n} \) are the initial values of the Heisenberg spin and pointer observables, respectively, and \( \hat{S}_f = U \hat{S} U \) respectively are the final Heisenberg pointer and spin direction observables after the measurement interaction.

The predictive fidelity is defined as

\[ \eta_d = \inf_{|\chi\rangle \in H_{\text{sys}}} \langle \psi | \otimes \langle \chi | \frac{1}{2} (\hat{n}_f \cdot \hat{S}_i + \hat{S}_i \cdot \hat{n}_f) | \psi \rangle \otimes | \chi \rangle. \]

The measurement disturbance is defined by

\[ \eta_d = \inf_{|\chi\rangle \in H_{\text{sys}}} \langle \psi | \otimes \langle \chi | S \hat{f}^\dagger (\hat{n}_f \cdot \hat{S}_i + \hat{S}_i \cdot \hat{n}_f) | \psi \rangle \otimes | \chi \rangle. \]

The intuition behind the above definitions is classical in the sense that it considers the alignment of the initial or final spin vector and the initial or final pointer direction. But the above fidelities were used to define maximal root mean square (rms) error of retrodiction,

\[ \Delta_{ef}S = \sup_{|\chi\rangle \in H_{\text{sys}}} \langle \psi | \otimes \langle \chi | \hat{n}_f \hat{S}_i - \hat{S}_i \hat{n}_f | \psi \rangle \otimes | \chi \rangle \]

\[ = (s + s^2 - \eta^2)^{\frac{1}{2}}, \]

maximal rms error of prediction,

\[ \Delta_{ef}S = \sup_{|\chi\rangle \in H_{\text{sys}}} \langle \psi | \otimes \langle \chi | \hat{n}_f \hat{S}_i - \hat{S}_i \hat{n}_f | \psi \rangle \otimes | \chi \rangle \]

\[ = (s + s^2 - \eta^2)^{\frac{1}{2}}, \]

and maximal rms disturbance

\[ \Delta_{d}S = \sup_{|\chi\rangle \in H_{\text{sys}}} \langle \psi | \otimes \langle \chi | \hat{S}_f \hat{S}_i - \hat{S}_i \hat{S}_f | \psi \rangle \otimes | \chi \rangle \]

\[ = \sqrt{2} (s + s^2 - \eta^2)^{\frac{1}{2}}, \]

where the spin \( s = \frac{1}{2} \) for our case. The quantities \( \Delta_{ef}S, \Delta_{ef}S \) and \( \Delta_{d}S \) were expected to play the same role for these measurements as similar quantities defined for the original Arthur–Kelly model. The authors in [21] also showed that the retrodictive and predictive fidelities reach their optimal values of \( s = \frac{1}{2} \) when the Krauss operators of the measurement POVM have a spin-coherent form defined using spin-coherent states. A physical model for such a measurement was given in [24].

No measurement uncertainties were derived in [21] and the question was left open for further investigation.
4. The model

We consider an instantaneous coupling interaction with the help of the Hamiltonian of the form

\[ H = -(\hat{q}_1 \otimes \sigma_x + \hat{q}_2 \otimes \sigma_y) \delta t \]  

(23)

(in \( \hbar = 1 \) units). Possible coupling constants in the above equation such as \( \lambda \) and \( \kappa \) in equation (2) have been absorbed by rescaling \( \hat{q}_1 \) and \( \hat{q}_2 \). Like the original Arthur–Kelly interaction (2), the idea is to entangle the detectors with the system through \( H \) and then perform a projective measurement of \( \hat{p}_1 \) and \( \hat{p}_2 \) to obtain the spin information. Now, as a consequence of the Ehrenfest theorem, the average momentum change of a particle carrying spin that experiences the above interaction is given by

\[ \langle \dot{\hat{p}}_1 \rangle = \langle \sigma_x \rangle \quad \text{and} \quad \langle \dot{\hat{p}}_2 \rangle = \langle \sigma_y \rangle. \]

Thus, an ensemble of particles whose spin state is \(|+\rangle_x\rangle\) and which has a symmetric distribution of \( p_1 \) before the interaction will have a greater probability of having a positive \( p_1 \) after it. The signs in equation (23) have been chosen so that this fact is true for both the \( x \) and \( y \) directions. The signs thus allow us to map the four quadrants of the momentum plane \((p_1, p_2)\) to the four outcomes of the joint measurement and take the signs of the momenta to correspond to the outcomes \((+, +), (+, -)\) and \((-+, -), (-, -)\) of the joint measurement. Note that for a Stern–Gerlach situation, the two terms in equation (23) should have opposite signs to satisfy divergenceless of the magnetic field.

Models similar to the above have been considered before, for example, in [22, 25, 26]. As shown in [25], this model naturally arises in the context of a Stern–Gerlach experiment with a linear magnetic field.

We further assume that the measurement interaction (23) is strong enough to dominate the other parts of the Hamiltonian during its presence (e.g., the kinetic energy part). In the Stern–Gerlach context, this would mean assuming the atoms carrying spin to be sufficiently massive.

The unitary evolution corresponding to the Hamiltonian of equation (23), obtained by integrating the time evolution operator, is given by

\[ U = \exp(i(\hat{q}_1 \otimes \sigma_x + \hat{q}_2 \otimes \sigma_y)). \]  

(24)

On direct expansion, \( U \) turns out to have the simple form

\[ U = e(\hat{q}_1, \hat{q}_2) \otimes 1_s + f(\hat{q}_1, \hat{q}_2) \otimes \sigma_s + g(\hat{q}_1, \hat{q}_2) \otimes \sigma_y, \]

with

\[ e(\hat{q}_1, \hat{q}_2) = \cos \left( \sqrt{(\hat{q}_1^2 + \hat{q}_2^2)} \right), \]

\[ f(\hat{q}_1, \hat{q}_2) = i\hat{q}_1 \frac{\sin \left( \sqrt{(\hat{q}_1^2 + \hat{q}_2^2)} \right)}{\sqrt{(\hat{q}_1^2 + \hat{q}_2^2)}} \]

\[ g(\hat{q}_1, \hat{q}_2) = i\hat{q}_2 \frac{\sin \left( \sqrt{(\hat{q}_1^2 + \hat{q}_2^2)} \right)}{\sqrt{(\hat{q}_1^2 + \hat{q}_2^2)}}. \]  

(25)
Thus, the final state of the system plus apparatuses is given by

\[
|\psi_f\rangle = \int_{q_1, q_2 = -\infty}^{+\infty} e(q_1, q_2)|q_1, q_2\rangle \psi_1(q_1) \psi_2(q_2) \, dq_1 \, dq_2 \otimes |\chi\rangle
\]

\[
+ \int_{q_1, q_2 = -\infty}^{+\infty} f(q_1, q_2)|q_1, q_2\rangle \psi_1(q_1) \psi_2(q_2) \, dq_1 \, dq_2 \otimes \sigma_i |\chi\rangle
\]

\[
+ \int_{q_1, q_2 = -\infty}^{+\infty} g(q_1, q_2)|q_1, q_2\rangle \psi_1(q_1) \psi_2(q_2) \, dq_1 \, dq_2 \otimes \sigma_j |\chi\rangle,
\]

with the initial state being \(|\psi_i\rangle = |\psi_1\rangle \otimes |\psi_2\rangle \otimes |\chi\rangle\). Let \(\rho = |\psi_f\rangle \langle \psi_f|\).

5. Approximate joint measurement in orthogonal directions

In this section, we consider the joint measurement of \(\sigma_x\) and \(\sigma_y\). We choose the observables \(\hat{p}_1\) and \(\hat{p}_2\) to serve as meters. As mentioned before, \((p_1 \geq 0, p_2 \geq 0)\) is taken to correspond to the outcome \((\sigma_x = 1, \sigma_y = 1) \equiv (+, +)\) of the joint measurement, \((p_1 \geq 0, p_2 \leq 0)\) to \((\sigma_x = 1, \sigma_y = -1) \equiv (+, -)\), \((p_1 \leq 0, p_2 \geq 0)\) to \((\sigma_x = -1, \sigma_y = 1) \equiv (-, +)\) and \((p_1 \leq 0, p_2 \leq 0)\) to \((\sigma_x = -1, \sigma_y = -1) \equiv (-, -)\).

After the interaction \(U\) between the system and meters, the projective measurement is performed separately for the observables \(\hat{p}_1\) and \(\hat{p}_2\). The probability of obtaining the outcome \((p_1, p_2)\) is given by

\[
p(p_1, p_2) = \text{Tr}((p_1, p_2)[\rho]_{(p_1, p_2)})
\]

\[
= \langle |\psi_1\rangle \psi_2 |\psi_1\rangle \psi_2 |\chi\rangle + 2 \text{Re}(f_0 e^{0^*}) \langle \chi | \sigma_i | \chi \rangle
\]

\[
+ 2 \text{Re}(g^0 e^{0^*}) \langle \chi | \sigma_j | \chi \rangle - 2 \text{Im}(g^0 f^{0^*}) \langle \chi | \sigma_i | \chi \rangle,
\]

with \(e^0\), \(f^0\) and \(g^0\) representing respectively the Fourier transforms of \(e \psi_1 \psi_2 = e(q_1, q_2) \psi_1(q_1) \psi_2(q_2)\), \(f \psi_1 \psi_2\) and \(g \psi_1 \psi_2\).

The initial pointer states are taken to be Gaussian,

\[
\psi_1(q_1) = \left[ \frac{1}{(a\sqrt{2\pi})} \exp\left[ -\frac{q_1^2}{2a^2} \right] \right]^1,
\]

\[
\psi_2(q_2) = \left[ \frac{1}{(b\sqrt{2\pi})} \exp\left[ -\frac{q_2^2}{2b^2} \right] \right]^1,
\]

satisfying \(\int_{-\infty}^{\infty} |\psi_j(q_j)|^2 \, dq_j = 1\) for \(j = 1, 2\).

We have chosen the initial states to be even in \(q_1\) and \(q_2\). Now as a Fourier transform of a real even function is a real even function and that of an imaginary odd function is a real odd function, we have \(f^0\) odd in \(p_1\) and even in \(p_2\), \(g^0\) the other way around and \(e^0\) even in both. Also, each of them is real. Thus, the \(\sigma_i\) term in (27) vanishes. Henceforth in this paper, we refer to these properties of \(e^0\), \(f^0\) and \(g^0\) as ‘parity properties’.

We have for the probability of outcome \((+, +)\) from equation (27)

\[
p(p_1 \geq 0, p_2 \geq 0) = \int_{p_1 = 0}^{\infty} \int_{p_2 = 0}^{\infty} \{\langle |\psi_1\rangle \psi_2 |\psi_1\rangle \psi_2 |\chi\rangle + 2 \text{Re}(f_0 e^{0^*}) \langle \chi | \sigma_i | \chi \rangle
\]

\[
+ 2 \text{Re}(g^0 e^{0^*}) \langle \chi | \sigma_j | \chi \rangle \} \, dp_1 \, dp_2.
\]

One also of course has to satisfy

\[
p(p_1 \geq 0) + p(p_1 \leq 0) = 1,
\]
which yields, due to the ‘parity properties’,

\[
\int_{p_1=0}^{\infty} \int_{p_2=0}^{\infty} (|e|^2 + |f|^2 + |g|^2) \, dp_1 \, dp_2 = \frac{1}{4}.
\]  

(31)

From equation (31) and ‘parity properties’, we have

\[
p(p_1 \geq 0, p_2 \geq 0) = \langle \chi | \frac{1}{4} + \frac{a'\sigma_x}{4} + \frac{b'\sigma_y}{4} | \chi \rangle = \langle \chi | G_{++} | \chi \rangle.
\]  

(32)

with \(G_{++} = \left( \frac{1}{4} + \frac{a'\sigma_x}{4} + \frac{b'\sigma_y}{4} \right)\) and

\[
a' = \int_{p_1=0}^{+\infty} \int_{p_2=-\infty}^{p_1=+\infty} 4|f^0e^0| \, dp_1 \, dp_2,
\]  

(33)

\[
b' = \int_{p_2=0}^{+\infty} \int_{p_1=-\infty}^{p_2=+\infty} 4|g|e^0 \, dp_1 \, dp_2.
\]  

(34)

For the other outcomes, we have from the consideration of the corresponding momentum probabilities,

\[
G_{+-} = \frac{1}{4} - \frac{a'\sigma_x}{4} - \frac{b'\sigma_y}{4},
\]  

\[
G_{-+} = \frac{1}{4} + \frac{a'\sigma_x}{4} - \frac{b'\sigma_y}{4},
\]  

(35)

\[
G_{-+} = \frac{1}{4} - \frac{a'\sigma_x}{4} + \frac{b'\sigma_y}{4}.
\]

From equations (32) and (35), we obtain the marginal unsharp observables as (see equation (1))

\[
\gamma_1^1 = \frac{1}{2}(I \pm a'\sigma_x),
\]  

\[
\gamma_2^1 = \frac{1}{2}(I \pm b'\sigma_x).
\]  

(36)

Thus, we see that the approximate observables \(\gamma_1^1\) and \(\gamma_2^1\) are unbiased and \(a'\) and \(b'\) themselves serve as measures of proximity to the sharp observables \(\frac{1}{2}(I \pm \sigma_x)\) and \(\frac{1}{2}(I \pm \sigma_y)\), respectively.

In section 8, we will see how the approximate joint measurement, characterized by the marginals in equation (36), arises as a consequence of symmetries of the Hamiltonian in equation (23) and that of the initial detector states \(\psi_1\) and \(\psi_2\) (given in equations (28) and (29)).

5.1. rdm of the system after pre-measurement

The interaction \(U\), described by equation (24), acting on the system and the detectors induces a completely positive map on the system. The rdm (reduced density matrix) of the system after the interaction is given by

\[
\rho_f = \text{Tr}_{1,2}(|\psi_f\rangle\langle\psi_f|)
\]  

(37)

\[
= \int_{q_1, q_2 = -\infty}^{+\infty} |e|^2|\psi_1|^2|\psi_2|^2 \, dq_1 \, dq_2 |\chi\rangle \langle \chi | + \int_{q_1, q_2 = -\infty}^{+\infty} |f|^2|\psi_1|^2|\psi_2|^2 \, dq_1 \, dq_2 |\sigma_x\rangle \langle \sigma_x | + \int_{q_1, q_2 = -\infty}^{+\infty} |g|^2|\psi_1|^2|\psi_2|^2 \, dq_1 \, dq_2 |\sigma_y\rangle \langle \sigma_y | \quad (38)
\]

\[
+ \int_{q_1, q_2 = -\infty}^{+\infty} |g|^2|\psi_1|^2|\psi_2|^2 \, dq_1 \, dq_2 |\sigma_y\rangle \langle \sigma_y | \quad (39)
\]
\[ = \sum_{j=1}^{3} K_j |\chi\rangle \langle \chi | K_j^\dagger. \quad (40) \]

The action of the measurement interaction on the system is thus that of an asymmetric depolarizing channel, with the Krauss operators given by

\[ K_1 = 2\sqrt{c_f} \sigma_1, \quad K_2 = 2\sqrt{c_g} \sigma_2 \quad \text{and} \quad K_3 = \sqrt{(1 - 4c_f - 4c_g)} I, \]

where

\[ c_f = \int_{q_1, q_2 = 0}^{+\infty} |f|^2 \psi_1^2 \psi_2^2 \, dq_1 \, dq_2, \]

\[ c_g = \int_{q_1, q_2 = 0}^{+\infty} |g|^2 \psi_1^2 \psi_2^2 \, dq_1 \, dq_2. \quad (41) \]

If \((x, y, z)\) denotes the Bloch vector of \(|\chi\rangle \langle \chi|\), then in the Bloch sphere representation, we have

\[ \rho_s^f = \frac{1}{2} (I + x(1 - 8c_g) \sigma_x + y(1 - 8c_f) \sigma_y). \quad (42) \]

### 6. Effect of initial detector states

The form of the POVM elements in equations (32) and (35) which leads to unbiased marginal observables (see section 3.1.1) only depends on the ‘parity properties’ and not on the Gaussian form of the initial detector states given by equations (28) and (29).

The integrals for computing the coefficients \(c_f\) and \(c_g\) have been performed using the Monte Carlo integrator included in GNU Scientific Library. The MISER algorithm has been used which uses stratified random sampling for performing ‘importance’ sampling [27].

**Results.** As we are considering the unbiased approximate joint measurement in orthogonal directions, equation (8) applied on the parameters \(a'\) and \(b'\) of the marginals in equation (36) yields

\[ a'^2 + b'^2 \leq 1. \quad (43) \]

#### 6.1. Symmetric case

In this case, the standard deviations \(a\) and \(b\) of the initial momentum wavefunctions in equations (28) and (29) are taken to be equal. As there is nothing else that differentiates between the detectors, one must have here \(a' = b'\). This is reflected in equation (33). Both the observables \(\sigma_x\) and \(\sigma_y\) are approximated equally well. From equation (43), we have \(a' \leq \frac{1}{\sqrt{2}} = 0.707\).

#### 6.1.1. Quality. In figure 1, we see that when \(a\) and \(b\) are close to zero, which corresponds to the fact that the initial momenta \(p_1\) and \(p_2\) have large spread, we have almost no information about the spin state by looking at the momentum. In this case, \(a'\) and \(b'\) are both close to zero and we have trivial marginal effects (\(\frac{1}{2}\)) which represent guessing the value of the corresponding sharp observable with equal probability for + and –. With an increase in \(a\), \(a'\) increases and touches the value 0.628 at \(a = b = 0.7\) (see figure 1(a)). The graph of \(a'\) versus \(a\) is identical to the one obtained in [25] for the same case. The maximum in the curve is a feature of the interaction used and is explained in the following section.
Figure 1. Effects of initial pointer states on the joint measurement and on the system state after the measurement interaction for \( a = b \). (a) A plot of \( a' \) versus \( a \). (b) A plot of \( 1 - 8c_f \) versus \( a \).

6.1.2. Disturbance due to the measurement interaction. Figure 1(b) shows the variation of \( \frac{\langle \sigma_i \rangle_f}{\langle \sigma_i \rangle_i} \) with \( a \), where \( \langle \sigma_i \rangle_f = \text{Tr}(\rho_f \sigma_i) \) and \( \langle \sigma_i \rangle_i = \langle \chi | \sigma_i | \chi \rangle \), and similarly for \( \sigma_j \) (see equation (42)).

Corresponding to the maximum in \( a' \), we also have a minimum in the disturbance of the state characterized by \( \frac{\langle \sigma_i \rangle_f}{\langle \sigma_i \rangle_i} = \frac{\langle \sigma_i \rangle_f}{\langle \sigma_i \rangle_i} \). Thus, a sharper measurement setting seems to disturb the state less, though there is a slight difference in the value of \( a \) at which the maximum and the minimum occur. This disturbance is even more prominent in the asymmetric case to be discussed next.
Figure 2. Effects of initial pointer states on the joint measurement and post-measurement system state for $b \geq a = 0.1$. (a) A plot of $a'$ and $b'$ versus $b$ for $a = 0.1$. (b) A plot of $(1 - 8c_f)$, $(1 - 8c_g)$ (dotted line) versus $b$.

6.2. $b \geq a$: POVM to PVM transition

Let us choose $a = 0.1$. The lhs of equation (43) starts off at a low value for $b = a$ and gradually increases and closes on the bound 1.0 as $b$ becomes much greater than $a$ (see figure 2(a)). $b$ much greater than $a$ reflects the situation where the initial momentum wavefunction of apparatus 2 is much sharper than that of apparatus 1. As shown clearly in figure 2(a), this marks a transition from a POVM measurement to a PVM in the sense that the unsharp measurement of $\sigma_j$ becomes almost sharp. The requirement of complementarity is satisfied
by the fact that the unsharp $\sigma_x$ measurement becomes almost trivial. The fact that $a'$ and $b'$ depend both on $a$ and $b$ is a reflection of the correlations between $p_1$ and $p_2$ brought about by the unitary evolution given by equation (24).

6.2.1. Disturbance due to the measurement. As seen in figure 2(b), as the measurement becomes sharper ($b$ increases with $a = 0.1$, $b'$ tends to 1 while $a'$ tends to 0), the $\langle \sigma_y \rangle$ information of the initial density matrix is almost kept intact ($\langle \sigma_y \rangle_f = (1 - 8c_f) \langle \chi | \sigma_y | \chi \rangle$), while the $\langle \sigma_x \rangle$ information gets very disturbed ($\langle \sigma_x \rangle_f = (1 - 8c_g) \langle \chi | \sigma_x | \chi \rangle$). This is reminiscent of a Heisenberg Gamma–Ray microscope kind of a situation where using a short wavelength light to reduce the uncertainty in the position measurement of an electron disturbs the momentum of the electron.

Next we choose $a = 1.0$. The conclusions are almost similar with the exception that the joint measurement uncertainty (i.e. the lhs of equation (8)) starts off at a much higher value compared to that of figure 2(a) (see figure 3).

7. Physics of the model and entanglement

In this section, we first try to understand the results obtained in the previous section. It is instructive to first look at a single approximate measurement arising from a von Neumann model. As mentioned before, the situation is almost like a Stern–Gerlach experiment. So we consider a neutral particle of mass $m$ carrying spin $\frac{1}{2}$ which is propagating in the $z$-direction and passes through a magnetic field $\vec{B} = 2B_0 \hat{x}$ ($\hat{x}$ is the unit vector along $x$ direction) for a time interval $\tau$. The position wavefunction of the particle before it enters the magnetic field is given by $\psi_1(x) = \left[ \frac{1}{\sqrt{\pi \sigma^2}} \exp[-x^2/2\sigma^2] \right]^\frac{1}{2}$. The interaction Hamiltonian is taken to be $H = \frac{\vec{p}^2}{2m} - \vec{S} \cdot \vec{B}$. The unitary evolution corresponding to this interaction is given by

$$U = \exp \left( iB_0 \tau \sigma_x \otimes \sigma_z - i \left( \frac{\vec{p}_0}{2m} \otimes 1 \right) \tau \right) \text{ (in } \hbar = 1 \text{ units).} \quad (44)$$
The strong impulsive coupling approximation that we employed in section 5 to neglect the kinetic energy term amounts to \( \frac{ma^2t}{2} \ll B_0a \). We assume that the particle is sufficiently massive for this to hold for the values of \( a \) we have considered. We also note that in a usual Stern–Gerlach setting, the initial position wavefunction is taken to be sharp and hence this assumption will breakdown for sufficiently small \( a \). After neglecting the kinetic energy, we have

\[
|\psi_f\rangle = U(|\psi\rangle \otimes |\chi\rangle) \\
= \int_{p_x=-\infty}^{\infty} dp_x \left( |+\rangle \psi(p_x - \lambda) |p_x\rangle \otimes |+\rangle_x + |\rangle \psi(p_x + \lambda) |p_x\rangle \otimes |\rangle_x \right),
\]

with \( \lambda = B_0 \tau \) and \( \sigma_x |\pm\rangle_x = \pm |\pm\rangle_x \).

If \( \lambda \geq \frac{1}{2} \), then we can distinguish between \(|+\rangle_x\) and \(|\rangle_x\) by looking at the momentum, as the Gaussian momentum distributions of width \( \frac{1}{2} \) about \( \pm \lambda \) do not essentially overlap. The mean of the distribution moves to \( \pm \lambda \) depending on whether the state is \(|\pm\rangle_x\). If we take \( p_x \geq 0 \), \( p_x \leq 0 \) to correspond to an unsharp measurement of \( \sigma_x \), then the POVM element \( G_+ \), characterizing the unsharp measurement, satisfies

\[
p(p_x \geq 0) = \langle \chi | G_+ | \chi \rangle \\
= \langle \chi | \left( I + a' \sigma_1 \right) | \chi \rangle,
\]

with \( a' = 2F(\lambda a) - 1 \) and \( F(x) = \int_{-\infty}^{x} \frac{e^{-t^2}}{\sqrt{\pi}} dt \). Again as expected, as \( \lambda a \) increases beyond 1, \( F(\lambda a) \) and \( a' \) move closer to 1 and the measurement becomes sharper.

In the other limit, \( \lambda \ll \frac{1}{a} \) distinguishability is lost and this is also reflected in \( a' \) going to zero.

Thus, a necessary condition for an approximate measurement of \( \hat{\sigma} \hat{\sigma} \) to be good can be taken to be its ability to distinguish between the eigenstates of \( \hat{\sigma} \hat{\sigma} \). Distinguishability in turn depends on the two length-scales \( \lambda \), the distance by which the mean of the distribution moves and the width of the distribution being \( \frac{1}{2} \). This is also reflected in the fact that in the limit \( \lambda \gg \frac{1}{a} \), equation (45) becomes a Schmidt decomposition and \( |\psi_f\rangle \) becomes maximally entangled.

### 7.1. Joint measurement

The Arthur–Kelly model that we have considered can be thought to come from a magnetic field \( \vec{B} = 2B_0 \hat{x}\hat{x} + 2B_0 \hat{y}\hat{y} \) \( (\lambda = B_0 \tau \) has been taken to be 1). As mentioned before, the non-zero divergence of this field is not a serious issue. We could also have taken \( \vec{B} = 2B_0 \hat{y}\hat{y} + 2B_0 \hat{z}\hat{z} \) and measured \( p_x, p_y \) in order to have an approximate joint measurement of \( \sigma_x \) and \( \sigma_z \).

After the measurement of momentum, the probability of obtaining \( p_x \) is given by (see equation (27))

\[
p(p_x) = \int_{-\infty}^{\infty} (|e^0|^2 + |f^0|^2 + |g^0|^2) dp_y + 2 \int_{-\infty}^{\infty} e^{0} f^{0*} dp_x \langle \chi | \sigma_x | \chi \rangle
\]

and

\[
p(p_x \geq 0) = \frac{1}{2} + \frac{a'}{2} \langle \chi | \sigma_x | \chi \rangle.
\]

with

\[
a' = \int_{p_x=0}^{+\infty} \int_{p_y=-\infty}^{+\infty} 4(f^0 e^0) dp_x dp_y.
\]
The origin of complementarity between \( \sigma_x \) and \( \sigma_y \) in this model can be understood from the way the effective length-scales governed by the movement of the mean momenta \( \langle p_x \rangle \) and \( \langle p_y \rangle \) change as we make one of the initial momentum wavefunctions sharper or broader. The Ehrenfest theorem applied on Hamiltonian (23) gives

\[
\langle \hat{p}_x \rangle = \langle \sigma_x \rangle, \\
\langle \hat{p}_y \rangle = \langle \sigma_y \rangle.
\]  

(50)

Now as we saw in sections 5.1 and 6.2.1, the effect of interaction (24) on the spin state of the system is that of an asymmetric depolarizing channel that disturbs the \( \langle \sigma_x \rangle \) component of the density matrix while keeping the \( \langle \sigma_y \rangle \) almost intact as the initial \( y \) momentum wavefunction is made much sharper than the \( x \) momentum one. Equation (50) shows that the rate of change of the average momentum in the \( x \) and \( y \) directions is similarly affected. Thus, with the increasing sharpness of the initial \( y \)-momentum wavefunction, the movement of the \( p_y \) mean moves very little, while the \( p_y \) mean moves about –\( \hat{p} \) also shows a minimum. The maximum in the \( x \) component of the initial density matrix is disturbed more and more, while the \( \langle \sigma_x \rangle \) component is almost kept intact.

In order to understand the asymmetric depolarizing action of the interaction on the spin state, we look at the rdm of the system after the interaction once more,

\[
\rho_y^i = \text{Tr}_{1,2}(|\psi_f\rangle\langle\psi_f|) \\
= \text{Tr}_{1,2}(e^{i(x_0\sigma_y + y_0\sigma_y)}|\psi\rangle\langle\psi| \otimes |\chi\rangle\langle\chi| e^{-i(x_0\sigma_y + y_0\sigma_y)}) \\
= \int_{x,y=-\infty}^{+\infty} e^{i(x-\hat{r}y)}|\chi\rangle\langle\chi| e^{-i(x_0\sigma_y + y_0\sigma_y)}|\psi_1(x)|^2|\psi_2(y)|^2 \, dx \, dy \\
= \int_{x,y=-\infty}^{+\infty} e^{i\hat{r}^2(2\pi)}|\chi\rangle\langle\chi| e^{-i\hat{r}^2(2\pi)} \frac{e^{-\frac{1}{2}(\hat{r}_x^2 + \hat{r}_y^2)}}{2\pi ab} \, dx \, dy,
\]  

(51)

with \( r = \sqrt{x^2 + y^2} \) and \( \hat{r} \) denoting the radius and the unit radial vector, respectively, in polar coordinates, while \( x_0 \) and \( y_0 \) are operators. As equation (51) shows, the rdm of the spin part of the system after the interaction is a mixture of rotated states about \( -\hat{r} \) by an angle \( (2\pi a \mod 2\pi \) ). The weight of a rotated state about \( -\hat{r} \) in the mixture is the Gaussian probability density of the initial position of the particle. Thus, as we increase \( b \) keeping \( a \) fixed, the weight of states which are rotated near about the \( y \) axis increases. Hence, the \( \langle \sigma_x \rangle \) component of the initial density matrix is disturbed more and more, while the \( \langle \sigma_y \rangle \) component is almost kept intact.

We earlier saw that the quality of unsharpness \( a' \) increases and shows a maximum in the symmetric case (figure 1(a)). We also saw that the disturbance due to the measurement governed by the plot of \( 1 - 8\epsilon_f \) versus \( a \) also shows a minimum. The maximum in the \( a' \) versus \( a \) curve was also there in [25].

We argue that the maxima are due to the \( 2\pi \) factor in equation (51). For smaller values of \( a \), the rotations are constrained to smaller angles. To understand this, we take a magnetic field of the form \( \vec{B} = -2B_0\hat{r} \). This removes the \( r \) factor in (51). \( a' \) is then seen to change very little with \( a \) (from .027 to .285). The disturbance characterized by \( \epsilon_f \) remains constant at .088 to about four orders of magnitude. The slight increase in \( a' \) is presumably due to the modification of the Ehrenfest equations (50) due to the nonlinearity in the magnetic field.
Figure 4. A plot of the probability of measuring $p_1$ after pre-measurement with $p_1$. $P(p_1)$ represents the probability of $p_1$ for the initial spin state $|+\rangle_x$ and $P_1(p_1)$ represents that for the initial spin state $|-\rangle_x$. (a) A plot of $P(p_1)$ and $P_1(p_1)$ for $a = 5.0, b = 1.0$. (b) A plot of $P(p_1)$ and $P_1(p_1)$ for $a = 5.0, b = 25.0$.

7.2. Entanglement between the detectors and the system

We next consider the entanglement between the system and the two detectors. The detectors are infinite dimensional, while the system of course is two dimensional. Although strictly speaking only correlations between the detectors and the system are required for the measurement on
the detectors to reflect the measurement statistics of the system, one expects entanglement to play a role in this kind of a scenario.

7.2.1. Entanglement between the joint detector system and the system. First, we consider the entanglement between the two detectors (considered as a single system) and the qubit system. As the final state $|\psi_f\rangle$ of the system and the detectors after pre-measurement is pure, this entanglement is simply given by the von Neumann entropy of the reduced density of the system after the detectors are traced out. As shown earlier in section 5.1, this density matrix is given by equation (42).

Taking $x, y = \frac{1}{2}$ and $z = 0$ in equation (42), the von Neumann entropy of the system is given in figure 5.

The basic feature is that when the sharpnesses of the initial momentum wavefunctions are low, the entanglement is relatively low, increasing as the sharpness increases. In both the symmetric and the asymmetric cases, maximal entanglement is not reached.

7.2.2. Entanglement between the detector and the system. In this subsection, we try to see the entanglement in the mixed state of one of the detectors and the system after the other detector is traced out. Now the situation is made difficult to handle by the fact that the momentum of the detector corresponds to an infinite-dimensional configuration degree of freedom. For this reason, we project the state of the detector into two-dimensional momentum subspaces. The average entanglement considering all such outcomes gives a lower bound on the entanglement of the state of the system and the detector as projection, being a local action, cannot increase the entanglement on the average.

In figure 6, we clearly see that the probability in equation (48) becomes increasingly greater than 0.5 as $a$ increases with respect to $b$ (for a $|+\rangle_x$ spin state), signifying an increase in $a'$ as well. Depending on whether the state is $|+\rangle_x$ or $|-\rangle_x$, the probability peaks around $p_x = \pm 1$ (as the initial distribution is symmetric and we have taken the time of interaction to be unity, this follows from equation (50) for almost sharp measurements). Thus, correlation is likely to be the highest between the system and the detector as projection, being a local action, cannot increase the entanglement on the average.

As a measure of entanglement we consider concurrence, which is defined below. We have considered projections in different two-dimensional momentum subspaces. The concurrence for states projected into different subspaces is qualitatively seen to follow the same behavior as the state projected into the $p_1 = \pm 1$ subspace. However, for $a$ much greater than $b$, numerics becomes difficult when we consider projections into subspaces far from $\pm 1$. This is due to the momentum distribution peaking around $p_1 = \pm 1$ (see figure 6) and we expect the entanglement to fall for $a$ much greater than $b$ in these subspaces. Also, as the probability of obtaining such projections falls, we do not expect the concurrence of states projected into momentum subspaces far from $\pm 1$ to contribute much to the average concurrence.

Thus, we take the concurrence of the projected mixed state into the $p_1 = \pm 1$ subspace to be an indicator of how the average concurrence for all possible projections onto two-dimensional subspaces behaves (the first concurrence multiplied by the probability of obtaining $p_1 = \pm 1$ is of course a lower bound for the average concurrence) as we increase the sharpness of one initial detector momentum wavefunction ($a$) keeping the other fixed ($b$) (equations (28) and (29)).

We consider the spin state to be the symmetric state $|\chi\rangle\langle\chi| = \frac{1}{2}(I + \frac{\sigma_x}{\sqrt{2}} + \frac{\sigma_y}{\sqrt{2}})$.
Figure 5. Entanglement between the joint detector system and the qubit system as reflected by the von Neumann entropy of the rdm of the system after the configuration degrees are traced out. (a) A plot of $S(\rho_s^f)$ versus $a = b$. (b) A plot of $S(\rho_s^f)$ versus $b$ for $a = 0.1$.

Let us first consider the entanglement between the first detector and the system. Let 

$$\rho_1^f = \text{Tr}_2(\ket{\psi_f}\bra{\psi_f})$$

be the final state of the system and the detectors after the measurement interaction. Because of 'parity properties', we have

$$\int_{-\infty}^{\infty} e^0(1, p_2)e^0(-1, p_2) \, dp_2 = \int_{-\infty}^{\infty} e^0(1, p_2)e^0(1, p_2) \, dp_2 := E^0(1),$$

$$\int_{-\infty}^{\infty} f^0(1, p_2)f^0(1, p_2) \, dp_2 = - \int_{-\infty}^{\infty} f^0(1, p_2)f^0(-1, p_2) \, dp_2 := F^0(1).$$
Figure 6. Probability distribution for obtaining $p_1$. $P(p_1)$ and $P1(p_1)$ denote, respectively, the probability of obtaining $p_1$ for the initial system states $|+\rangle_x$ and $|-\rangle_x$. (a) A plot of $P(p_1)$, $P1(p_1)$ versus $p_1$ for $a = b = 0.1$. (b) A plot of $P(p_1)$, $P1(p_1)$ versus $p_1$ for $a = 0.5$, $b = 0.1$.

\[
\begin{align*}
\int_{-\infty}^{\infty} \mathcal{g}(1, p_2) \mathcal{g}(1, p_2) \, dp_2 &= \int_{-\infty}^{\infty} \mathcal{g}(1, p_2) \mathcal{g}(-1, p_2) \, dp_2 := G(1), \\
\int_{-\infty}^{\infty} \mathcal{e}(1, p_2) \mathcal{f}(1, p_2) \, dp_2 &= -\int_{-\infty}^{\infty} \mathcal{e}(1, p_2) \mathcal{f}(-1, p_2) \, dp_2 := E_0 F(1).
\end{align*}
\]

Let $P = |p_1 = 1\rangle \langle p_1 = 1| + |p_1 = -1\rangle \langle p_1 = -1|$. We have

\[
\rho^1 = P \rho^1 P = E^0(1) |1\rangle\langle 1| + |1\rangle\langle -1| + | -1\rangle\langle 1| + | -1\rangle\langle -1| \otimes \chi\langle \chi| \\
+ F^0(1) |1\rangle\langle 1| + | -1\rangle\langle -1| + |1\rangle\langle -1| + | -1\rangle\langle 1| \otimes \sigma_1\langle \chi| \sigma_1
\]
Let H be an Arthur–Kelly-like Hamiltonian of the form

\[ H = f(\hat{q}_1, \hat{q}_2) \otimes \sigma_z + g(\hat{q}_1, \hat{q}_2) \otimes \sigma_y \]

which has a symmetry given by \( [A \otimes B, H] = 0 \). Here A and B are unitaries acting respectively on the joint detector space and the spin space. Let \( P_{\rho}(p_1, p_2) \) represent the probability of obtaining the momenta values \( p_1, p_2 \) if the system is initially in the state \( |\chi\rangle \).

**Further**, let the initial joint detector state \( |\psi\rangle \) have the symmetry A so that \( A|\psi\rangle = e^{i\theta}|\psi\rangle \).

Then

\[
\rho_{1} = 0.5(\rho_{2}(1) + F_{0}(1) + G_{0}(1)) + 0.707iE_{0}F_{0}(1),
\]

\[
\rho_{2} = 0.5(\rho_{1}(1) - F_{0}(1) + G_{0}(1)) + 0.707iE_{0}F_{0}(1),
\]

\[
\rho_{3} = 0.353(1 - i)E_{0}(1) + 0.353(1 + i)F_{0}(1) - 0.353(1 + i)G_{0}(1) + E_{0}F_{0}(1),
\]

\[
\rho_{4} = 0.353(1 - i)E_{0}(1) - 0.353(1 + i)F_{0}(1) - 0.353(1 + i)G_{0}(1),
\]

where \( |1\rangle \equiv |p_1 = 1, p_2 = 1\rangle \) and \( |-1\rangle \equiv |p_1 = 1, p_2 = -1\rangle \).

Now for the spin state we have considered, we have the unnormalized density matrix \( \rho^\ast \) in the basis

\[
|1\rangle \equiv |p_1 = 1\rangle \otimes |0\rangle, |2\rangle \equiv |p_1 = 1\rangle \otimes |1\rangle, |3\rangle \equiv |p_1 = -1\rangle \otimes |0\rangle \text{ and } |4\rangle \equiv |p_1 = -1\rangle \otimes |1\rangle \text{ given by}
\]

\[
\rho_{1} = 0.5(\rho_{2}(1) + F_{0}(1) + G_{0}(1)) + 0.707E_{0}F_{0}(1),
\]

\[
\rho_{2} = 0.5(\rho_{1}(1) - F_{0}(1) + G_{0}(1)) + 0.707iE_{0}F_{0}(1),
\]

\[
\rho_{3} = 0.353(1 - i)E_{0}(1) + 0.353(1 + i)F_{0}(1) - 0.353(1 + i)G_{0}(1) + E_{0}F_{0}(1),
\]

\[
\rho_{4} = 0.353(1 - i)E_{0}(1) - 0.353(1 + i)F_{0}(1) - 0.353(1 + i)G_{0}(1),
\]

with \( \text{Tr}(\rho_{1}) = 2(\rho_{2}(1) + F_{0}(1) + G_{0}(1)) \).

After normalization, we apply the Peres–Horodecki PPT criterion to check for entanglement [29]. As a measure of entanglement, we use the concurrence which for a 2 \( \otimes 2 \) density matrix is defined as

\[
C(\rho) = \max \{ 0, \lambda_1 - \lambda_2 - \lambda_3 - \lambda_4 \}, \]

where \( \lambda_i \) denote in decreasing order the square root of the eigenvalues of the non-Hermitian matrix \( \hat{\rho} \), with

\[
\hat{\rho} = (\sigma_z \otimes \sigma_z) \rho (\sigma_z \otimes \sigma_z) \] [30].

In figure 7, we see that the concurrence increases as \( a \) increases with fixed \( b \), i.e. as the \( \sigma_z \) measurement becomes sharper. We expect the entanglement between the system and the detector to behave similarly.

8. Effect of the symmetries of the underlying Hamiltonian on the POVM elements

It was shown in [22] in the context of a Stern–Gerlach-type Hamiltonian how the symmetries of the underlying Hamiltonian, like the one given in equation (23), can be used to perform some particular unsharp measurement. The various cases considered there can be seen to follow from the following lemma.

Consider an Arthur–Kelly-like measurement process.

**Lemma 1.** Let \( H \) be an Arthur–Kelly-like Hamiltonian of the form

\[
H = f(\hat{q}_1, \hat{q}_2) \otimes \sigma_z + g(\hat{q}_1, \hat{q}_2) \otimes \sigma_y,
\]

which has a symmetry given by \( [A \otimes B, H] = 0 \). Here A and B are unitaries acting respectively on the joint detector space and the spin space. Let \( P_{\rho}(p_1, p_2) \) represent the probability of obtaining the momenta values \( p_1, p_2 \) if the system is initially in the state \( |\chi\rangle \).

Further, let the initial joint detector state \( |\psi\rangle \) have the symmetry A so that \( A|\psi\rangle = e^{i\theta}|\psi\rangle \). Then
Figure 7. Concurrence of the rdm of the system and the first detector after projection into the subspace $p_1 = \pm 1$ versus $a$ for $b = 0.1$.

for the initial system state $|\chi\rangle$ and a new basis in the joint detector space $|p'_1, p'_2\rangle = A|p_1, p_2\rangle$, we have

$$P_{|\chi\rangle}(p_1, p_2) = P_{B|\chi\rangle}(p'_1, p'_2).$$  \hspace{1cm} (55)

**Proof.**

$$P_{|\chi\rangle}(p_1, p_2) = \text{Tr}_{1,2}[P_{|p_1, p_2\rangle}\text{Tr}_S(e^{-iH}P_{|\psi\rangle \otimes |\chi\rangle}e^{iH})]$$  \hspace{1cm} (56)

with $P_{|\eta\rangle}$ denoting the projector on $|\eta\rangle$.

Now replacing $e^{-iH}$ by $(A^\dagger \otimes B^\dagger)e^{-iH}(A \otimes B)$ in (56) and using $A|\psi\rangle = e^{i\phi}|\psi\rangle$, we have

$$P_{|\chi\rangle}(p_1, p_2) = \text{Tr}_{1,2}[P_{|p_1, p_2\rangle}\text{Tr}_S((A^\dagger \otimes B^\dagger) e^{-iH}(|\psi\rangle \langle\psi| \otimes B|\chi\rangle \langle\chi| B^\dagger) e^{iH} (A \otimes B))].$$  \hspace{1cm} (57)

Let

$$e^{-iH}(|\psi\rangle \langle\psi| \otimes B|\chi\rangle \langle\chi| B^\dagger) e^{iH} = e^{-iH}(P_{|\psi\rangle \otimes B|\chi\rangle}) e^{iH} = \sum_j C_j \otimes D_j,$$  \hspace{1cm} (58)

where the index $j$ runs over a countable set.

$C_j$s and $D_j$s are operators acting respectively on the joint Hilbert space of the two detectors and the system Hilbert space.

Using $B^\dagger B = I$, we have from equations (56)–(58)

$$\text{Tr}_S(e^{-iH}P_{|\psi\rangle \otimes |\chi\rangle}e^{iH}) = \sum_j \text{Tr}_S(D_j)(A^\dagger C_j A).$$  \hspace{1cm} (59)
\[ P(\rho) = \sum_j \text{Tr}_1,2(\rho p_j^{(1)} p_2^{(2)} |p_j^{(1)} p_2^{(2)} \text{Tr}_2(D_j) (A^H C_j A)] \]
\[ = \sum_j \rho p_j^{(1)} p_2^{(2)} |p_j^{(1)} p_2^{(2)} \text{Tr}_2(D_j) \]
\[ = \sum_j \rho p_j^{(1)} p_2^{(2)} |C_j p_j^{(1)} p_2^{(2)} \text{Tr}_2(D_j). \quad (60) \]

Again,
\[ P_{\rho \sigma}(\rho_j^{(1)}, p_2^{(2)}) = \text{Tr}_1,2(\rho p_j^{(1)} p_2^{(2)} |p_j^{(1)} p_2^{(2)} \text{Tr}_2(e^{-iHt} \rho p_j^{(1)} p_2^{(2)} e^{iHt})] \]
\[ = \text{Tr}_1,2 \left[ \rho p_j^{(1)} p_2^{(2)} |p_j^{(1)} p_2^{(2)} \text{Tr}_2 \left( \sum_j C_j \otimes D_j \right) \right] \text{ (from equation (58))} \]
\[ = \sum_j \rho p_j^{(1)} p_2^{(2)} |C_j p_j^{(1)} p_2^{(2)} \text{Tr}_2(D_j). \quad (61) \]

9. Approximate joint measurement in arbitrary directions

Let \( I_i (i = 1, 2) \) denote the operation of reflection in the detector space about the \( q_i \)-axis (where \( q_1 = x \) and \( q_2 = y \)). Consider a Hamiltonian of the entire system (i.e. two detectors plus the qubit) which satisfies \([I_i \otimes \sigma_s, H] = 0 \) and an initial state \( \psi \) of the two detectors jointly which satisfies \( I_1 |\psi\rangle = |\psi\rangle \). Then, as a corollary of lemma 1, we have for the probability \( P(p_1 \geq 0) \),
\[ P(p_1 \geq 0) = \langle \chi | E(p_1 \geq 0) | \chi \rangle = \langle \chi | \frac{1}{2}(\alpha I + \beta \sigma_s) | \chi \rangle, \]
with constant \( \alpha \) and \( \beta \).

Proof: From equation (55), we have
\[ P(\rho) = \rho \sigma_1 \rho \sigma_2. \]
So \( \langle \chi | E(p_1 \geq 0) | \chi \rangle = \langle \chi | \sigma_1 E(p_1 \geq 0) \sigma_2 | \chi \rangle, \)
i.e. \( E(p_1, p_2) = \sigma_1 E(p_1, -p_2) \sigma_2, \)
and integrating over \( p_2 \), \( [E(p_1), \sigma_s] = 0. \)
So we can write \( E(p_1) = \frac{1}{2}(\alpha(p_1) I + \beta(p_1) \sigma_s), \)
where \( \alpha(p_1) \) and \( \beta(p_1) \) are real numbers. Hence, we have
\[ E(p_1 \geq 0) = \frac{1}{2}(\alpha I + \beta \sigma_s), \quad (62) \]
with the constants \( \alpha = \int_{p_1=0}^{\infty} \alpha(p_1) dp_1 \) and \( \beta = \int_{p_1=0}^{\infty} \beta(p_1) dp_1. \)

As the Hamiltonian \( H \) of equation (23) satisfies \([I_i \otimes \sigma_s, H] = 0 \) and as the initial joint detector state \( |\psi\rangle = |\psi_1 \rangle \otimes |\psi_2 \rangle \), where \( |\psi_1 \rangle \) and \( |\psi_2 \rangle \) are given respectively by equations (28) and (29), satisfies both \( I_1 |\psi_1 \rangle = |\psi_1 \rangle \) and \( I_2 |\psi_2 \rangle = |\psi_2 \rangle \), it is clear from equation (62) how the approximate joint measurement of \( \sigma_s \) and \( \sigma_s \) with the marginals given by equation (36) arises out of the measurement of \( p_1 \) and \( p_2 \).

Further consider an \( H \) which in addition to having the symmetry mentioned in the beginning of this section has a further rotational symmetry: \([R(\theta) \otimes S_c(\theta), H] = 0 \), where \( R(\theta) \) and \( S_c(\theta) \), respectively, denote rotation by an angle \( \theta \) in the detector Hilbert space (i.e.
the Hilbert space operation corresponding to rotation in the \( q_1, q_2 \) plane) and the spin space (about the \( z \)-axis). From equation (55), we have

\[
\kappa_e(p_1, p_2) = \kappa_e(p_1', p_2') \tag{63}
\]

with \((p_1', p_2')^\top = R(\theta)(p_1, p_2)^\top\).

Integrating both sides of equation (63) over the region \( p_1 \geq 0, -\infty \leq p_2 \leq \infty \), we have

\[
\kappa_e(p_1, p_2) = \kappa_e(p_1', p_2'),
\]
i.e. \( \kappa_e(p_1, p_2) = \kappa_e(p_1', p_2') \leq \frac{\sqrt{2}}{4} \).

Now, using equation (62), we have from equation (64)

\[
\kappa_e(p_1', p_2') = \langle \chi | \sigma_1' \rangle \langle \sigma_1' | \chi \rangle,
\]
i.e. \( \langle \chi | E(p_1', p_2') \rangle = \langle \chi | \sigma_1' \rangle \langle \sigma_1' | \chi \rangle \),

where \( \sigma_1' = \hat{\sigma} \cos(\theta) + \hat{\sigma} \sin(\theta) \), with \( \theta \) being the polar angle of a point in the \( x-z \) plane. In the last but one line, we have also used \( \sigma_z = | \chi \rangle \langle \chi | \).

The Hamiltonian in equation (23) satisfies both the above-mentioned reflection and the rotational symmetry properties. Choosing the detector state to be a symmetric Gaussian, i.e. \( b = a \) in equation (29), we have the required rotational invariance for all angles. Hence, we get an approximate joint measurement of spin in any direction by measuring the detector momentum in that direction.

9.1. POVM elements

Consider the approximate joint measurement of \( \sigma_z \) and \( \sigma_.\hat{\sigma}_z \). We have already shown that the marginal probabilities of the joint measurement will be given by

\[
p(p_1', p_2') = \langle \chi | \frac{1}{2} (1 + a' \sigma_z) | \chi \rangle,
\]
with \( p_1' \) and \( p_2' \) denoting momenta in the \( \hat{x} \) and \( \hat{\sigma}_z \) directions, respectively (\( \hat{x} \) is along the positive \( x \)-axis and \( \hat{\sigma}_z \) along \( \hat{x} \cos(\theta) + \hat{y} \sin(\theta) \) in the momentum plane \((p_1, p_2)\)).

9.1.1. Angular dependence of \( e^0, f^0 \) and \( g^0 \). The rotational invariance of the initial state allows one to extract the polar angular dependence of \( e^0, f^0 \) and \( g^0 \) (introduced in section 5) in the detector space \((p_1, p_2)\) \((p_1 = p \cos(\theta_p), p_2 = p \sin(\theta_p))\):

\[
e^0(p, \theta_p) = \frac{1}{2\pi|q^2|} \int_0^{2\pi} e^{-ipq^2 \cos(\theta_p)} e^{-ipq^2 \sin(\theta_p)} \cos(\sqrt{q^2 + q^2}) e^{q^2/2} dq_1 dq_2. \tag{68}
\]

Putting \( q_1 = r \cos(\theta) \) and \( q_2 = r \sin(\theta) \), we have

\[
e^0(p, \theta_p) = C \int_0^{2\pi} \int_0^{\infty} e^{-ipr \cos(\theta_p)} r e^{r^2/2} dr d\theta = C \int_0^{\infty} \int_0^{2\pi} e^{-ipr \cos(\theta_p)} d\theta_1 dr,
\]

with \( C \) being the constant part, \( g_1(r) = r \cos(r) e^{r^2/2} \). Now taking \( \theta = \theta_p = \theta' \), the theta integral in equation (69) becomes

\[
\int_0^{2\pi} e^{-ipr \cos(\theta_p)} d\theta = \int_{-\pi}^{\pi} e^{-ipr \cos(\theta)} d\theta' \tag{70}
\]

where \( \theta \) and \( \theta' \) are the polar angles in the momentum plane \((p_1, p_2)\) and \((p_1', p_2')\).
The last equation follows from the fact that the integral of a periodic function over its period is independent of the limits of integration. Thus, $e^0$ is only a function of $p$:

$$e^0 = e_1(p).$$

(71)

Again, similar transformation for $f^0$ yields

$$f^0(p, \theta_p) = C' \int_0^{2\pi} \int_{-\theta_p}^{2\pi-\theta_p} e^{-ipq\cos(\theta')}(\cos(\theta)\sin(\theta) e^{-ir}) d\theta,$$

(72)

with $C'$ being the constant part. Using $\theta - \theta_p = \theta'$, the theta integral in (72) becomes

$$\int_{-\theta_p}^{2\pi-\theta_p} e^{-ipq\cos(\theta')}(\cos(\theta')\cos(\theta_p)\sin(\theta)\sin(\theta_p) d\theta'.

Now,

$$-\int_{-\theta_p}^{2\pi-\theta_p} e^{-ipq\cos(\theta')}\sin(\theta') d\theta' = -\int_0^{2\pi} e^{-ipq\cos(\theta')}\sin(\theta') d\theta'.$$

(73)

Taking $\cos(\theta') = x$, the above equation becomes $\int_{-1}^1 e^{-ipq} dx + \int_{-1}^1 e^{-ipq} dx = 0$. Therefore,

$$f^0(p, \theta_p) = f_1(p) \cos(\theta_p).$$

(74)

Proceeding exactly similarly, one can show that

$$g^0(p, \theta_p) = f_1(p) \sin(\theta_p).$$

(75)

### 9.1.2. Joint measurement probabilities.

Choosing the $p'_1$ and $p'_2$ axes according to figure 8, the joint probability is given by

$$p(p'_1 \geq 0, p'_2 \geq 0) = \int_{p'_1, p'_2=0}^{\text{GOD}} p(p_1, p_2) dp_1 dp_2 + \int_{p'_1, p'_2=0}^{\text{DOE}} p(p_1, p_2) dp_1 dp_2,$$

(76)

with $p(p_1, p_2)$ given by (27), where the last integral represents the integral over the region $\text{DOE}$ in figure 8, given in polar coordinates by $\{(p, \theta') : 0 \leq p \leq \infty, -(\pi/2) \leq \theta' \leq 0\}$.

So, using equation (30) in (76), we have

$$p(p'_1 \geq 0, p'_2 \geq 0) = \langle \chi | \left[ \frac{1}{4} + a \frac{\sigma_x}{4} + a \frac{\sigma_y}{4} \right] | \chi \rangle + \int_{\text{DOE}} p(p_1, p_2) dp_1 dp_2.$$

(77)

Now noting the $\theta$ dependence of $e^0, f^0$ and $g^0$ from (71), (74) and (75), respectively, and using it in (27), we have

$$\int_{\text{DOE}} p(p_1, p_2) dp_1 dp_2$$

$$= \langle \chi | \left( \int_0^{\infty} \int_0^{-(\pi/2)} \left[ (|e_1(p)|^2 + |f_1(p)|^2 \cos^2(\theta) + |f_1(p)|^2 \sin^2(\theta)) 1, \right. \\
+ 2e_1(p)f_1(p) \cos(\theta) \sigma_1 + 2e_1(p)f_1(p) \sin(\theta) \sigma_2 \right] dp d\theta) | \chi \rangle.$$

(78)

Now, from equation (31), we have

$$\int_0^{\infty} \int_0^{-(\pi/2)} (e_1^2 + f_1^2) 2\pi p dp = 1.$$

(79)
We also have from the definition of $a'$ (equation (33))

$$\int_0^{\infty} e_{f_1} p \, dp = \frac{a'}{8}. \quad (80)$$

Hence,

\begin{align*}
p(p'_1 \geq 0, p'_2 \geq 0) &= \langle \chi | \left( \frac{1}{2} - \frac{\theta}{2\pi} \right) 1_s + \frac{a'}{4} (1 + \cos(\theta))\sigma_x + \frac{a'}{4} \sin(\theta)\sigma_y \right| \chi \rangle \quad (81) \\
p(p'_1 \leq 0, p'_2 \geq 0) &= \langle \chi | \left( \frac{\theta}{2\pi} 1_s + \frac{a'}{4} (1 - \cos(\theta))\sigma_x + \frac{a'}{4} \sin(\theta)\sigma_y \right| \chi \rangle, \quad (82) \\
p(p'_1 \geq 0, p'_2 \leq 0) &= \langle \chi | \left( \frac{-\theta}{2\pi} 1_s + \frac{a'}{4} (-\cos(\theta) + 1)\sigma_x - \frac{a'}{4} \sin(\theta)\sigma_y \right| \chi \rangle, \quad (83) \\
p(p'_1 \leq 0, p'_2 \leq 0) &= \langle \chi | \left( \frac{1}{2} - \frac{-\theta}{2\pi} \right) 1_s - \frac{a'}{4} (1 + \cos(\theta))\sigma_x - \frac{a'}{4} \sin(\theta)\sigma_y \right| \chi \rangle. \quad (84)
\end{align*}

At $\theta = \frac{\pi}{2}$, we get back the POVM elements for the orthogonal case as derived in section 5. For $\theta \to 0$, i.e. the joint measurement along two almost same directions, the probabilities $p(p'_1 \leq 0, p'_2 \geq 0)$ and $p(p'_1 \geq 0, p'_2 \leq 0)$ vanish, while $p(p'_1 \geq 0, p'_2 \geq 0)$ and $p(p'_1 \leq 0, p'_2 \leq 0)$ give the probabilities for the single unsharp measurement along $\hat{x}$, as expected.

**Remark:** The joint measurement uncertainty relation is given by equation (8),

$$||\vec{a}_1 + \vec{a}_2|| + ||\vec{a}_1 - \vec{a}_2|| \leq 2. \quad (85)$$
If $\theta$ is the angle between $\vec{a}_1$ and $\vec{a}_2^*$, in the case $|\vec{a}_1| = |\vec{a}_2^*|$ (like above), one has
\[
|\vec{a}_1| \leq \frac{1}{\sin \left(\frac{\theta}{2}\right) + \cos \left(\frac{\theta}{2}\right)}.
\] (86)

The denominator in equation (86) is positive in $\theta \in [0, \pi]$ with the maximum value $\sqrt{2}$ at $\theta = \pi/2$. This is also our bound in the approximate joint measurement we have implemented, because for the symmetric initial detector state case (see section 6.1), the approximate joint measurement occurs in orthogonal directions as well. In fact we just showed that measuring momenta in any two directions yields the approximate joint measurement of spin in those two directions. We saw earlier that in our case (see section 6), $|\vec{a}_1|$ is able to reach about 0.628 (see figure 1(a)). Thus, it is possible that one can have a different scheme with the same but better quality of the unsharp measurement in both the directions. For example, for $\theta = \pi/4$ the bound by the joint measurement inequality (86) is about 0.765.

10. Spin direction fidelities

In this section, we consider the spin direction fidelities, defined in [21]. We consider a type-1 measurement as defined by the author in [21] in which the pointer observables are taken to be the commuting components of some unit vector $\hat{n}$. The measurement scheme considered by us in section 5 yields the momentum values $p_1, p_2$ (which can also be considered as the components of $\vec{\rho}$). Alternatively, in polar coordinates we can look at the magnitude of the momentum $\vec{\rho}$ and its angle $\theta$ with the $x$-axis in the momentum plane. This angle uniquely fixes the direction of the normalized momentum meter $\hat{\rho}$, and this is what we take for $\hat{n}$ in this section.

The joint unsharp measurement on the system through the measurement on the commuting observables after the meters have interacted with the system yields the POVM described by equations (32) and (35). The effect of the measurement on the system (a completely positive map acting on the system density matrix) can be described by Krauss operators for the measurement $T(\vec{\rho})$. They are defined as follows by considering the average density matrix of the system $\rho_f$ after the measurement on the meters, considering all possible outcomes:
\[
\rho_f = \int dp d\theta T(p, \theta) |\chi\rangle |\chi\rangle T^\dagger(p, \theta).
\] (87)

The probability of registering the outcome at an angle $\theta$ in the $(p_1, p_2)$ plane is given by $p(\theta) = \langle \chi | E(\theta) | \chi \rangle$, with the ‘angle’ POVM
\[
E(\theta) = \int_0^\infty T^\dagger(p, \theta) T(p, \theta) \rho_f dp.
\] (88)

The fidelities were verified to be expressible in terms of $E$ and $T$ as (see [21])
\[
\eta_s = \inf_{|\chi\rangle \in H_{sys}} \langle \chi | \left[ \int_0^{2\pi} \frac{1}{2} (E(\theta) S^\dagger \vec{\rho} + S \vec{\rho} (E(\theta)) ) d\theta \right] |\chi\rangle,
\] (89)

\[
\eta_f = \inf_{|\chi\rangle \in H_{sys}} \langle \chi | \left[ \int T^\dagger (\vec{n} \cdot \vec{S}) T \rho_f dp d\theta \right] |\chi\rangle,
\] (90)

\[
\eta_d = \inf_{|\chi\rangle \in H_{sys}} \langle \chi | \left[ \int \sum_{i=1}^3 (\chi | \frac{1}{2} (T^\dagger S_i T + S_i T^\dagger T) \rho_f dp d\theta \right] |\chi\rangle,
\] (91)

with $T$ in our case given by
\[
T(\vec{\rho}) = \langle \vec{\rho} | e | \psi \rangle 1_x + \langle \vec{\rho} | f | \psi \rangle \sigma_x + \langle \vec{\rho} | g | \psi \rangle \sigma_y = e^0 1_x + f^0 \sigma_x + g^0 \sigma_y.
\] (92)
From (88), we have
\[
E(\theta) = \int_{0}^{+\infty} (|e^{0}|^2 + |f^{0}|^2 + |g^{0}|^2)p \, dp + \int_{0}^{+\infty} 2 \text{Re}(e^{0}f^{0*})p \, dp \sigma_{x} + \int_{0}^{+\infty} 2 \text{Re}(e^{0}g^{0*})p \, dp \sigma_{y}
= a(\theta)1_{x} + b(\theta)\sigma_{x} + c(\theta)\sigma_{y} \text{(say)}. \tag{93}
\]
Hence,
\[
\frac{1}{2}(E(\hat{n})\hat{n}_{S} + \hat{n}_{S}E(\hat{n})) = \frac{1}{2}[(b(\theta)\cos(\theta) + c(\theta)\sin(\theta))1_{x} + a(\theta)\cos(\theta)\sigma_{x} + a(\theta)\sin(\theta)\sigma_{y}]. \tag{94}
\]
So from equation (89),
\[
\eta_{i} = \inf_{(\chi)\in\mathcal{U}_{\chi}} \int_{0}^{2\pi} d\theta \left[ \frac{1}{2} (b(\theta)\cos(\theta) + c(\theta)\sin(\theta)) \\
+ \frac{1}{2} a(\theta)\cos(\theta)\langle \chi | \sigma_{x} | \chi \rangle + \frac{1}{2} a(\theta)\sin(\theta)\langle \chi | \sigma_{y} | \chi \rangle \right]
= \int_{0}^{\infty} \int_{0}^{2\pi} e^{0}f^{0*} \cos(\theta) + e^{0}g^{0*} \sin(\theta) \, dp \, d\theta \\
+ \int_{0}^{\infty} \int_{0}^{2\pi} (|e^{0}|^2 + |f^{0}|^2 + |g^{0}|^2) \cos(\theta) \, dp \, d\theta \langle \chi | \sigma_{x} | \chi \rangle \\
+ \int_{0}^{\infty} \int_{0}^{2\pi} (|e^{0}|^2 + |f^{0}|^2 + |g^{0}|^2) \sin(\theta) \, dp \, d\theta \langle \chi | \sigma_{y} | \chi \rangle. \tag{95}
\]
Changing back to the Cartesian coordinates, we have
\[
\eta_{i} = \int_{-\infty}^{\infty} e^{0}f^{0*} \frac{p_{1}}{(p_{1}^{2} + p_{2}^{2})^{\frac{3}{2}}} \, dp_{1} \, dp_{2} + \int_{-\infty}^{\infty} e^{0}g^{0*} \frac{p_{2}}{(p_{1}^{2} + p_{2}^{2})^{\frac{3}{2}}} \, dp_{1} \, dp_{2}. \tag{96}
\]
The last two terms on the rhs of equation (95) vanish because the third term is odd in \(p_{1}\) and the fourth one is odd in \(p_{2}\).
Now using (80), we have
\[
\eta_{i} = \int e^{0}f^{0} \cos(\theta) \, dp \, d\theta + \int e^{0}g^{0} \sin(\theta) \, dp \, d\theta \\
= \int e_{1}f_{1} \, dp \int_{\theta=0}^{2\pi} \cos^{2}(\theta) \, d\theta + \int e_{1}f_{1} \, dp \int_{\theta=0}^{2\pi} \sin^{2}(\theta) \, d\theta \\
= 2\pi \int e_{1}f_{1} \, dp. \tag{97}
\]
Thus,
\[
\eta_{i} = \frac{\pi}{4} a'. \tag{98}
\]

**Bound on** \(a'\): It was shown in [21] that \(\eta_{i}\) is bounded by the value \(s = \frac{1}{2}\). From here we obtain the bound on \(a'\) to be about 0.64 which is almost exactly what is obtained in the simulation of \(a'\) (0.6292 in figure 1(a)). The joint measurement uncertainty relation (43) allows \(a'\) to go till 0.707. This is because equation (43) refers to the most general approximate joint measurement in orthogonal directions without reference to any Arthur–Kelly kind of an implementation.
Also equation (98) shows that the error of retrodiction (given in equation (20)) falls as the measurement becomes sharp.

A similar calculation shows that  
\[ \eta_f = \inf_{|\chi\rangle \in H_{sys}} \int p \, d\theta \, dp \langle \chi | \left[ \cos(\theta) \left( |e_0|^2 + |f_0|^2 - |g_0|^2 \right) + 2 \text{Re}(f_0^* g_0) \sin(\theta) \sigma_x + \text{Re}(f_0^* g_0^*) \cos(\theta) \sigma_y \right] |\chi\rangle \]  
(99)

and  
\[ \eta_d = \inf_{|\chi\rangle \in H_{sys}} \int p \, d\theta \, dp \langle \chi | \left[ \left| \frac{3}{4} |e_0|^2 + |f_0|^2 + |g_0|^2 \right| L_x + \frac{\text{Re}(f_0^* g_0^*)}{2} \sigma_x + \frac{\text{Re}(f_0^* g_0^*)}{2} \sigma_y \right] |\chi\rangle \]  
(100)

Thus, as in the case of the original Arthur–Kelly model, the fidelities are independent of the initial system state [19].

Again, because of the parity of various terms present there, \( \eta_f \) turns out to be the same as \( \eta_i \).

From the probability normalization condition given in equation (31) and parities, we have \( \eta_d = \frac{3}{4}, \) independent of the initial apparatus state and \( \eta_i, \eta_f \).

This shows that any error-disturbance relationship between the error of retrodiction/prediction and the error of disturbance does not hold for all choices of the direction pointer observable.

11. Approximate joint measurement for three qubit observables

An approximate joint measurement observable for three unsharp qubit observables is defined by extending conditions given in (1) in the obvious way. Unlike the two-observable case, no necessary–sufficient condition for the approximate joint measurement on the marginal effect parameters in \( \mathcal{R}^4 \) (like the ones derived for example in [9]) is known. No measurement uncertainty relation like equation (9), but which is stronger than the relations applied pairwise, is known. Here, we derive a necessary condition on the approximate joint measurement of three qubit observables which yields the familiar necessary–sufficient condition known in the literature for the case of mutually orthogonal unbiased observables (to be discussed below). We show that it holds even when the observables are biased. Also when one of the measurements is trivial, i.e. the corresponding marginal effect is \( \frac{I}{2} \), representing equiprobable guesses of the values of the observable, our condition reproduces equation (8).

The eight joint effects \( G_{+++}, G_{++-}, \ldots, G_{---} \) have to satisfy six marginality conditions like  
\[ \Upsilon_+^1 = G_{+++} + G_{+--} + G_{++-} + G_{---} \]  
(101)

A joint measurement of three observables gives rise to three two-observable joint measurements. Let \( G_{12}^{++} \) denote the joint measurement marginal effect corresponding to the outcome (++) in directions 1 and 2. Then seven of the effects (\( G_{+++}, G_{++-}, \) etc) can be written in terms of the three single-observable marginal effects, three pairwise joint measurement marginal effects and \( G_{+++} \), as follows:

\[ G_{++-} = G_{12}^{++} - G_{+++} \]  
(102)

\[ G_{+--} = G_{23}^{-+} - G_{++-} + G_{+++} \]  
(103)

\[ G_{---} = G_{13}^{--} + G_{23}^{-+} + G_{+++} - \Upsilon_+^2 \]  
(104)
\[ G_{++} = \gamma^2_+ - G^3_{++} - G_{+++} , \]  
\[ G_{++} = G^3_{++} - G_{+++} , \]  
\[ G_{+} = G^3_{+} + G^2_{+} - G_{+++} , \]  
\[ G_{++} = G^2_{++} - G_{+++} . \]  

It was shown in [9] that any set of effects for the joint measurement of two observables (defined in equation (1)) can be written in the form
\[ G_{sgn(a)sgn(b)}(Z, \bar{Z}) = \frac{(1 + ax + by + abZ)I + (ab\bar{z} + a\bar{n} + b\bar{t})\sigma}{4} , \]  
with \( a, b \in \{1, -1\} \), all other scalars in \( R \), all vectors in \( R^3 \) and the two single-qubit marginal effects given by \( \bar{y}_1(x, \bar{m}) = \frac{1}{2}((1 \pm x)I \pm m\sigma) \) and \( \bar{y}_2(y, \bar{m}) = \frac{1}{2}((1 \pm y)I \pm n\sigma) \) with the positivity constraint being \(|x| + m \leq 1 \) and \(|y| + n \leq 1 \). Condition (109) is true because for the two-observable case all the effects can be expressed in terms of one effect (say \( G_{++} \)) and the marginals [6]. The freedom in \((Z, \bar{Z})\) suffices to specify an arbitrary \( G_{+++} \).

It then follows from equation (109) that the set of joint effects for the three observable case is of the form
\[ G_{sgn(\alpha)sgn(\beta)sgn(\gamma)}([Z], \bar{Z}) = ((1 + ax + by + cz + abZ)I + (ab\bar{z} + c\bar{m} + b\bar{n})\sigma) / 8 , \]  
with \( a, b, c \in \{1, -1\} \), all other scalars in \( R \), all other vectors in \( R^3 \) as before, the two-marginals \( G_{12} = G(Z_1, \bar{z}_1), G_{23} = G(Z_2, \bar{z}_2), G_{13} = G(Z_3, \bar{z}_3) \) given by equation (109), and the one-marginals being \( \bar{y}_1(s, \bar{l}) = \frac{1}{2}((1 \pm x)I \pm \bar{l}\sigma) \), \( \bar{y}_2(s, \bar{m}) = \frac{1}{2}((1 \pm y)I \pm \bar{m}\sigma) \), \( \bar{y}_3(s, \bar{t}) = \frac{1}{2}((1 \pm x)I \pm \bar{t}\sigma) \). Given the positivity of the one-marginals, the positivity of the eight joint effects places restrictions on the marginal effect parameters which are interpreted as measurement uncertainty relations in contrast to the usual state uncertainty relations.

### 11.1. Necessary condition

From equation (110), \( G_{+++} \geq 0 \) gives
\[ |l^2 + m^2 + n^2 + z^2 + z_1 + z_2 + z_3| \leq |1 + x + y + z + Z_1 + Z_2 + Z_3 + Z_4| , \]  
while \( G_{+++} \geq 0 \) gives
\[ |l^2 + m^2 + n^2 + z^2 + z_1 + z_2 + z_3 + z_4| \leq |1 - x - y - z + Z_1 + Z_2 + Z_3 + Z_4| . \]  

Equations (111) and (112) together can be interpreted as the collection of all points \((z_1, z_2, z_3, z_4) \in R^3 \), each of which lies within two spheres in \( R^3 \) with their centers at \((\bar{l}^2 + \bar{m}^2 + \bar{n}^2 + \bar{z}^2) \) and \((\bar{t}^2 + \bar{t}^2 + \bar{n}^2 + \bar{z}^2) \) and radii \(|1 - x - y - z + Z_1 + Z_2 + Z_3 + Z_4| \) and \(|1 + x + y + z + Z_1 + Z_2 + Z_3 + Z_4| \), respectively. Thus, the distance between their centers should be less than or equal to the sum of their radii, implying
\[ |l^2 + m^2 + n^2 + z_1 + z_2 + z_3| \leq |1 + Z_1 + Z_2 + Z_3| . \]  

Similarly, by consideration of the other complementary pairs \( G_{pq} \) and \( G_{(-p)(-q)(-r)} \) for \( p, q, r = \pm \), we obtain three other equations,
\[ |l^2 + m^2 + n^2 + z_2| + |l^2 + m^2 + n^2 + z_3| \leq |1 + Z_1 + Z_2 + Z_3| , \]  
\[ |l^2 + m^2 + n^2 + z_1 + z_2| \leq |1 + Z_1 + Z_2 + Z_3| , \]  
\[ |l^2 + m^2 + n^2 + z_1 + z_3| \leq |1 + Z_1 + Z_2 + Z_3| . \]  

This completes the proof of the necessary condition. 

\[ \]
11.2. Geometric interpretation

The above inequalities (113)–(116) imply that we have a cuboid with edges 2⃗{l}, 2⃗{m} and 2⃗{n} with the origin at the center of this cuboid and the solid spheres about the points A ≡ (−⃗{l}−⃗{m}−⃗{n}), B ≡ (⃗{l}+⃗{m}−⃗{n}), C ≡ (−⃗{l}+⃗{m}+⃗{n}) and D ≡ (⃗{l}−⃗{m}+⃗{n}) have to intersect so that the sum of the radii of the spheres is 4 (see figure 9). That is, we necessarily have

\[ ||−⃗{l}−⃗{m}−⃗{n}−⃗{z}_4|| + ||⃗{l}+⃗{m}−⃗{n}−⃗{z}_4|| + ||⃗{l}−⃗{m}+⃗{n}−⃗{z}_4|| + ||−⃗{l}+⃗{m}+⃗{n}−⃗{z}_4|| \leq 4, \]  

(117)

Now, for any given set of points in \( \mathbb{R}^3 \), the Fermat–Toricelli (F–T) point of the set is defined to be the point which minimizes the sum of distances from that point in \( \mathbb{R}^3 \) to the points of the set. The F–T point has been studied for quite long and its properties for any set of four non-coplanar points are well known [28]. For example, it is known that the F–T point either belongs to the set itself or else it is the point at which the gradient of the sum of distances vanishes. Thus, choosing ⃗{z}_4 to be the F–T point of the set of points A, B, C, D, one necessarily has (according to the definition of the F–T point)

\[ ||−⃗{l}−⃗{m}−⃗{n}−⃗{z}_{4_F}|| + ||⃗{l}+⃗{m}−⃗{n}−⃗{z}_{4_F}|| + ||⃗{l}−⃗{m}+⃗{n}−⃗{z}_{4_F}|| + ||−⃗{l}+⃗{m}+⃗{n}−⃗{z}_{4_F}|| \leq 4, \]  

(118)
11.3. Sufficiency

In order to evaluate the F–T point for a particular set of points A, B, C and D defined above, we will use the following theorem from [28] (theorem 2.1), due to Lorentz Lindelf and Sturm.

**Theorem 2.** Suppose that $z_{A}^{4}$ is the F–T point for a set of points $S_{n} \equiv \{ \vec{a}_{i} \in \mathcal{R}^{3} : i = 1, 2, \ldots, n \}$. Then either $z_{A}^{4}$ belongs to the set $S_{n}$ or $z_{A}^{4} \notin S_{n}$.

(i) If $z_{A}^{4} \in S_{n}$, then for $z_{A}^{4} = \vec{a}_{j}$, for some $j \in \{ 1, 2, \ldots, n \}$, $\| \sum_{i=1(i \neq j)}^{n} (\vec{a}_{i} - \vec{a}_{j}) \| \leq 1$.

(ii) If $z_{A}^{4} \notin S_{n}$, then $z_{A}^{4}$ is the point at which $\| \sum_{i}^{n} (\vec{a}_{i} - \vec{z}_{A}^{4}) \| = 0$.

Thus, for the first case, the resultant of unit vectors to the F–T point from other points of the set has magnitude less than 1. In the second case, the unit vectors from the F–T point to the points of the set add up to the null vector. The condition for the second case is also the condition for the gradient of the function representing the sum of distances from $z_{A}^{4}$ to $\vec{a}_{j}$, for $i = 1, 2, \ldots, n$, to vanish at $z_{A}^{4}$.

11.3.1. $\vec{l}, \vec{m}, \vec{n}$ mutually orthogonal. Suppose that we are considering the case of the approximate joint measurement in three orthogonal directions, $\vec{l}, \vec{m}$ and $\vec{n}$. In this case, we have $\| - \vec{l} + \vec{m} - \vec{n} \| = \| \vec{l} + \vec{m} + \vec{n} \| = \| - \vec{l} + \vec{m} + \vec{n} \|$. Hence, at the origin ($z_{A}^{4} = 0$) of the cuboid (see figure 9), we have

$$
\sum_{i} (\vec{a}_{i} - \vec{z}_{A}^{4}) \| \vec{a}_{i} - \vec{z}_{A}^{4} \| = (\vec{l} - \vec{m} + \vec{n} - \vec{l} + \vec{m} - \vec{n})/((l^{2} + m^{2} + n^{2}) = 0.
$$

Hence, equation (118) gives

$$
l^{2} + m^{2} + n^{2} \leq 1.
$$

(120)

This condition was shown to be sufficient in [4]. The necessity of this condition was shown in [13] assuming unbiasedness of the marginals (i.e. $x = y = z = 0$ in equation (110)) by considering measurements by two parties on a singlet state. We have shown above that this is true for biased cases as well. Sufficiency of the above condition (120) shows that condition (118) is sufficient for the case of the approximate joint measurement in three orthogonal directions. Also note that condition (120) is stronger than pairwise conditions for the two-observable joint measurement in orthogonal directions like equation (43) which when added produces the bound of 1.5 on the lhs of equation (120).

11.3.2. Reduction to the two-observable inequality. Suppose that our joint measurement scheme is such that some approximate joint measurement on two observables is performed, while the value of the third observable is guessed with the probability of ‘+’ being $\frac{1+z}{2}$ and that of ‘−’ being $\frac{1-z}{2}$. This will correspond to $\vec{n} = 0$, i.e. the corresponding marginal $\Upsilon_{\pm}(z, \vec{n}) = \frac{(1 \pm z)l}{2}$. In this case, the points $A \equiv (-\vec{l} - \vec{m}), B \equiv (\vec{l} + \vec{m}), C \equiv (\vec{l} - \vec{m})$ and $D \equiv (-\vec{l} + \vec{m})$ form a parallelogram of sides of lengths $|2\vec{l}|$ and $|2\vec{m}|$ about the origin O. OA and OB lie opposite to each other and so does OC and OD. Thus, condition (ii) of theorem 2 is satisfied and the origin is the F–T point. Hence, condition (118) reproduces equation (8) which shows that the bound on the unsharpness is not more stringent than the two-observable case, as to be expected.
11.4. Arthur–Kelly model

The Arthur–Kelly model for the case of three qubits proceeds exactly as that of two qubits with the Hamiltonian of the measurement interaction of the form \( H = -(q_1 \otimes \sigma_x + q_2 \otimes \sigma_x + q_3 \otimes \sigma_z) \).

The nature of the interaction is impulsive as in equation (23) and the corresponding unitary evolution is given by

\[
U = e(q_1, q_2, q_3) \otimes 1_s + f(q_1, q_2, q_3) \otimes \sigma_x + g(q_1, q_2, q_3) \otimes \sigma_y + h(q_1, q_2, q_3) \otimes \sigma_z, \tag{121}
\]

with

\[
e(q_1, q_2, q_3) = \cos \left( \sqrt{q_1^2 + q_2^2 + q_3^2} \right), \tag{122}
\]

\[
f(q_1, q_2, q_3) = i q_1 \frac{\sin \left( \sqrt{q_1^2 + q_2^2 + q_3^2} \right)}{\sqrt{q_1^2 + q_2^2 + q_3^2}}, \tag{123}
\]

\[
g(q_1, q_2, q_3) = i q_2 \frac{\sin \left( \sqrt{q_1^2 + q_2^2 + q_3^2} \right)}{\sqrt{q_1^2 + q_2^2 + q_3^2}}, \tag{124}
\]

\[
h(q_1, q_2, q_3) = i q_3 \frac{\sin \left( \sqrt{q_1^2 + q_2^2 + q_3^2} \right)}{\sqrt{q_1^2 + q_2^2 + q_3^2}}. \tag{125}
\]

The initial joint state of the three detectors and system is \( |\psi_i\rangle = |\psi_1\rangle \otimes |\psi_2\rangle \otimes |\psi_3\rangle \otimes |\chi\rangle \).

As before, the initial detector states are Gaussians: \( |\psi_1(q_1)\rangle = \left[ \frac{1}{(4\pi)^{3/2}} \exp[-q_1^2/2a^2] \right]^2 \),

\( |\psi_2(q_2)\rangle = \left[ \frac{1}{(4\pi)^{3/2}} \exp[-q_2^2/2b^2] \right]^2 \) and \( |\psi_3(q_3)\rangle = \left[ \frac{1}{(4\pi)^{3/2}} \exp[-q_3^2/2c^2] \right]^2 \).

The detector outcome \((p_1 \geq 0, p_2 \geq 0, p_3 \geq 0)\) is taken here to correspond to the outcome \((+, +, +)\) for the joint unsharp measurement of the system observables \( \sigma_x, \sigma_y, \) and \( \sigma_z, (p_1 \geq 0, p_2 \geq 0, p_3 \leq 0)\) to \((+, +, -)\) and so on. The POVM elements corresponding to the outcomes \((+, +, +), (+, +, -),\) etc of the joint unsharp measurement of \( \sigma_x, \sigma_y, \) and \( \sigma_z \) have a similar structure to that for the case of two observables: \((+, +, +) \leftrightarrow \frac{1}{8} (I + \frac{1}{2} \sigma_x + \frac{1}{2} \sigma_y + \frac{1}{2} \sigma_z), (+, +, -) \leftrightarrow \frac{1}{8} (I + \frac{1}{2} \sigma_x + \frac{1}{2} \sigma_y - \frac{1}{2} \sigma_z), \ldots, (-, -, -) \leftrightarrow \frac{1}{8} (I - \frac{1}{2} \sigma_x - \frac{1}{2} \sigma_y - \frac{1}{2} \sigma_z),\)

\[
a' = \int_{p_1=0}^{+\infty} \int_{p_2=-\infty}^{+\infty} \int_{p_3=-\infty}^{+\infty} 4(f^0 e^0) \, dp_1 \, dp_2 \, dp_3, \tag{126}
\]

\[
b' = \int_{p_2=0}^{+\infty} \int_{p_1=-\infty}^{+\infty} \int_{p_3=-\infty}^{+\infty} 4(g^0 e^0) \, dp_1 \, dp_2 \, dp_3, \tag{127}
\]

\[
c' = \int_{p_3=0}^{+\infty} \int_{p_2=-\infty}^{+\infty} \int_{p_1=-\infty}^{+\infty} 4(h^0 e^0) \, dp_1 \, dp_2 \, dp_3, \tag{128}
\]

with \( e^0, f^0, g^0, h^0 \) being the Fourier transforms of \( e, f, g, h \), respectively.
11.4.1. Symmetric case with a=b=c. Choosing the initial joint pointer state to be a symmetric Gaussian, as in the two-observable case, we have for the unsharp measurement along any three directions $\hat{n}, \hat{m}, \hat{l}$ the marginal effects being given by

\[ \Upsilon_{\hat{n}}^{+} = \frac{1}{2} (I + a' \hat{n} \cdot \vec{\sigma}) \]

\[ \Upsilon_{\hat{m}}^{+} = \frac{1}{2} (I + a' \hat{m} \cdot \vec{\sigma}) \]

\[ \Upsilon_{\hat{l}}^{+} = \frac{1}{2} (I + a' \hat{l} \cdot \vec{\sigma}) \]

Thus, from equation (120), we have \( a' \) to be upper bounded by \( \frac{1}{\sqrt{3}} = 0.577 \). Numerically, for our scheme, \( a' \) is seen to be able to reach up to about 0.49. It cannot reach the bound 0.577, as for two-observable joint measurement (see figure 10).

For the approximate joint measurement in the directions $\hat{l} \equiv (1, 0, 0)$, $\hat{m} \equiv (\cos(\phi), \sin(\phi), 0)$, $\hat{n} \equiv (\sin(\theta) \cos(\phi_1), \sin(\theta) \sin(\phi_1), \cos(\theta))$ with $\theta = 0.414\pi$, $\phi_1 = 0.159\pi$ and $\phi = 0.477\pi$, the F–T point of the set of points A, B, C, D, as defined before, is seen to be at $\hat{l} + \hat{m} - \hat{n}$. This yields, from inequality (118), $a' \leq 0.667$. Thus, as for the two-observable case, considerably more freedom is available for the unbiased measurement in non-orthogonal directions which our scheme cannot take advantage of.

12. Conclusions

Non-commuting observables cannot be measured jointly. However, it is possible in the POVM formalism to perform joint measurements of (generally non-commuting) observables which are approximations of the actual non-commuting observables and called unsharp observables. We have considered in this paper the approximate joint measurement of two- and three-qubit observables separately, through an Arthur–Kelly-like model for qubit observables. This model comes naturally when one considers a Stern–Gerlach setup with a linear magnetic field. In the Stern–Gerlach setup, the momenta of the atoms act as pointer observables for their spin degrees of freedom. Considering the approximate joint measurement of $\sigma_i$ and $\sigma_j$ through this model, we have shown here numerically that the measurement uncertainty relations derived elsewhere (see [6]) hold. It has also been shown how increasing the relative sharpness of the initial momentum wavefunctions of the detectors leads the measurement of one observable to become almost sharp while that of the other to become almost trivial. The effect of initial detector states on the post-measurement state of the system has also been
considered. The action of the measurement interaction on the system turns out to be that of an asymmetric depolarizing channel. This forms the basis of a physical understanding of the origin of complementarity (between the $\sigma_x$ and $\sigma_y$ measurement) in the model. We also see an indication of the entanglement between the system and the detectors increasing as one of the measurements becomes sharper.

We have considered two different characterizations of unsharpness. First, by comparing the probability distribution of the values of the observable to be approximately measured with that of the approximate observable. Second, by considering the alignment of the momentum direction with the spin observable in the Heisenberg picture. We have shown that for the case in which both the observables are approximated equally well, the corresponding measures of unsharpness are proportional. For our choice of the pointer observable, the measures checking the alignment and disturbance due to measurement do not seem to satisfy an error-disturbance relationship.

We have expounded the connection between the symmetries of the underlying Hamiltonian for measurement interaction and initial detector states with the joint measurement in a lemma. This was first stated by Martens et al [22]. It has then been used to perform the approximate joint measurement in arbitrary directions. The POVM elements calculated for the same match with those which were found earlier in the orthogonal case. They also turn as expected to that of a single unsharp measurement when the directions are taken to be almost the same.

For the case of the joint unsharp measurement of three-qubit observables, we have given a necessary condition to be satisfied by the parameters of the marginal POVM elements. This condition has been derived from certain geometric considerations involving the so-called Fermat–Toricelli point. The condition is sufficient for the case of three orthogonal observables and identical to the only necessary–sufficient condition known for three-observable joint measurements. Our proof shows that this also holds for biased unsharp measurements, namely those in which the probability of obtaining ‘up’ or ‘down’ is different for a maximally mixed state.

The measurement scheme employed by us for the joint measurement in non-orthogonal directions cannot take advantage of the greater freedom available for better approximation compared to the orthogonal case. It will be interesting to see such a scheme in the Arthur–Kelly setup that is able to get close to the bound set by equation (8) for arbitrary directions. We have shown that the retrodictive fidelity $\eta_i$ and the unsharpness $a'$ are proportional for the symmetric joint unsharp measurement case that we have considered. It may be possible to connect the two pictures in a more general setting starting with certain symmetries of the Hamiltonian and initial detector states. The problem of determining necessary–sufficient conditions for the most general joint measurement of three observables by the extension of our approach or otherwise is also left open.

Acknowledgments

We would like to thank both referees for providing their valuable comments/suggestions to revise the earlier version of the manuscript. We also want to thank Syed Raghib Hasaan and Rajeev Singh for useful discussions.

References

[1] Busch P, Lahti P J and Mittelstaedt P 1996 The Quantum Theory of Measurement (Berlin: Springer)
[2] Braginsky V B and Khalili F Y 1992 Quantum Measurement (Cambridge: Cambridge University Press)
[3] Arthurs E and Kelly J R 1965 Bell Syst. Tech. J. 44 725

35
