## Review Article

# Applications of Quantum Probability Amplitude in Decision Support Systems 

Shahram Payandeh (ㅁ)<br>Networked Robotics and Sensing Laboratory, School of Engineering Science, Simon Fraser University, Burnaby, British Columbia V5A 1S6, Canada<br>Correspondence should be addressed to Shahram Payandeh; payandeh@sfu.ca

Received 15 May 2023; Revised 3 July 2023; Accepted 17 November 2023; Published 7 December 2023
Academic Editor: Anandakumar Haldorai
Copyright © 2023 Shahram Payandeh. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Establishing various frameworks for managing uncertainties in decision-making systems have been posing many fundamental challenges to the system design engineers. Quantum paradigm has been introduced to the area of decision and control communities as a possible supporting platform in such uncertainty management. This paper presents an overview of how a quantum framework and, in particular, probability amplitude has been proposed and utilized in the literature to complement two classical probabilistic decision-making approaches. The first such framework is based in the Bayesian network, and the second is based on an element of Dempster-Shafer (DS) theory using the definition of mass function. The paper first presents a summary of these classical approaches, followed by a review of their preliminary enhancements using the quantum model framework. Particular attention was given on how the notion of probability amplitude is utilized in such extensions to the quantum-like framework. Numerical walk-through examples are combined with the presentation of each method in order to better demonstrate the extensions of the proposed frameworks. The main objective is to better define and develop a common platform in order to further explore and experiment with this alternative framework as a part of a decision support system.

## 1. Introduction

Design of autonomous systems or systems which can assist or collaborate with people in a semiautonomous framework has been gaining considerable attention for the past decades [1, 2]. For the majority of these systems, a typical system architecture involves a layer of sequenced sensing and perception, followed by a layer of decision-making or actuations. Sensing and sequences of sensing in general gives noisy measurement of information from the environment which can then be used and be interpreted by the next stages of decision-making [3, 4].

In order to construct a basis of such decision-making frameworks in presence of uncertainties, two different sets of axioms were proposed in the area of probability theory. One formulation is based on Kolmogorov axioms (Kolmogorov, 1933/1950) and the other is based on von Neumann axioms (von Neumann, 1932/1955). Former organized the principles underlying the classical probabilistic applications, and
the other was based on the probabilistic interpretation of laws underlying quantum mechanics. At the conceptual level, a key difference is that the classical probability theory relies on a set-theoretic representation, whereas quantum theory used a vector space representation. There have been a number of interpretations of quantum mechanics in the context of classical decision-making. In this paper, we present an overview of two classical frameworks which have been used in the context of decision-making and presents how the quantum model and, in particular, quantum probability amplitudes have been used within these two classical models.
1.1. Background Review. There exist a number of methodologies proposed in the literature for decision-making under uncertainties. In this section, we highlight a number of these methods. Expected utility theory (EUT) is a traditional method of decision-making that takes uncertainty into
account by tying probabilities to potential outcomes. It entails figuring out each alternative's expected utility based on its likelihoods and potential rewards. The choice is then made based on which option has the most projected usefulness. Prospect Theory: By taking into account how people view and assess uncertainties, prospect theory expands on the conventional expected utility theory. It takes into account the concepts of risk aversion and loss aversion, where people choose losses over equal rewards. To account for these biases, prospect theory includes value functions and probability weighting. Fuzzy logic allows for degrees of truth or membership rather than binary values in order to deal with uncertainty. By designating linguistic phrases to express uncertainty, it enables decision-makers to handle hazy or ambiguous information. Fuzzy logic makes it easier to reason and make decisions in challenging, unpredictable situations. Making Robust Decisions: Robust decisionmaking techniques look for tactics that work effectively in a variety of conceivable future scenarios. These techniques take into account the robustness or resilience of judgements under diverse uncertainty rather than optimising for a particular result. Techniques like robust optimisation and sensitivity analysis are frequently applied in this situation. Decision Trees: Decision trees are visual representations of decisions and their outcomes in the form of a tree-like structure. Decision trees make it easier to evaluate options when there are uncertainties by assigning probability and values to various outcomes. Decision trees can be used in conjunction with methods such as expected value analysis and Monte Carlo simulation. Bayesian Decision Analysis: Bayesian decision analysis uses decision theory and Bayesian inference to make decisions in the face of uncertainty. It entails re-evaluating prior assumptions in light of new information to derive posterior probabilities, which are then used to determine the best course of action based on variables like expected utility or expected regret. A mathematical optimisation technique called stochastic programming takes into account uncertainties modelled by probabilistic distributions. It creates decision models with specified probabilistic aims and restrictions and then solves them to get solid conclusions that hold up under a variety of circumstances.

In the still-emerging discipline of quantum decisionmaking, theories from quantum mechanics are applied to the problem of decision-making. It offers alternative models for decision-making processes by utilising ideas and mathematical frameworks from quantum physics. Traditional approaches to decision-making frequently rely on conventional probabilities and logic. However, quantum decision-making contends that the use of quantum phenomena like superposition, entanglement, and interference can be advantageous for decision-making processes. There has been a number of work establishing fundamentals of quantum mechanics within the notion decision-making under uncertainties. For example, recently [5] applied the quantum computational model to represents modelling the human affective system and applying lessons learned to human-robot interaction, in handling ambiguous emotional states and probabilistic decisions using quantum logic on
fuzzy-sets hardware. Reference [6] presents some results in exploring the notion of interference effect within the context of Bayesian network and fuzzy set theory. In addition, they have utilized the Dempster-Shafer's evidence theory to transform fuzzy number into probability. Reference [7] argues that human decision-making follows Bayes probabilistic reasoning and there are relationships between Bayes probability, fuzzy sets, and quantum probability.

This paper presents a tutorial overview of how the notion of quantum probability has been integrated within the two classical approaches in the decision-making, namely, Bayesian network and Dempster-Shafer theory (in particular the notion of the mass functions). Through the step-by-step and break-down of basic approaches, it is hoped that the reader can gain a better understanding of such quantum extensions and allow for further study of future extensions.

The paper is organized as follows: Section 2 presents the main results of this paper which starts with the Bayesian network. It contains a definition of naive Bayesian network and the associated marginalization, followed by an example of how quantum model can be integrated within the classical definitions. This section also presents an overview of Dempster-Shafer theory of combining information. In particular, how the notion of the mass function, which is the cornerstone of this formulation is used in the literature as a basis of its extension to the quantum framework. Section 3 presents some discussions and concluding remarks. Throughout the presentation, simple numerical examples are used in order to better demonstrate both of these classical methods and their extensions to the quantum framework and quantum probability amplitudes. It is hoped that this paper can provide a better understanding of some of these example extensions in order to further investigate their properties.

## 2. Proposed Approach: Overview of Probability Amplitude Applications

This section presents the main results of this paper. The section is divided into two parts: Bayesian network and Dempster-Shafer theory. The contribution of this section is a step-by-step discretion and walk-through examples of the general methods which have been proposed in the literature in defining how the notion of quantum probability amplitude can be used to enhance these two classical methods.
2.1. Bayesian Network. This section presents an application of the quantum paradigm and probability amplitude which have been proposed in the literature in the construction of the Bayesian network. First, we present a background review of the general Bayesian networks in context of decisionmaking, followed by an extension to the quantum-enhanced Bayesian network.

The Bayesian method offers an approach for combining evidence according to probability theory [8, 9]. Uncertainties are represented using the conditional probability and through Bayes rules. The framework can also be defined
in a sequential scheme where the degree of belief in a hypothesis can be updated based on new information [10-12].

Naive Bayes model is proposed to represent the conditional independence of parameters. This model states that the random variables are conditionally independent given the instances of the parent in the context of a graph model representation. For example, the instance of the parent node $Y$ is given by the propagation instances of its children's random variables $\left\{X_{1}, X_{2}, \cdot, X_{n}\right\}$ which can be represented by the following conditional probability:

$$
\begin{equation*}
\operatorname{Pr}\left(Y \mid X_{1}, X_{2}, \cdots X_{n}\right) \tag{1}
\end{equation*}
$$

which can further be expanded as follows:

$$
\begin{equation*}
\operatorname{Pr}\left(Y \mid X_{1}, X_{2}, \cdots X_{n}\right)=\frac{\operatorname{Pr}\left(X_{1}, X_{2}, \cdots X_{n} \mid Y\right) \operatorname{Pr}(Y)}{\operatorname{Pr}\left(X_{1}, X_{2}, \cdots X_{n}\right)} \tag{2}
\end{equation*}
$$

where the probability representation presented in the denominator can be assumed to be constant of proportionality due to the fact that this joint probability do not dependent on the instance of the parent node $Y$. As a result, equation (2) can further be expressed as the reduced condition of joint distributions, or

$$
\begin{equation*}
\operatorname{Pr}\left(Y \mid X_{1}, X_{2}, \cdots X_{n}\right) \propto \operatorname{Pr}\left(X_{1}, X_{2}, \cdots X_{n} \mid Y\right) \operatorname{Pr}(Y)=\operatorname{Pr}\left(Y, X_{1}, X_{2}, \cdots X_{n}\right) \tag{3}
\end{equation*}
$$

Give the above joint probability distribution of equation (3), it can be further expanded using the chain rule to obtain

$$
\begin{align*}
\operatorname{Pr}\left(Y, X_{1}, X_{2}, \cdots X_{n}\right)= & \operatorname{Pr}(Y) \operatorname{Pr}\left(X_{1} \mid Y\right) \operatorname{Pr}\left(X_{2} \mid Y, X_{1}\right) \operatorname{Pr}\left(X_{3} \mid Y, X_{1}, X_{2}\right) \\
& \left.\cdots \operatorname{Pr}\left(X_{n}\right) \mid Y, X_{1}, X_{2}, X_{n-1}\right) . \tag{4}
\end{align*}
$$

The Naive Bayesian assumption states that each feature $X_{i}$ defined in the above is conditionally independent of every other feature $X_{j}, \forall i \neq j$, given the instance of their parent class $Y$. This as a result implies that

$$
\begin{equation*}
\operatorname{Pr}\left(X_{i} \mid Y, X_{j}\right)=\operatorname{Pr}\left(X_{i} \mid Y\right) \forall i \neq j . \tag{5}
\end{equation*}
$$

With the above assumption, the joint probability distribution defined in equation (3) can be written as follows:

$$
\begin{align*}
& \operatorname{Pr}\left(Y \mid X_{1}, X_{2}, \cdots X_{n}\right) \propto \\
& \quad \operatorname{Pr}(Y) \operatorname{Pr}\left(X_{1} \mid Y\right) \operatorname{Pr}\left(X_{1} \mid Y\right) \operatorname{Pr}\left(X_{3} \mid Y\right), \cdots \operatorname{Pr}\left(X_{n}\right) \mid Y \propto \\
& \quad \operatorname{Pr}(Y) \prod_{i=1}^{n} \operatorname{Pr}\left(X_{i} \mid Y\right) \tag{6}
\end{align*}
$$

where by using the naive independence assumption again in the normalization factor, equation (6) can be written as follows:

$$
\begin{equation*}
\operatorname{Pr}\left(Y \mid X_{1}, X_{2}, \cdots X_{n}\right)=\alpha \operatorname{Pr}(Y) \prod_{i=1}^{n} \operatorname{Pr}\left(X_{i} \mid Y\right) \tag{7}
\end{equation*}
$$

The above equation can be extended to the case of the Bayesian network which uses the conditional independence and the Markov assumption [9]. Let $X=\left\{X_{1}, X_{2}, \ldots, X_{n}\right\}$ be a set of $n$ random variables of a Bayesian network graph structure. Let Parents $\left(X_{i}\right)$ be the parent of the random variable $X_{i}$ and NonDescendants $\left(X_{i}\right)$ be the variables in the graph that are nondescendants of $X_{i}$. The Markov assumption states that each variable $X_{i}$ is independent of its nondescendant given its parent $X_{i} \Pi$ NonDescendents $\left(X_{i}\right) \mid$ Parents $\left(X_{i}\right)$. Combining
the naive Bayes formula given in equation (7) with the definition of local independences, we can write the following relationship:

$$
\begin{equation*}
\operatorname{Pr}\left(X_{1}, X_{2}, \ldots, X_{n}\right)=\beta \prod_{i-1}^{n} \operatorname{Pr}\left(X_{i} \mid \operatorname{Parents}\left(X_{i}\right)\right) \tag{8}
\end{equation*}
$$

2.1.1. Quantum-Enhanced Bayesian Network. Within the past decades, there have been many interpretations of how some of the basic definitions of quantum mechanics can be extended to the framework of Bayesian reasoning and decision-making; see [13-17]. Using the Naive Bayesian network reasoning, interpretations of quantum mechanics have allowed various exploratory studies in generalization of the Bayesian decision-making. Such studies can allow formalization of general tools which can further be enhanced and complemented within the autonomous and/or human-in-the-loop decision-making process.

In this section and through a descriptive example, we explore how some of the basic reasoning in the classical Bayesian network can be complemented with the quantum type representation. Classical Bayesian network is represented by a directed acyclic graph structure where each node is a representation of a random variable and each edge represents a direct influence of the parent on the child node. Another interpretation of the graph is its representation of the independence relationship in terms of conditional probability over the values of a node given each possible joint assignment of values of its parent.

As an example, let us considered a two-node Bayesian network ( $N_{1}$ is the parent node and $N_{2}$ the child node),
where the probability distribution of node $N_{2}$ is directly influenced by node $N_{1}$. For example, node $N_{1}$ can be associated with the state of a robot detecting an overhead light and observing its color changes as it passes a certain location along its discrete trajectory (e.g., light turning blue or orange). Node $N_{2}$ would be the decision that the robot will face in turning left or right Table 1. Here, we are assuming that the light will have an equal probability of changing its color to either blue or orange.

The right column of Table 2 shows the joint distributions on the conditions of this two node example. As it can be seen from the table, the full value of the sum of this joint distributions over all the values are equal to 1 .

In general, and given two random variables $X$ and $Y$, the marginal probability of $X$ is simply the probability of $X$ averaging over the information about $Y$ [9]. For the discrete random variables, such marginalization can be defined as follows:

$$
\begin{align*}
\operatorname{Pr}(X=x) & =\sum_{y} \operatorname{Pr}(X=x, Y=y) \\
& =\sum_{y} \operatorname{Pr}(X=x \mid Y=y) \operatorname{Pr}(Y=y) . \tag{9}
\end{align*}
$$

The full joint distribution of the Bayesian network is defined based on equation (8), where $X_{i}$ is a list of random variables and Parent $\left(X_{i}\right)$ corresponds to all parent nodes of $X_{i}$. For example, after computing the joint distributions of the two-node Bayesian network, we need to sum out all the variables that are unknown, say the outcome of the $N_{1}$ (i.e., light turning blue or orange). This can be accomplished by applying the probability expansion shown in equation (9). For example, the probability of the robot taking either of the actions given that the observed information was the light turning blue, e.g., $\operatorname{Pr}\left(N_{2}=\right.$ Left $\mid N_{1}=$ blue $)$ can be computed as follows:

$$
\begin{align*}
\operatorname{Pr}\left(N_{2}=\text { Left } \mid N_{1}=\text { blue }\right) & =\frac{\operatorname{Pr}\left(N_{1}=\text { blue }, N_{2}=\text { left }\right)}{\operatorname{Pr}\left(N_{1}=\text { blue }\right)}  \tag{10}\\
& =\frac{0.355}{0.355+0.145}=0.71
\end{align*}
$$

with the negation probability of

$$
\begin{align*}
\operatorname{Pr}\left(N_{2}=\text { right } \mid N_{1}=\text { blue }\right) & =\frac{\operatorname{Pr}\left(N_{1}=\text { blue }, N_{2}=\text { right }\right)}{\operatorname{Pr}\left(N_{1}=\text { blue }\right)} \\
& =\frac{0.145}{0.355+0.145}=0.29, \tag{11}
\end{align*}
$$

which satisfying the requirement of: $\operatorname{Pr}\left(N_{2}=\right.$ left $\mid N_{1}=$ blue $)+\operatorname{Pr}\left(N_{2}=\right.$ right $\mid N_{1}=$ blue $)=1$.

Table 1: Distribution of the action variables $N_{2}$ given the observed variables $N_{1}$.

|  | $\operatorname{Pr}\left(N_{2}=\right.$ left $)$ | $\operatorname{Pr}\left(N_{2}=\right.$ right $)$ |
| :--- | :---: | :---: |
| $\operatorname{Pr}\left(N_{1}=\right.$ blue $)$ | 0.71 | 0.29 |
| $\operatorname{Pr}\left(N_{1}=\right.$ orange $)$ | 0.5 | 0.5 |

Table 2: Full joint distribution of the Bayesian network of Table 1.

| $N_{1}$ | $N_{2}$ | $\operatorname{Pr}\left(N_{1}, N_{2}\right)$ |
| :--- | :---: | :---: |
| Blue | Left | $0.5 \times 0.71=0.355$ |
| Blue | Right | $0.5 \times 0.29=0.145$ |
| Orange | Left | $0.5 \times 0.5=0.25$ |
| Orange | Right | $0.5 \times 0.5=0.25$ |

The marginal probability, e.g., the robot turning left or turning right without detecting or knowing the outcome of the observation which would be color change of light can be computed from equation (9) as follows:

$$
\begin{align*}
\operatorname{Pr}\left(N_{2}=\text { left }\right)= & \sum_{n \in N_{1}} \operatorname{Pr}\left(N_{1}=n, N_{2}=\text { left }\right) \\
= & \sum_{n \in N_{1}} \operatorname{Pr}\left(N_{1}=n\right) \operatorname{Pr}\left(N_{2}=\text { left } \mid N_{1}=n\right) \\
= & \operatorname{Pr}\left(N_{1}=\text { blue }\right) \operatorname{Pr}\left(N_{2}=\text { left } \mid N_{1}=\text { blue }\right) \\
& +\operatorname{Pr}\left(N_{1}=\text { orange }\right) \operatorname{Pr}\left(N_{2}=\text { left } \mid N_{1}=\text { orange }\right) \\
= & 0.355+0.25=0.605 . \tag{12}
\end{align*}
$$

Similarly, we can compute $\operatorname{Pr}\left(N_{2}=\right.$ right $)=0.145+$ $0.25=0.395$ which also satisfied condition of $\operatorname{Pr}\left(N_{2}=\right.$ left $)+$ $\operatorname{Pr}\left(N_{2}=\right.$ right $)=1$.

In quantum representation of the probability, the finite choices in the example which are available at each node of the graph can be represented by a state vector and as a superposition of the corresponding basis representing each of the choices (Appendix A-E), [18]). For example, for each of the nodes of our simple Bayesian network, the corresponding state vector in each of the Hilbert spaces $\mathscr{H}_{1}$ and $\mathscr{H}_{2}$ can be defined at each node as follows:

$$
\begin{align*}
& \left.\left.|\Psi\rangle_{N_{1}}=e^{j \theta_{\text {blue }}} \mid \text { blue }\right\rangle+e^{j \theta_{\text {orange }}} \mid \text { orange }\right\rangle, \\
& \left.\left.|\Psi\rangle_{N_{2}}=e^{j \theta_{\text {left }}} \mid \text { left }\right\rangle+e^{j \theta_{\text {right }}} \mid \text { right }\right\rangle, \tag{13}
\end{align*}
$$

which can be represented the combined Hilbert space $\mathscr{H}$ of the each of the subspaces as follows: $\mathscr{H}=\mathscr{H}_{1} \otimes \mathscr{H}_{2}$. The basis of this combined space can be defined using the following tensor product (outer product) or

$$
\begin{equation*}
\left.\left.\left.\left.\left.\left.\left.\mid \text { blue }\rangle_{\mathscr{N}_{1}} \otimes \mid \text { left }\right\rangle_{\mathscr{N}_{2}} ; \mid \text { blue }\right\rangle_{\mathscr{N}_{1}} \otimes \mid \text { right }\right\rangle_{\mathcal{N}_{2}} ; \mid \text { orange }\right\rangle_{\mathscr{N}_{1}} \otimes \mid \text { left }\right\rangle_{\mathscr{N}_{2}} ; \mid \text { orange }\right\rangle_{\mathscr{N}_{1}} \otimes \mid \text { right }\right\rangle_{\mathscr{N}_{2} .} \tag{14}
\end{equation*}
$$

The new combined state vector description $|\Psi\rangle \in \mathscr{H}$ can then be defined as follows:

$$
\begin{align*}
|\Psi\rangle= & \left.\left.\left.\left.\omega_{1} \mid \text { blue }\right\rangle_{N_{1}} \otimes \mid \text { left }\right\rangle_{\mathcal{N}_{2}}+\omega_{2} \mid \text { blue }\right\rangle_{\mathcal{N}_{1}} \otimes \mid \text { right }\right\rangle_{\mathscr{N}_{2}} \\
& \left.\left.\left.\left.+\omega_{3} \mid \text { orange }\right\rangle_{\mathscr{N}_{1}} \otimes \mid \text { left }\right\rangle_{\mathcal{N}_{2}}+\omega_{4} \mid \text { orange }\right\rangle_{\mathcal{N}_{1}} \otimes \mid \text { right }\right\rangle_{\mathcal{N}_{2}}, \tag{15}
\end{align*}
$$

where $\omega_{i}$ represents the corresponding probability amplitude along each of the combined basis subspaces in the new space. Referring to example defined in Table 2 and the Bayesian model, the following probability amplitudes can be defined:

$$
\begin{align*}
& \omega_{1}=\sqrt{0.5} e^{\left(\theta_{\text {blue }}\right)} \times \sqrt{0.71} e^{\left(\theta_{\text {left }}\right)}, \\
& \omega_{2}=\sqrt{0.5} e^{\left(\theta_{\text {blue }}\right)} \times \sqrt{0.29} e^{\left(\theta_{\text {right }}\right)}, \\
& \omega_{3}=\sqrt{0.5} e^{\left(\theta_{\text {orange }}\right)} \times \sqrt{0.5} e^{\left(\theta_{\text {left }}\right)},  \tag{16}\\
& \omega_{4}=\sqrt{0.5} e^{\left(\theta_{\text {orange }}\right)} \times \sqrt{0.5} e^{\left(\theta_{\text {left }}\right)},
\end{align*}
$$

along the corresponding bases space

$$
\begin{align*}
\mid \text { blue left }\rangle & =(1000)^{T}, \\
\mid \text { blue right }\rangle & =(0100)^{T}, \\
\mid \text { orange left }\rangle & =(0010)^{T},  \tag{17}\\
\mid \text { orange right }\rangle & =(0001)^{T},
\end{align*}
$$

$$
\begin{align*}
\operatorname{Pr}\left(N_{2}=\text { left }\right) & =\left.\left|\sum_{n \in N_{1}}\right| \Psi\left(N_{1}=n, N_{2}=\text { left }\right\rangle\right|^{2} \\
& =\| \Psi\rangle_{N_{1}=\text { blue }, N_{2}=\text { left }}+\left.|\Psi\rangle_{N_{1}=\text { orange }, N_{2}=\text { left }}\right|^{2}  \tag{18}\\
& =\left|\sqrt{0.355} e^{j \theta_{1}}+\sqrt{0.25} e^{j \theta_{3}}\right|^{2} \\
& =|\sqrt{0.355}|^{2}+|\sqrt{0.25}|^{2}+2 \sqrt{0.355} \sqrt{0.25} \cos \left(\theta_{1}-\theta_{3}\right) .
\end{align*}
$$

Comparing the above computation with the classical example of computing the marginal probability solved previously, it can be seen that for the case when $\cos \left(\theta_{1}-\theta_{3}\right)=0$, the quantum enhanced marginal probability is the same as the classical computation. In the above expansion, the term $2 \sqrt{0.355} \sqrt{0.25} \cos \left(\theta_{1}-\theta_{3}\right)$ is referred
to as the interference term.
2.2. Dempster-Shafer (DS) Formulation of Mass Function. Another broad approach which has been proposed in the literature in combining conditional information obtained through various sources of sensing modalities is based on Dempster-Shafer evidence theory, which also referred to as the generalization of the Bayesian theory [19, 20]. It has been used in various applications of data fusion in a sensor network, e.g., [21]. It provides a formalism that could be used

(20.
$|\Psi\rangle$ represents a quantum superposition over all possible states of the Bayesian network. Such superposition can be view as a quantum-like joint full distribution. Table 3 shows such distribution associated with the working example of this section.

In the quantum-enhanced Bayesian network, marginal probability can be carried by following a similar procedural approach as highlighted above. However, in the quantum framework, the procedure for computing marginal probability can follow the definition of Feynman's second rule for combining probabilities in the associated probability path diagram (Appendix A-E). The second rule states that the probability amplitude of transition from a state which is defined as the initial node to the final state node, taking multiple indistinguishable paths, is given by the sum of the amplitudes of each path. For example, for the two-node example in Table 3, the quantum-enhanced marginal probability of the robot turning left can be computed as follows:
to represent incomplete knowledge, updating beliefs, and a combination of evidence for explicitly representing uncertainty. In the following, we present first an overview of DS theory and its interpretation in the context of multiple sensor fusion. This can offer the reader a better sense of how the enhanced model using quantum formulation is utilized within a component of the DS formulation, namely the mass function.

In Dempster-Shafer (DS) theory, each inquiry into a fact has a degree of support between 0 and 1 (i.e., 0 no support for the fact and 1 full support for the fact). Let a set of possible conclusions be a mutually exclusive and exhaustive set of all possible outcomes be given as: (for example, these can be measurements from a collection of sensors monitoring a common event).

$$
\begin{equation*}
\mathcal{S}=\left(s_{1}, s_{2}, \ldots, s_{n}\right) \tag{19}
\end{equation*}
$$

Table 3: Quantum-enhanced full joint distribution of the Bayesian network.

| $\mathcal{N}_{1}$ | $\mathcal{N}_{2}$ | $\omega$ |
| :--- | :---: | :---: |
| Blue | Left | $\sqrt{0.5} \times \sqrt{0.71} e^{j\left(\theta_{\text {blue }}+\theta_{\text {left }}\right)}=\sqrt{0.355} e^{j \theta_{1}}$ |
| Blue | Right | $\sqrt{0.5} \times \sqrt{0.29} e^{j\left(\theta_{\text {blue }}+\theta_{\text {right }}\right)}=\sqrt{0.145} e^{j \theta_{2}}$ |
| Orange | Left | $\sqrt{0.5} \times \sqrt{0.5} e^{j\left(\theta_{\text {orange }}+\theta_{\text {leff }}\right)}=\sqrt{0.25} e^{j \theta_{3}}$ |
| Orange | Right | $\sqrt{0.5} \times \sqrt{0.5} e^{j\left(\theta_{\text {orange }}+\theta_{\text {right }}\right)}=\sqrt{0.25} e^{j \theta_{4}}$ |

where $\mathcal{\delta}$ is referred to as the universe or frame of discernment, where at least one of the $s_{i}$ must be true. DS theory is concerned with pieces of evidence which support subsets of outcomes in $\mathcal{S}$ which have $2^{\mathcal{S}}$ elements represented in its power set over the frame of discernment. For example, for a three element membership set defined in relation to equation (19), the associated power set can be written as follows:

$$
\begin{equation*}
A=2^{\mathcal{S}}=\left\{\emptyset, s_{1}, s_{2}, s_{3},\left(s_{1}, s_{2}\right),\left(s_{2}, s_{3}\right),\left(s_{1}, s_{3}\right),\left(s_{1}, s_{2}, s_{3}\right)\right\} \tag{20}
\end{equation*}
$$

where $\emptyset$ represents the empty set and has the probability of 0 , and each of the other elements in the power set has a probability between 0 and 1 .

Mass function for $\mathcal{\delta}$ is defined as a mapping of $m$ from 0 to $1: m: 2^{\mathcal{S}} \longrightarrow[0,1]$. Mass function of a member $A_{i}$ of the power set, or $\left(A_{i}\right)$, is equal to the portion of all evidence that supports the existence of the element $A_{i}$ of the power set. The value of each $m\left(A_{i}\right)$ is between 0 and 1 , where $\sum_{i}^{n} m\left(A_{i}\right)=1$. This also implies the evidence tells us the truth is in $A_{i} \subseteq \mathcal{S}$ for sure and we also have a logical or categorical mass function. Each element of the mass function is called a focal element. $m$ is said to be Bayesian if all focal sets of $m$ are singletons. The value $m\left(A_{i}\right)$ represents the allocation of belief to the possibility that looks for the true value of $s_{i}$ belonging to $\mathcal{S}$. As an example, for our three sensor model example, i.e., $\mathcal{S}=\left(s_{1}, s_{2}, s_{3}\right)$, let us define the following values for the members of the power set defined in equation (20): $m\left(\left\{s_{1}\right\}\right)=0.4, m\left(\left\{s_{2}\right\}\right)=0.3$, and $m(\mathcal{S})=0.3$.

The belief $\beta$ in one of the members of the power set $A$ is defined as the sum of the masses of subset of power set (including the set itself) that represents the existing evidence supporting $A$. In the above example, $\beta\left(\left\{s_{1}\right\}\right)=m\left(\left\{s_{1}\right\}\right)=0.4$.

The plausibility of $A, \pi(A)$, is defined as the sum of all the masses in the power set that intersect with the set $A$. For example, $\pi\left(\left\{s_{1}\right\}\right)=m\left(\left\{s_{1}\right\}\right)+m(\{\mathcal{S}\})=0.4+0.3=0.7$. The certainty which we can assign to a given subset of A is defined by what is referred to as the interval, $[\beta(A) \pi(A)]$. For the working example, this is defined to be [0.40.7]. This will imply that the probability of $A, \operatorname{Pr}(A)$ falls somewhere between $\beta(A)$ and $\pi(A)$.
2.2.1. Combining Mass Functions. When different decisionmaking agents are responsible in gathering information from a shared sensed or observed data, various methodologies have been proposed in the literature in combining this information based on the DS framework. For example, if two mass functions $m_{1}$ and $m_{2}$ from two reliable and distinct sources of information are available, conjunctive rule of combination can be followed. If the sources are distinct but only one of the sources is reliable (and we do not know which one), the disjunctive rule of combination can be followed to construct a combined mass function [22, 23].

For the cases, where $m_{1}$ and $m_{2}$ are two reliable and distinct mass functions over $C \subseteq \mathcal{S}$, Dempster's rule of conjunctive of combination offers an approach for defining such combination as follows [24]:

$$
\begin{align*}
m(C) & =\left(m_{1} \oplus m_{2}\right)(\mathcal{S}) \\
& =m_{1 \oplus 2}(\mathcal{S})  \tag{21}\\
& =\frac{1}{1-\kappa} \sum_{A_{i} \cap B_{j}=C} m_{1}\left(A_{i}\right) \times m_{2}\left(B_{j}\right), \quad \forall \mathcal{S} \neq \emptyset
\end{align*}
$$

where $\kappa$ is referred to as the degree of conflict defined as follows:

$$
\begin{equation*}
\kappa=\sum_{A_{i} \cap b_{j} \neq \emptyset} m_{1}\left(A_{i}\right) \times m_{2}\left(B_{j}\right) . \tag{22}
\end{equation*}
$$

As an example of utilization of the above rule, let us assume that we have again three sensors, i.e., $\mathcal{S}=\left\{s_{1}, s_{2}, s_{3}\right\}$ and let us assume that there exists two decision-making modules which need to be fused in order to obtain their combined information for determining the presence of an object in the common monitoring. Let us define the two mass functions based on the power set description of the sensor set as follows:

$$
\begin{align*}
& m_{1}=\left(m_{1}\left\{s_{1}\right\}=0.3, m_{1}\left\{s_{2}\right\}=0.4, m_{1}\left\{s_{1}, s_{2}, s_{3}\right\}=0.3\right) \\
& m_{2}=\left(m_{2}\left\{s_{1}, s_{2}\right\}=0.5, m_{2}\left\{s_{1}, s_{3}\right\}=0.1, m_{2}\left\{s_{1}, s_{2}, s_{3}\right\}=0.4\right) \tag{23}
\end{align*}
$$

For this example, a measure (or degree) of conflict $\kappa$ defined in equation (22) can be computed to be:

$$
\begin{align*}
\kappa & =\sum_{A_{i} \cap b_{j} \neq \emptyset} m_{1}\left(A_{i}\right) \times m_{2}\left(B_{j}\right) \\
& =\left(m_{1}\left\{s_{1}\right\} \times m_{2}\left\{s_{1}, s_{3}\right\}\right)  \tag{24}\\
& =0.4 \times 0.1 .
\end{align*}
$$

A measure of common shared belief of the two decisionmakers given the measure or observed value of a sensor one, i.e., $s_{1}$ is computed as follows:

$$
\begin{align*}
\left(m_{1} \oplus m_{2}\right)\left(s_{1}\right) & =\frac{1}{1-\kappa} \times\left[\left(m_{1}\left\{s_{1}\right\} \times m_{2}\left\{s_{1}, s_{2}\right\}+m_{1}\left\{s_{1}\right\} \times m_{2}\left\{s_{1}, s_{3}\right\}+m_{1}\left\{s_{1}\right\} \times m_{2}\left\{s_{1}, s_{2}, s_{3}\right\}\right]\right. \\
& =\frac{1}{1-0.04} \times[(0.3) \times(0.5)+(0.3) \times(0.1)+(0.3) \times(0.4)]=0.312 \tag{25}
\end{align*}
$$

similarly, the shared belief for the second sensor $s_{2}$ is computed as follows:

$$
\begin{align*}
\left(m_{1} \oplus m_{2}\right)\left(s_{2}\right) & =\frac{1}{1-\kappa} \times\left[\left(m_{1}\left\{s_{2}\right\} \times m_{2}\left\{s_{1}, s_{2}\right\}+m_{1}\left\{s_{2}\right\} \times m_{2}\left\{s_{1}, s_{2}, s_{3}\right\}\right]\right.  \tag{26}\\
& =\frac{1}{1-0.04} \times[(0.4) \times(0.5)+(0.4) \times(0.4)=0.375
\end{align*}
$$

Table 4 summarizes the complete combined shared believes using the definition of the two mass functions.
2.2.2. Quantum-Enhanced Mass Function. There have been a number of approaches proposed in the literature to extend the framework of DS evidence theory and, in particular, the notion of the mass function within the context of quantum probability construction [25-28].

Here we present an overview of one of the approaches where the state description captures the quantized representation of the mass function over the basis defined by the Dempster combination rule of combining mass functions [29]. The representation is defined as a quantum state with a well-defined quantum superposition.

Referring to the previous section, given a set of observations/measurements defined over $\mathcal{S}$, its quantum representation is written as: $\mathcal{S}_{q}=|\mathcal{S}\rangle=\left\{\left|s_{1}\right\rangle,\left|s_{2}\right\rangle, \ldots,\left|s_{n}\right\rangle\right\}$ (Appendix A-E). This quantum state can further be written as follows: $|\mathcal{S}\rangle=\sum_{i=1}^{n} \omega_{i}\left|s_{i}\right\rangle$. The probability of each basis state $\left|s_{i}\right\rangle$ is $\left|\omega_{i}\right|^{2}$ where $\sum_{i=1}^{n}\left|\omega_{i}\right|^{2}=1$.

The quantum mass function $m$ is defined on $\mathcal{S}_{q}$ as follows:

$$
\begin{equation*}
|\mathscr{M}\rangle=\rho_{1} e^{j \theta_{1}}\left|s_{1}\right\rangle+\rho_{2} e^{j \theta_{2}}\left|s_{2}\right\rangle+\cdots+\rho_{n} e^{j \theta_{n}}\left|s_{n}\right\rangle \tag{27}
\end{equation*}
$$

where $\quad \rho_{i} \geq 0, \quad \theta_{i} \in[-\pi, \pi], \quad \sum_{s_{i} \in \mathcal{S}} \rho_{i}^{2}=1, \quad$ and $\beta(A) \leq \sum_{s_{i} \in A} \rho_{i}^{2} \leq \pi(A), \forall A \subseteq \mathcal{S}$ (where from the previous section we have $\beta($.$) as the belief and \pi($.$) as the plausibility$ of an event) and $\rho_{i} e^{j \theta_{i}}$ is the corresponding probability amplitude.

Equation (27) represents the mass function $|\mathscr{M}\rangle$ as a superposition of state composed of $\left|s_{1}\right\rangle,\left|s_{2}\right\rangle, \ldots\left|s_{n}\right\rangle$ in which $\rho_{i} e^{j \theta_{i}}$ is the probability amplitude associated with basis state $\left|s_{i}\right\rangle$, and $\rho_{i}$ and $\theta_{i}$ are the modulus and phase angle of this probability amplitude, respectively (Appendix $\mathrm{A}-\mathrm{E}) . \sum_{s_{i} \in \mathcal{S}} \rho_{i}^{2}=1$ represents the normalization condition required by the definition of standard quantum state. In addition, the modulus of each probability amplitude, i.e., $\rho_{i}, i=1, \ldots, n$, has to meet the DS constraint defined in the previous section over the associated power set such that $\beta(A) \leq \sum_{s_{i} \in A} \rho_{i}^{2} \leq \pi(A), \forall A \subseteq \mathcal{S}$.

As an example, let us again define a frame of discernment consists of outcome of three sensors $\mathcal{S}=\left(s_{1}, s_{2}, s_{3}\right)$ with the associated power set defined in the previous section. Let us defined the following two mass functions over $\mathcal{S}$ as follows: $m_{1}\left(s_{1}\right)=0.7, m_{1}\left(s_{1}, s_{2}, s_{3}\right)=0.3$ and $m_{2}\left(s_{1}\right)=0.4, m_{2}\left(s_{2}\right)=$ $0.5, m_{2}\left(s_{3}\right)=0.1$.

From definition of equation (27), the first mass function in its quantum representation, $\left|\mathscr{M}_{1}\right\rangle$ over $\mathcal{S}_{q}$ is defined as follows:

$$
\begin{equation*}
\left|\mathscr{M}_{1}\right\rangle=\rho_{11} e^{j \theta_{11}}\left|s_{1}\right\rangle+\rho_{12} e^{j \theta_{12}}\left|s_{2}\right\rangle+\rho_{13} e^{j \theta_{13}}\left|s_{3}\right\rangle \tag{28}
\end{equation*}
$$

with the following constraints on the modulus of each probability amplitudes (i.e., using the belief and plausibility constraints), we have $0.7 \leq \rho_{11}^{2} \leq 1,0 \leq \rho_{12}^{2} \leq 0.3,0 \leq \rho_{13}^{2} \leq 0.3$, $0.7 \leq \rho_{11}^{2}+\rho_{12}^{2} \leq 1,0.7 \leq \rho_{11}^{2}+\rho_{13}^{2} \leq 1,0.0 \leq \rho_{12}^{2}+\rho_{13}^{2} \leq 0.4$, and $\rho_{11}^{2}+\rho_{12}^{2}+\rho_{13}^{2}=1 \quad$ with $\quad \rho_{11} \geq 0, \rho_{12} \geq 0, \rho_{13} \geq 0, \quad$ and $\theta_{11}, \theta_{12}, \theta_{13} \in[-\pi, \pi]$.

The second example mass function $\left|\mathscr{M}_{2}\right\rangle$ over $\mathcal{S}$ is defined as follows:

$$
\begin{equation*}
\left|\mathscr{M}_{2}\right\rangle=\rho_{21} e^{j \theta_{21}}\left|s_{1}\right\rangle+\rho_{22} e^{j \theta_{22}}\left|s_{2}\right\rangle+\rho_{23} e^{j \theta_{23}}\left|s_{3}\right\rangle \tag{29}
\end{equation*}
$$

with the following constraints on the modulus of the probability amplitudes: $\rho_{21}=\sqrt{0.4}, \rho_{22}=\sqrt{0.5}$ and $\rho_{23}=\sqrt{0.1}$ with $\theta_{21}, \theta_{22}, \theta_{23} \in[-\pi, \pi]$.
2.2.3. Combining Quantum Mass Functions. There exist a number of approaches for obtaining the combined (average) description of the quantum mass functions which are extensions of classical methods [30]. This section presents an overview example of a method which is used to obtain such combined (average) description of the quantum mass function proposed by [29]. Let the frame of discernment be define by $\mathcal{S}=\left\{s_{1}, s_{2}, \ldots, s_{n}\right\}$ and similar to the previous section, let us define the following description of the quantum mass functions $\left|\mathscr{M}_{1}\right\rangle$ and $\left|\mathscr{M}_{2}\right\rangle$ of the two pieces of evidence $m_{1}$ and $m_{2}$ with the associated constraints on their modulus of the probability amplitudes. Let us also define $w_{1}$ and $w_{2}$ which correspond to weight assignments to these information evidences, Table 5.

Table 4: Example of Dempster's rule of combining two mass functions.

|  | $m_{1}$ | $m_{2}$ |  |
| :--- | :---: | :---: | :---: |
| $\emptyset$ |  |  | $m_{1} \oplus m_{2}$ |
| $s_{1}$ | 0.3 |  |  |
| $s_{2}$ | 0.4 | 0.312 |  |
| $s_{1}, s_{2}$ |  | 0.5 |  |
| $s_{3}$ |  | 0.1 |  |
| $s_{1}, s_{3}$ | 0.3 | 0.156 |  |
| $s_{2}, s_{3}$ | 0.4 |  |  |
| $s_{1}, s_{2}, s_{3}$ |  | 0.031 |  |

Table 5: Definition of two quantum mass functions and constraints in their probability amplitudes.

| First pieces of evedance | Second pieces of evedance |
| :--- | ---: |
| $\left\|\mathscr{M}_{1}\right\rangle=\sum_{i=1}^{n} \rho_{1 i} \theta^{j \theta_{1 i}}\left\|s_{i}\right\rangle$ Subject to: | $\left\|\mathscr{M}_{2}\right\rangle=\sum_{i=1}^{n} \rho_{2 i} e^{j \theta_{2 i}\left\|s_{i}\right\rangle \text { Subject to: }}$ |
| $\rho_{1 i} \geq 0, \theta_{1 i} \in[-\pi, \pi], \sum_{s_{i} \in \mathcal{S}} \rho_{1 i=1}^{2}$ | $\rho_{2 i} \geq 0, \theta_{2 i} \in[-\pi, \pi], \sum_{s_{i} \in \delta} \mathcal{S}_{2 i}^{2}=1$ |
| $\beta(A) \leq \sum_{s_{i} \in A} \rho_{1 i}^{2} \leq \pi(A), \forall A \subseteq \mathcal{S}$ | $\beta(A) \leq \sum_{s_{i} \in A} \rho_{2 i}^{2} \leq \pi(A), \forall A \subseteq \mathcal{S}$ |

A weighted combination of the two quantum mass functions can be defined as follows:

$$
\begin{equation*}
|\mathscr{M}\rangle=\frac{1}{\psi} \sum_{i=1}^{n} \rho_{i} e^{j \theta_{i}}\left|s_{i}\right\rangle \tag{30}
\end{equation*}
$$

$$
\begin{align*}
\psi^{2} & =\sum_{i=1}^{n}\left(w_{1}^{2} \rho_{1 i}^{2}+w_{2}^{2} \rho_{2 i}^{2}+2 w_{1} w_{2} \rho_{1 i} \rho_{2 i} \cos \left(\theta_{1 i}-\theta_{2 i}\right)\right) \forall i \in\{1,2, \ldots, n\}  \tag{31}\\
\rho_{i}^{2} & =w_{1}^{2} \rho_{1 i}^{2}+w_{2}^{2} \rho_{2 i}^{2}+2 w_{1} w_{2} \rho_{1 i} \rho_{2 i} \cos \left(\theta_{1 i}-\theta_{2 i}\right)  \tag{32}\\
\theta_{i} & =\arctan \left(\frac{w_{1} \rho_{1 i} \sin \theta_{1 i}+w_{2} \rho_{2 i} \sin \theta_{2 i}}{w_{1} \rho_{1 i} \cos \theta_{1 i}+w_{2} \rho_{2 i} \cos \theta_{2 i}}\right) \tag{33}
\end{align*}
$$

The probability amplitudes of the basis state $\left|s_{i}\right\rangle$ in $\mathscr{M}_{1}$ and $\mathscr{M}_{2}$ are $\rho_{1 i} e^{j \theta_{1 i}}$ and $\rho_{2 i} e^{j \theta_{2 i}}$, respectively. The quantum averaging makes the weighted average of $\rho_{1 i} e^{j \theta_{1 i}}$ and $\rho_{2 i} e^{j \theta_{2 i}}$ in order to generate the combined probability amplitude of $\left|s_{i}\right\rangle$, i.e., $\rho_{i} e^{j \theta_{i}}=w_{1} \rho_{1 i} e^{j \theta_{1 i}}+w_{2} \rho_{2 i} e^{j \theta_{2 i}}$. This in turn can further be defined as the square of the probability amplitude defined in the relationship: $\rho_{i}^{2}=w_{1}^{2} \rho_{1 i}^{2}+w_{2}^{2} \rho_{2 i}^{2}+2 w_{1}$ $w_{2} \rho_{1 i} \rho_{2 i} \cos \left(\theta_{1 i}-\theta_{2 i}\right)$.

The term $\epsilon=2 \rho_{1 i} \rho_{2 i} \cos \left(\theta_{1 i}-\theta_{2 i}\right)$ define in equation (32) is referred to as the interference term between $\left|\mathscr{M}_{1}\right\rangle$ and $\left|\mathscr{M}_{2}\right\rangle$ on $\left|s_{i}\right\rangle$ (Appendix A-E). The phase difference $\mid \theta_{1 i}-$ $\theta_{2 i} \mid$ represents the interaction of the mass functions on the same sensed event $\left|s_{i}\right\rangle$ which cannot be determined through individual mass function. For the case when $\left|\theta_{1 i}-\theta_{2 i}\right|=(\pi / 2)$, the above averaging reduced to the classical averaging techniques. With reduction of this angle, the effect of interference between the two mass function increases. When $\left|\theta_{1 i}-\theta_{2 i}\right|=0$, this can imply that there exist a maximum information for mutual support between the two mass functions. For the purpose of demonstration, in the following example we assume that there exists maximum information and mutual support between the
mass functions. The quantum averaging of equation (30) can be written as follows:

$$
\begin{equation*}
|\mathscr{M}\rangle=\frac{1}{\psi} \sum_{i=1}^{n} \rho_{i}\left|s_{i}\right\rangle \tag{34}
\end{equation*}
$$

subject to the constraints associated with the probability amplitudes of each of the mass functions and simplified constraint relationships defined in equations (31) and (32), or

$$
\begin{align*}
& \psi^{2}=\sum_{i=1}^{n} w_{1}^{2} \rho_{1 i}^{2}+w_{2}^{2} \rho_{2 i}^{2}+2 w_{1} w_{2} \rho_{1 i} \rho_{2 i} \forall i \in\{1,2, \ldots, n\} \\
& \rho_{i}^{2}=w_{1}^{2} \rho_{1 i}^{2}+w_{2}^{2} \rho_{2 i}^{2}+2 w_{1} w_{2} \rho_{1 i} \rho_{2 i} \tag{35}
\end{align*}
$$

The following example presents a demonstration of above algorithm in order to combine two quantum presentation of mass functions. Let $\mathcal{S}=\left(s_{1}, s_{2}, s_{3}\right)$ be the frame of discernment be where we have defined the two mass functions $m_{1}$ and $m_{2}$, where $m_{1}\left(s_{1}\right)=0.7, m_{1}\left(s_{2}\right)=0.1$, $m_{1}\left(s_{1}, s_{2}, s_{3}\right)=0.2, \quad m_{2}\left(s_{1}\right)=0.5$, and $\quad m_{2}\left(s_{2}, s_{3}\right)=0.5$.

Table 6: Numerical description example of two quantum mass functions.

| $\sum_{i=1}^{3} \rho_{11} e^{j \theta_{1 i}}\left\|s_{i}\right\rangle$ | $\sum_{i=1}^{3} \rho_{2 i 2} e^{j \theta_{2 i}\left\|s_{i}\right\rangle}$ |
| :--- | ---: |
| Subject to: | Subject to: |
| $0.7 \leq \rho_{11}^{2} \leq 0.9$ | $0.5 \leq \rho_{21}^{2} \leq 1$ |
| $0.1 \leq \rho_{12}^{2} \leq .3$ | $0 \leq \rho_{22}^{2} \leq 0.5$ |
| $0 . \leq \rho_{13}^{2} \leq 0.2$ | $0 . \leq \rho_{23}^{2} \leq 0.5$ |
| $0.7 \leq \rho_{11}^{2}+\rho_{12}^{2} \leq 0.9$ | $0.5 \leq \rho_{12}^{2}+\rho_{22}^{2} \leq 1$ |
| $0.7 \leq \rho_{11}^{2}+\rho_{13}^{2} \leq 1$ | $0.5 \leq \rho_{21}^{2}+\rho_{23}^{2} \leq 1$ |
| $0.1 \leq \rho_{12}^{2}+\rho_{13}^{2} \leq 0.2$ | $0 . \leq \rho_{22}^{2}+\rho_{23}^{2} \leq 0.5$ |
| $\rho_{11}^{2}+\rho_{12}^{2}+\rho_{13}^{2}=1$ | $\rho_{21}^{2}+\rho_{22}^{2}+\rho_{23}^{2}=1$ |



Figure 1: Probability path diagram starting with state initial $\mathcal{S}$ mapped to probabilistic measurements of $\mathscr{M}$ which is then mapped to final probabilistic state $\mathscr{D}$. This diagram encapsulates the double-slit experiment which was proposed by [37]. In the example of this section, the state probability $\mathcal{S}$, or $|\Psi\rangle$ can have some measure or information available to it, namely, $|\mathrm{Lo}\rangle_{m}$ and $|\mathrm{La}\rangle_{m}$ while deciding between the two choices of motion final movements, namely, $|L o\rangle$ and $|L a\rangle$.

Following the description of Table 5 the quantum representation of these mass functions can be written as follows:

Using the definition of the combined quantum mass functions (equations (34) and (35)) with the weight $w_{1}=w_{2}=0.5$, we can define an instant of the combined quantum mass functions by satisfying the constraints associated with the individual and combined probability amplitudes.

The quantum combination of the two mass functions can be written as follows:

$$
\begin{equation*}
|\mathscr{M}\rangle=\frac{1}{\psi}\left(\rho_{1}\left|s_{1}\right\rangle+\rho_{2}\left|s_{2}\right\rangle+\rho_{3}\left|s_{3}\right\rangle\right) \tag{36}
\end{equation*}
$$

with the following coefficients (Table 6)

$$
\begin{array}{lll}
\rho_{11}^{2}=0.8 ; & \rho_{21}^{2}=0.6 ; & \rho_{1}^{2}=(0.5)^{2}(0.8)+(0.5)^{2}(0.6)+2(0.5)(0.5)(0.8)(0.6), \\
\rho_{12}^{2}=0.1 ; & \rho_{22}^{2}=0.3 ; & \rho_{2}^{2}=(0.5)^{2}(0.1)+(0.5)^{2}(0.3)+2(0.5)(0.5)(0.1)(0.3),  \tag{37}\\
\rho_{13}^{2}=0.1 ; & \rho_{23}^{2}=0.2 ; & \rho_{3}^{2}=(0.5)^{2}(0.1)+(0.5)^{2}(0.2)+2(0.5)(0.5)(0.1)(0.2) .
\end{array}
$$

In the expansion of the above equation (and as highlighted in equations (31) and (35)), the term $2 w_{1} w_{2} \rho_{1 i} \rho_{2 i}$ represents the interference term in the quantum combination of mass functions.

## 3. Discussion and Conclusions

Managing uncertainties in a decision-making process which needs to incorporate many sources of information/sensing
as a part of the process is a challenging task. The difficulty of this challenge is further compounded when, for example, dealing with a design of an autonomous decision-making system involving multiple sensing modalities mounted on robots responsible for monitoring movements and activities of people. Historically, a number of decision-making frameworks have been proposed in the literature based on deterministic reasoning and various extensions of Bayesian/ fuzzy reasoning. This paper presents an overview of examples proposed in the literature in extending the basis of the classical framework of decision-making with various interpretations of quantum mechanics. These extensions which may be viewed as parsimonious models can offer frameworks to further study various similarities, limitations of the classical decision-making methods. In this paper, we focus only on review of two of the classical methods, namely, the Bayesian network and Dempster-Shafer theory. For both approaches, we first present a review of each method, followed by their proposed quantum counterpart. The review of the quantum extensions presented here also only focuses on the probability amplitude part defined in quantum framework. The paper highlights through various walk-through examples on where the quantum framework differ from the classical one. There still remain many questions regarding the feasibility of the quantum framework within the decision-making paradigms. Recently, there have been many initiatives which try to explore relationships between human way of reasoning, fuzzy concepts, and quantum mechanics.

## Appendix

## A. Quantum State

Outcomes or events which are associated with random phenomenon or experiments are defined within the sample space which is usually represented by $\Omega$. As a descriptive example, let us consider a planar movement of an object is constrained to move in the horizontal plane and we are interested in tracking this object. The sample space can be defined as the instantiation (probability) of movements that the object can take which can be either of the independent directions of movements namely, movement to the left/right (lateral) (La) or movement to the left/right (longitudinal) (Lo) incremental (discrete) movements, e.g., $\Omega=\{\mathrm{La}, \mathrm{Lo}\}$. We refer to these instances of example movements as mutually exclusive events. This means that the choices that the tracked object has in selecting either direction is mutually exclusive or does not dependent on each other. Or in another words, at any instant, the decision that the object makes are in a superposition of these basic primitives (states).

In the quantum framework, outcomes and events are defined within the finite dimension Hilbert space $\mathscr{H}$. In general, this would be a vector space of complex numbers embedded with the structure of inner product which allows defining a distance function. Similar to Euclidean space, this space is also defined within the span of the basis vectors which are orthonormal to each other. For the working example of this paper, this can be the choices of movement if
an object in a room where its space is defined as $\mathscr{H}=\{|\mathrm{La}\rangle,|\mathrm{Lo}\rangle\}(|()$.$\rangle is a column vector and is referred$ to as a ket or, $|\mathrm{La}\rangle=\left[\begin{array}{ll}1 & 0\end{array}\right]^{T}$ and $|\mathrm{Lo}\rangle=\left[\begin{array}{ll}0 & 1\end{array}\right]^{T}$ are the two bases that defines the instantaneous choices of movements.

Outcomes and events are given geometrical meaning. Uncertainties in selection for the possible direction of movements by the object is represented by a state vector defined by $|\Psi\rangle$, which can capture the occurrence of all events [31]. This would be in contrary to classical probability theory where at any instance, the occurrence of each event is represented in a sequential order belonging to a set. In the quantum framework, we can represent all possible occurrences of events at the same time through vector representation. Similar to classical theory, the mutual exclusiveness of events is represented by orthonormal vectors. In our example, an instance of state $|\Psi\rangle$ can be defined as a superposition of choices for the probability of movements.

$$
\begin{equation*}
|\Psi\rangle=\frac{e^{i \theta_{\mathrm{La}}}}{\sqrt{4}}|\mathrm{La}\rangle+\frac{\sqrt{3} e^{i \theta_{\mathrm{Lo}}}}{\sqrt{4}}|\mathrm{Lo}\rangle \tag{A.1}
\end{equation*}
$$

where $\theta_{\mathrm{La}}$ and $\theta_{\mathrm{Lo}}$ are the instances of the relative position angles that the state can make with respect to each of the basis in the unit sphere (ball) description. $\left(e^{i \theta_{\mathrm{La}}} / \sqrt{4}\right)$ and $\left(\sqrt{3} e^{i \theta_{\mathrm{Lo}}} / \sqrt{4}\right)$ are generally referred to as probability amplitudes along each of the basis space. These amplitudes are now can be used to define the probability of subject in selecting either of the directions. The magnitude of these probabilities is obtained by squaring the associated projected magnitude along a given basis subspace through utilization of Born's rule (see the following). To compute such probability, each of the probability amplitudes can be multiplied by its corresponding complex conjugate. For example, the probability (Pr) of the tracked subject moving in the lateral direction, i.e., $\operatorname{Pr}(\mathrm{La})$, at the instant of measurement can be obtained as follows [18, 32]:

$$
\begin{align*}
\operatorname{Pr}(\mathrm{La}) & =\left|\frac{e^{i \theta_{\mathrm{La}}}}{\sqrt{4}}\right|^{2} \\
& =\left(\frac{e^{i \theta_{\mathrm{La}}}}{\sqrt{4}}\right) \cdot\left(\frac{e^{i \theta_{\mathrm{La}}}}{\sqrt{4}}\right)^{*}  \tag{A.2}\\
& =\frac{e^{i \theta_{\mathrm{La}}}}{\sqrt{4}} \cdot \frac{e^{-i \theta_{\mathrm{La}}}}{\sqrt{4}} \\
& \left.=e^{\left(\theta_{\mathrm{La}}-\theta_{\mathrm{La}}\right.}\right)\left(\frac{1}{\sqrt{4}}\right)^{2}=0.25
\end{align*}
$$

The quantum state description, such as the one described above, is rotating in a unit circle (as seen from the quantum theory framework) and requires by one of its axioms to have the sum of the squared magnitudes of each amplitude (projections) to be equal to one, or

$$
\begin{equation*}
\left|\frac{e^{i \theta_{\mathrm{La}}}}{\sqrt{4}}\right|^{2}+\left|\frac{\sqrt{3} e^{i \theta_{\mathrm{Lo}}}}{\sqrt{4}}\right|^{2}=1 \tag{A.3}
\end{equation*}
$$

This axiom is called the normalization axiom (Unitarily) [33] and also corresponds to the classical probability theory that constraint that the probability of all events in a sample space to sum-up to one. Give the quantum state $|\Psi\rangle$, the probability of an event along any of the independent subspaces can also be obtained by squaring the magnitude of the projection of the state onto the corresponding subspace. As an example, for the quantum state defined above, projection operators $P_{\mathrm{La}}$ and $P_{\mathrm{Lo}}$ of the state onto each of the subspaces are defined by exterior product, or

$$
\begin{align*}
& P_{\mathrm{La}}=|\mathrm{La}\rangle\langle\mathrm{La}|=\left[\begin{array}{ll}
1 & 0 \\
0 & 0
\end{array}\right],  \tag{A.4}\\
& P_{\mathrm{Lo}}=|\mathrm{Lo}\rangle\langle\mathrm{Lo}|=\left[\begin{array}{ll}
0 & 0 \\
0 & 1
\end{array}\right],
\end{align*}
$$

which are the outer product (tensor product) of the basis states (subspaces) $|\mathrm{La}\rangle$ and $|\mathrm{Lo}\rangle$, respectively. The square measure of the length of the vector (i.e., probability measure), which is the square of the resultant projection of the superposition state vector $|\Psi\rangle$ to a desired subspace can now be computed. For example, the projection of the state onto subspace $|\mathrm{Lo}\rangle$ is [32]

$$
\begin{align*}
\left.\left|P_{\mathrm{Lo}}\right| \Psi\right\rangle\left.\right|^{2} & =\left(P_{\mathrm{Lo}}|\Psi\rangle\right)^{*}\left(P_{\mathrm{Lo}}|\Psi\rangle\right)  \tag{A.5}\\
& =|\Psi\rangle P_{\mathrm{Lo}}|\Psi\rangle
\end{align*}
$$

Using complex angle description and expansion, the probability of the along the longitudinal direction is

$$
\begin{align*}
\operatorname{Pr}(\mathrm{Lo}) & \left.=\left|P_{\mathrm{Lo}}\right| \Psi\right\rangle\left.\right|^{2} \\
& =\left|\frac{\sqrt{3} e^{i \theta_{\mathrm{Lo}}}}{\sqrt{4}}\right|^{2}=0.75 . \tag{A.6}
\end{align*}
$$

The collapse of the state $|\Psi\rangle$ (or the wave function) onto a subspace gives the probability of associating a measurement (or observation) along the subspace. An interpretation of the above can be that until the autonomous object committee to a direction of movement, it is in the state of superposition of all the choices (subspaces) (for example, the probability of choosing a movement along the Lo direction would be $0.75 \%$ ).

## B. Lüder's Rule

Probabilities are assigned to histories or sequences of events based on an interpretation Lüder's Rule [34]. This rule states that in order to compute the probability of an ordered events, say events $\mathbf{E}_{1}$ and $\mathbf{E}_{2}$ which are spanned by two different basis in two spaces $\mathscr{H}_{1}$ and $\mathscr{H}_{2}$, one should compute the probability of first observing/measuring the event $\mathbf{E}_{1}$ and then calculating the probability of occurrence of $\mathbf{E}_{1}$, followed by $\mathbf{E}_{2}$. The following extensions presents the associated sequences for such computation which are based on the results presented in [35, 36].

Let $\mathbf{E}_{1}$ be the first sensor having basis of $|\mathrm{La}\rangle_{\mathscr{H}_{1}}$ and $|\mathrm{Lo}\rangle_{\mathscr{H}}$ that are spanned in $\mathscr{H}_{1}$ and let $\mathbf{E}_{2}$ be the sensor which has spanned basis of $|\mathrm{La}\rangle_{\mathscr{H}_{2}}$ and $|\mathrm{Lo}\rangle_{\mathscr{H}_{2}}$ spanned in
$\mathscr{H}_{2}$. After observing/measuring the movement of the subject $|\Psi\rangle$ corresponding to the event $\mathbf{E}_{1}$, we can compute the following probability which would be also the new definition of the revised state of the subject, i.e., $|\Psi\rangle_{\mathscr{H}_{1}}$ :

$$
\begin{equation*}
\left.\operatorname{Pr}(\mathrm{La})_{\mathscr{H}_{1}}=\left|P_{L a_{\mathscr{H}_{1}}}\right| \Psi\right\rangle\left.\right|^{2} \tag{B.1}
\end{equation*}
$$

with the revised state

$$
\begin{equation*}
\left|\Psi_{\mathscr{H}_{1}}\right\rangle=\frac{P_{\mathrm{La}_{\mathscr{H}_{1}}}|\Psi\rangle}{\left.\left|P_{\mathrm{La}_{\mathscr{H}_{1}}}\right| \Psi\right\rangle \mid} \tag{B.2}
\end{equation*}
$$

The probability of an event $\mathbf{E}_{2}$ given that we have observed/measured event $\mathbf{E}_{1}$ is given by square of the projection of the revised state vector $|\Psi\rangle_{\mathscr{H}_{1}}$ corresponding to event $\mathbf{E}_{1}$, onto the subspace related to event $\mathbf{E}_{2}$. This is equivalent to $\left.\operatorname{Pr}\left(\mathbf{E}_{2} \mid \mathbf{E}_{1}\right)=\left|P_{\mathrm{Lo}_{\mathscr{H}}}\right| \Psi_{\mathscr{H}_{1}}\right\rangle\left.\right|^{2}$. According to Lüder s rule, the probability of $\mathbf{E}_{2}$, followed by $\mathbf{E}_{1}$ is given by $\mathrm{P}\left(\mathbf{E}_{1}\right) \cdot \mathrm{P}\left(\mathbf{E}_{2} \mid \Psi_{\mathscr{H}_{1}}\right)$, or

$$
\begin{align*}
& \left.\left.\operatorname{Pr}\left(\mathbf{E}_{1}\right) \cdot \operatorname{Pr}\left(\mathbf{E}_{2} \mid \Psi_{\mathscr{H}_{1}}\right)=\left|P_{L_{\mathscr{O}_{1}}}\right| \Psi\right\rangle\left.\right|^{2} \cdot\left|P_{\text {Lo }_{\mathscr{F}_{2}}}\right| \Psi\right\rangle\left._{\mathscr{H}_{1}}\right|^{2} \\
& \left.=\left|P_{L a_{x_{1}}}\right| \Psi\right\rangle\left.\right|^{2} \cdot\left|P_{L_{x_{2}}} \frac{P_{\mathrm{La}_{x_{1}}}|\Psi\rangle}{}\right| P_{\mathrm{La}_{\varkappa_{1}}}|\Psi\rangle| |^{2} \\
& \left.\left.=\left|P_{L a_{Y_{1}}}\right| \Psi\right\rangle\left.\right|^{2} \cdot \frac{1}{\left.\left|P_{L a_{x_{1}}}\right| \Psi\right\rangle\left.\right|^{2}} \cdot\left|P_{L o_{x_{2}}} P_{L a_{Y_{1}}}\right| \Psi\right\rangle\left.\right|^{2} \\
& \left.=\left|P_{\mathrm{Lo}_{\mathscr{K}_{2}}} P_{\mathrm{La}_{\varkappa_{1}}}\right| \Psi\right\rangle\left.\right|^{2}, \tag{B.3}
\end{align*}
$$

the above equation is the projection of the initial state $|\Psi\rangle$ onto the subspace of the first sensor measurement through the projection operator $P_{\mathrm{La} \mathscr{\mathscr { C }}}$. This is then results in the new state $|\Psi\rangle_{\mathscr{H}}$ which is then followed by the projection of this new state onto the subspace of the second sensor $|\mathrm{Lo}\rangle_{\mathscr{H}_{2}}$ given be the projector operator $P_{\mathrm{Lo}_{\mathscr{H}_{2}}}$.

## C. Probability Path Diagrams

Path diagrams are very important visualization approach for describing the interconnection and dependencies between set of event uncertainty variables. It is a description which was originally inspired by the effect of interference observed using the famous double slit experiment and the associated probability distributions [36, 37].

Using the example set-up of this paper, an interpretation of the interference effect can be stated as follows. Given the state of the tracked subject as described above $\mathcal{S}$, suppose that the subject is at the instance of selecting between the directions of movement $\mathscr{D}$ and has available to it information/measurement $\mathscr{M}$ (with a probability distribution) (here, by default we are assuming that outcome of measurement from any sensing modality has associate to it some uncertainty (probability distribution)) that can be used to assist in deciding between the choices in the incremental movements. Given this information, the subject will then
decide on whether or not to use it in order to decide on the direction of movement. This interpretation can also be extended to the case where the subject decides on the direction of movement not knowing (or having access to) to such available information or simply ignores the availability and presence of such distribution of the information [31, 36].

In order to capture and compute the probability model of the subject given various path configurations which can arise from the above scenarios (see Figure 1). Feynman [37] devices number of rules for asserting the probability of transition between various nodes. For example, the first rule for asserting the probability for a sequence of events along a single path state that, the probability amplitude of the path is the product of the amplitudes of each of the transition paths. This is equivalent to the classical product of the conditional probability $P(\cdot \mid \cdot)$ of transitions between each of the node along a single path, or

$$
\begin{align*}
\mathcal{S} \longrightarrow & \mathscr{M}_{1} \longrightarrow \mathscr{D}_{1} \\
& \operatorname{Pr}\left(\mathscr{M}_{1} \mid \mathcal{S}\right) \cdot \operatorname{Pr}\left(\mathscr{D}_{1} \mid \mathscr{M}_{1}\right), \tag{C.1}
\end{align*}
$$

or the computation of the probability using the associate projection operator $P_{(.)}$can be written as follows:

$$
\begin{align*}
& \left.\left|P_{\mathrm{Lo}_{m}}\right| \Psi\right\rangle\left.\right|^{2} \cdot \mid\left. P_{\mathrm{lo}}\left(P_{\mathrm{Lo}_{m}}|\Psi\rangle\right)\right|^{2} \text { or } \\
& \left.\left.\quad\left|\psi_{\mathscr{M}_{1}}\right| \delta\right|^{2} \cdot\left|\psi_{\mathscr{D}_{1} \mid}\right| \mathscr{M}_{1}\right|^{2}, \tag{C.2}
\end{align*}
$$

where the symbol $\psi$ is used to represent complex probability amplitude. For such single paths, the probability of the path is similar to the Markov model where the square of amplitudes replacing the magnitudes for the conditional probabilities.

An indistinguishable path is referred to the case when the subject arrives on decision on moving either to the lateral direction or longitudinal direction having some measure of information available to it in regard to the selection of the direction. But the subject does not consider using these information (or not being aware of the presence of such information). Referring to Figure 1, an example can correspond to the case where the subject starts at initial state $\mathcal{S}$ and arrives at the final state $\mathscr{D}_{1}$ by transiting from multiple possible paths without knowing for certain which path was taken to reach the goal state.

Quantum probability theory states that when the path is undistinguishable, then the goal state can be reached through a superposition of paths trajectories. This is known as Feynman's second rule. It states, for example, that the amplitude of transition from an initial state $\mathcal{S}$ to the final state $\mathscr{D}_{1}$, taking multiple indistinguishable paths, is given by the sum of all amplitudes for each path. This rule is in accordance with the law of total amplitude and the probability and computed by taking the squared magnitude of this sum, or

$$
\begin{equation*}
\left.\operatorname{Pr}\left(\mathcal{S} \longrightarrow \mathscr{D}_{1}\right)=\left|P_{\mathrm{Lo}_{m}}\right| \Psi\right\rangle \cdot P_{\mathrm{Lo}}\left(P_{\mathrm{Lo}_{m}}|\Psi\rangle\right)+\left.P_{\mathrm{La}_{m}}|\Psi\rangle \cdot P_{\mathrm{Lo}}\left(P_{\mathrm{La}_{m}}|\Psi\rangle\right)\right|^{2} \tag{C.3}
\end{equation*}
$$

Contrary to the above decomposition, if the above mentioned paths were observed/measured, the definition of the quantum probability theory is the same as Markov model (which is also known as Feynman's third rule). The rule states that the probability amplitude of observed/measured
multiple path trajectories is the sum of the amplitudes of each individual paths. Following Born's rule, the probability of each path is the squared magnitude along the path. For the case example mentioned above, this can be written as follows:

$$
\begin{equation*}
\left.\left.\operatorname{Pr}\left(\mathcal{S} \longrightarrow \mathscr{D}_{1}\right)=\operatorname{Pr}\left(\mathscr{M}_{1} \mid \mathcal{S}\right) \cdot \operatorname{Pr}\left(\mathscr{D}_{1} \mid \mathscr{M}_{1}\right)+\operatorname{Pr}\left(\mathscr{M}_{2} \mid \mathcal{S}\right) \cdot=\operatorname{Pr}\left(\mathscr{D}_{1} \mid \mathscr{M}_{2}\right)\left|P_{\mathrm{Lo}_{m}}\right| \Psi\right\rangle\left.\cdot P_{\mathrm{Lo}}\left(P_{\mathrm{Lo}_{m}}|\Psi\rangle\right)\right|^{2}+\left|P_{\mathrm{La}_{m}}\right| \Psi\right\rangle\left.\cdot P_{\mathrm{Lo}^{\prime}}\left(P_{\mathrm{La}_{m}}|\Psi\rangle\right)\right|^{2} \tag{C.4}
\end{equation*}
$$

## D. Born's Rule

The following interpretation of the Born rule is partly based on the results presented in [38]. Let $\mathscr{M}$ be a quantummechanical model of a system which is observable and can be represented by a self-adjoint operator on $\mathscr{H}$ with inner product $\langle | \cdot\rangle$. Assume that $\mathscr{M}$ has nondegenerated discrete spectrum which in our case of self-adjoint operator which implies that $\mathscr{M}$ has an orthonormal basis of eigenvectors $e_{i}$ with corresponding eigenvalues $\lambda_{i}$, i.e., $\mathscr{M} e_{i}=\lambda_{i} e_{i}$.

A fundamental assumption underlying the Born rule is that in $\mathscr{M}$ and measurement of the observable $\mathscr{M}$ will
produce one of its eigenvalues $\lambda_{i}$. Let $\Psi \in \mathscr{H}$ be a unit vector, then the Born rule states: if the system in a state $\Psi$, then the probability $\operatorname{Pr}\left(\mathscr{M}=\lambda_{i},|\Psi\rangle\right)$ that the eigenvalue $\lambda_{i}$ of $\mathscr{M}$ is found when $\mathscr{M}$ is measured is

$$
\begin{equation*}
\left.\operatorname{Pr}\left(\mathscr{M}=\lambda_{i},|\Psi\rangle\right)=\left|e_{i}\right| \Psi\right\rangle\left.\right|^{2} \tag{D.1}
\end{equation*}
$$

In other words, if $\left.|\Psi\rangle=\sum_{i=1} \omega_{i} e_{i}\right\rangle\left(\right.$ with $\left.\sum_{i=1}\left|\omega_{i}\right|^{2}=1\right)$, then,

$$
\begin{equation*}
\operatorname{Pr}\left(\mathscr{M}=\lambda_{i},|\Psi\rangle\right)=\left|\omega_{i}\right|^{2} \tag{D.2}
\end{equation*}
$$

## E. Interference Effects

As stated previously, the relationship between the classical probability density function and a quantum probability amplitude is given by Born's rule, or:

$$
\begin{equation*}
\operatorname{Pr}(A)=\left|e^{i \theta_{A}} \psi(A)\right|^{2} \tag{E.1}
\end{equation*}
$$

$\left|e^{i \theta_{A}} \psi(A)\right|^{2}$ is the square magnitude of a complex amplitude which is obtained as follows: $\left|e^{i \theta_{A}} \psi(A)\right|^{2}=e^{i \theta_{A}} \psi$ (A) $e^{-i \theta_{A}} \psi(A)$.

Given a set of disjointed events $A_{1}, A_{2}, \cdots A_{n}$, the law of total probability can be formulated as follows [39]:

$$
\begin{equation*}
\operatorname{Pr}(B)=\sum_{i=1}^{n} \operatorname{Pr}\left(A_{i}\right) \operatorname{Pr}\left(B \mid A_{i}\right) \text { where } \sum_{i=1}^{N} A_{i}=1 \tag{E.2}
\end{equation*}
$$

The quantum probability equivalent of the above total probability can be written as follows:

$$
\begin{equation*}
\operatorname{Pr}(B)=\left|\sum_{i=1}^{n} e^{i \theta_{i}} \Psi\left(A_{i}\right) \Psi\left(B \mid A_{i}\right)\right|^{2} \text {, with } \sum_{i=1}^{n}\left|e^{i \theta_{i}} \psi\left(A_{i}\right)\right|^{2}=1 \tag{E.3}
\end{equation*}
$$

Multiplying the expression with its complex conjugate and simplifying, the expression for $\mathrm{P}(B)$ can be written as follows:

$$
\begin{equation*}
\mathrm{P}(B)=\sum_{i=1}^{n}\left|\psi\left(A_{i}\right) \psi\left(B \mid A_{i}\right)\right|^{2}+2 \sum_{i=1}^{n-1} \sum_{j=i+1}^{n}\left|\psi\left(A_{i}\right) \psi\left(B \mid A_{i}\right)\right|\left|\psi\left(A_{j}\right) \psi\left(B \mid A_{j}\right)\right| \cos \left(\theta_{i}-\theta_{j}\right) \tag{E.4}
\end{equation*}
$$

As it can be seen from the above, when $\cos \left(\theta_{i}-\theta_{j}\right)=0$, the quantum probability theory converges to its classical counterpart because the interference term will be zero. For nonzero values, this interference term can affect destructively the classical probability or constructively.

## Data Availability

The data that used to support the findings of this study are available from the corresponding author upon request.

## Conflicts of Interest

The author declares that there are no conflicts of interest regarding the publication of this paper.

## Acknowledgments

This work was partially supported by the Discovery Grant through Natural Sciences and Engineering Research Council of Canada and Simon Fraser University.

## References

[1] Y. Tang, C. Zhao, J. Wang et al., "Perception and navigation in autonomous systems in the era of learning: a survey," IEEE Transactions on Neural Networks and Learning Systems, vol. 8, pp. 1-21, 2022.
[2] M. Mast, B. Graf, F. Weisshardt et al., Design Of the HumanRobot Interaction For a Semi-autonomous Service Robot To Assist Elderly People, Ambient Assisted Living, Advanced Technologies and Societal Change, Springer, Berlin, Germany, 2015.
[3] H. Chen, Y. Wen, M. Zhu et al., "From automation System to autonomous system: an architecture perspective," Journal of Marine Science and Engineering, vol. 9, no. 6, p. 645, 2021.
[4] J. Sifakis, "Autonomous systems-an architectural characterization," in Models, Languages, and Tools for Concurrent and Distributed Programming, M. Boreale, F. Corradini, M. Loreti, and R. Pugliese, Eds., vol. 11665, Berlin, Germany, Springer, 2019.
[5] J. K. W. Ho and J. F. Hoorn, "Quantum affective processes for multidimensional decision-making," Scientific Reports, vol. 12, no. 1, Article ID 20468, 2022.
[6] X. Gao, L. Pan, D. Pelusi, and Y. Deng, "Fuzzy Markov decision-making model for interference effects," IEEE Transactions on Fuzzy Systems, vol. 31, no. 1, pp. 199-212, 2023.
[7] P. L. Gentili, "Establishing a new Link between fuzzy logic, neuroscience, and quantum Mechanics through bayesian probability: Perspectives in artificial Intelligence and unconventional computing," Molecules, vol. 26, no. 19, p. 5987, 2021.
[8] T. Koski and J. Noble, Bayesian Networks: An Introduction, Wiley, New York, NY, USA, 2009.
[9] S. Russel and P. Norvig, Artificial Intelligence: A Modern Approach, Pearson, New York, NY, USA, 2010.
[10] S. Challa, M. R. Morelande, and R. J. Evans, Fundamentals Of Object Tracking, Cambridge University Press, Cambridge, MA, USA, 2011.
[11] S. Payandeh, Visual Tracking In Conventional Minimally Invasive Surgery, CRC Press, Boca Raton, FL, USA, 2017.
[12] S. Payandeh, "On importance sampling in sequential Bayesian tracking of elderly," in Proceedings of the 2016 Annual IEEE Systems Conference (SysCon), pp. 1-6, Orlando, FL, USA, April 2016.
[13] C. M. Caves, C. A. Fuchs, and R. Schack, "Quantum Probabilities as bayesian probabilities," Physical Review A, vol. 65, no. 2, Article ID 22305, 2002.
[14] C. G. Timpson, "Quantum bayesianism: a study," Studies In History and Philosophy of Modern Physics, vol. 39, no. 3, pp. 579-609, 2008.
[15] C. Moreira and A. Wichert, "Quantum-like bayesian Networks for modeling decision making," Frontiers in Psychology, vol. 7, pp. 11-2016, 2016.
[16] C. Moreira, P. Tiwari, H. M. Pandey, P. Bruza, and A. Wichert, "Quantum-like influence Diagrams for decision-making," Neural Networks, vol. 132, pp. 190-210, 2020.
[17] M. De Oliveira and L. S. Barbosa, "Quantum bayesian decision-making," Foundations of Science, vol. 28, no. 1, pp. 21-41, 2023.
[18] P. C. Moreira, Quantum probabilistic graphical models for cognition and decision, Ph.D Thesis, Universidade De Lisboa, Lisbon, Portugal, 2017.
[19] A. P. Dempster, "Upper and lower probabilities induced by a multivalued mapping," The Annals of Mathematical Statistics, vol. 38, no. 2, pp. 325-339, 1967.
[20] G. Shafer, A Mathematical Theory Of Evidence, Princeton University Press, Princeton, NY, USA, 1976.
[21] A. Perez, H. Tabia, D. Declercq, and A. Zanotti, "Using the Conflict in dempster-shafer evidence Theory as a rejection Criterion in classifier output Combination for 3D human action recognition," Image and Vision Computing, vol. 55, no. 2, pp. 149-157, 2016.
[22] K. Sentz and S. Ferson, "Combination Of Evidence In Dempster-Shafer Theory," Technical Report, Sandia National Laboratories, Albuquerque, NM, USA, 2002.
[23] N. Zlatoff, B. Tellez, and A. Baskurt, "Region-based perceptual grouping: a cooperative approach based on Dempster-Shafer theory," SPIE Proceedings, vol. 6064, 2006.
[24] G. Shafer, "Dempster's Rule of combination," International Journal of Approximate Reasoning, vol. 79, pp. 26-40, 2016.
[25] X. Gao and Y. Deng, "Quantum Model of mass function," International Journal of Intelligent Systems, vol. 35, no. 2, pp. 267-282, 2020.
[26] F. Xiao, "CED: a distance for complex mass functions," IEEE Transactions on Neural Networks and Learning Systems, vol. 32, no. 4, pp. 1525-1535, 2021.
[27] X. Deng, T. Iwagawa, M. Fukushima, Y. Suzuki, and S. Watanabe, "Setd1a plays pivotal roles for the survival and proliferation of retinal progenitors via histone modifications of Uhrf1," Investigative Ophthalmology and Visual Science, vol. 62, pp. 1-6, 2021.
[28] A. Vourdas, "Quantum Probabilities as dempster-shafer Probabilities in the Lattice of subspaces," Journal of Mathematical Physics, vol. 55, no. 8, 2014.
[29] X. Deng, S. Xue, and W. Jiang, "A novel quantum Model of mass Function for uncertain information fusion," Information Fusion, vol. 89, no. 2023, pp. 619-631, 2023.
[30] R. Yager, "On the Fusion of non-independent belief structures," International Journal of General Systems, vol. 38, no. 5, pp. 505-531, 2009.
[31] S. Payandeh, "A quantum Model for decision Support in a sensor network," in Intelligent Systems and Applications, K. Arai, S. Kapoor, and R. Bhatia, Eds., vol. 1252, Cham, Switzerland, Springer, 2020.
[32] M. Fannes, "An Introduction to quantum probability," in Theoretical Foundations of Quantum Information Processing and Communication, E. Brüning and F. Petruccione, Eds., vol. 787, Berlin, Germany, Springer, 2020.
[33] L. Hardy, "Quantum theory from five reasonable axioms," 2001, https://arxiv.org/abs/quant-ph/0101012.
[34] G. Niestegge, "An Approach to quantum Mechanics via conditional probabilities," Foundations of Physics, vol. 38, no. 3, pp. 241-256, 2008.
[35] M. Nielsen and I. Chuang, Quantum Computation And Quantum Information: 10th Anniversary Edition, Cambridge University Press, Cambridge, MA, USA, 2010.
[36] J. Busemeyer and P. Bruza, Quantum Model Of Cognition And Decision, Cambridge University Press, Cambridge, MA, USA, 2012.
[37] R. Feynma, "The feynman lecture on physics: vol. III," 2021, https://www.feynmanlectures.caltech.edu/.
[38] N. P. Landsman, "Born rule and its interpretation," in Compendium of Quantum Physics, D. Greenberger, K. Hentschel, and F. Weinert, Eds., pp. 65-70, Springer, Berlin, Germany, 2009.
[39] C. Moreira and A. Wichert, "Interference Effects in quantum belief networks," Applied Soft Computing, vol. 25, pp. 64-85, 2014.

