

# Quantum Knowledge in Phase Space

Davi Geiger 

Courant Institute of Mathematical Sciences, New York University, New York, NY 10012, USA; dg1@nyu.edu

**Abstract:** Quantum physics through the lens of Bayesian statistics considers probability to be a degree of belief and subjective. A Bayesian derivation of the probability density function in phase space is presented. Then, a Kullback–Liebler divergence in phase space is introduced to define interference and entanglement. Comparisons between each of these two quantities and the entropy are made. A brief presentation of entanglement in phase space to the spin degree of freedom and an extension to mixed states completes the work.

**Keywords:** Bayesian statistics; interference; entanglement; entropy; Kullback–Liebler divergence

## 1. Introduction

This journal issue celebrates Claude Shannon’s 1948 formulation of “lost information” in phone-line signals [1]. It is curious that when von Neumann asked Shannon how he was getting on with their information theory, Shannon replied (according to [2]) “The theory was in excellent shape, except that he needed a good name for “missing information”. “Why don’t you call it entropy”, von Neumann suggested. “In the first place, a mathematical development very much like yours already exists in Boltzmann’s statistical mechanics, and in the second place, no one understands entropy very well, so in any discussion you will be in a position of advantage”.

A quantification of entanglement and interference in quantum phase space for pure states is proposed through Shannon entropy and related concepts (see Appendix B for a brief review). In order to arrive there a phase space probability density must be derived and leads to the next topic that has captivated much of quantum physics discussions since its first days, namely the role of measurement in physics. Measurements are associated with events in statistics, since through measurements full knowledge of a physical variable is acquired. The role of knowledge in quantum physics is then visited.

### 1.1. Bayesian Knowledge in Quantum Physics

In the field of statistics two views offered by the Bayesian [3] and the frequentist [4,5] divide the experts. Bayesian thinking is based on the idea that probabilities are degrees of belief about the events while the frequentist approach is based on the idea that the probability of an event occurring is equal to the long-run frequency with which that event occurs.

### 1.2. Measurements and Knowledge

A measurement in a statistical theory is modeled as an event, since full knowledge of the variable being measured is acquired.

In quantum physics the outcome of a measurement is associated with eigenvalues of a chosen Hermitian operator. For example, if one measures the position or the spin along the  $z$ -direction of a state, such a position or spin value becomes an event and the state becomes the eigenstate associated with the measured eigenvalue.

In both classical probability and quantum physics, there are scenarios where it is possible to infer full knowledge of a variable without the direct measurement of such a variable. A simple example in classical probability starts with two balls in a bag, one is red



**Citation:** Geiger, D. Quantum Knowledge in Phase Space. *Entropy* **2023**, *25*, 1227. <https://doi.org/10.3390/e25081227>

Academic Editors: Josip Pečarić, Đilda Pečarić and Muhammad Adil Khan

Received: 24 July 2023

Revised: 13 August 2023

Accepted: 15 August 2023

Published: 17 August 2023



**Copyright:** © 2023 by the author. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

and the other is white. By randomly drawing one of them and upon the knowledge of the outcome, one can infer the other ball color immediately without any further measurement. The prior knowledge that there were two different color balls in the bag is combined with the evidence from the observation to give certainty, immediately, on the other ball's color. This knowledge acquired about the other ball is not causal since if another player is only told the color of the ball that was left in the bag, the player would immediately infer the color of the ball that was drawn. Reasoning with knowledge in this case is invariant with respect to the time of occurrence. In quantum physics entangled states offer an analogous scenario. In teleportation experiments [6–9] with entangled fermions, the spin  $z$ -direction of one fermion can immediately be inferred by measuring the spin  $z$ -direction component of the other fermion. The prior knowledge that both fermions must have opposite spin combined with the observation allows for such inference. However, such knowledge is not causal, it is acquired by combining the prior knowledge with an observation, like in the classical probability example.

The Bayesian view of quantum physics also says that knowledge associated with a state is subjective to the observer. For example, two observers conducting an experiment of quantum teleportation, even if they have the same prior knowledge of the set up, will have different knowledge about the experiment according to when and where a measurement is obtained. According to the special theory of relativity, information or knowledge cannot be transferred instantaneously and so observers of an entangled pair will possess different knowledge of the variables at the time one of them is measured and so their predictions about the outcome of the other variable will differ, e.g., see [10,11]. These experiments with the entangled pair traveling at long distances suggest that quantum physics is best described as a Bayesian theory. It is worth stressing that a theory of knowledge is not necessarily a theory of cause and consequence.

One may wonder, is there a causal explanation for the quantum phenomena? Revisiting the Bohr vs. Einstein debate, e.g., see [12,13] and references, the view put forward here follows the epistemological view of Bohr: quantum theory is today the best model for predictions. The view put forward here also resonates with the ontological concern of Einstein adding the question: "Is there a causal theory that accounts for the quanta phenomena?" The quest here is for a better understanding of quantum theory as a Bayesian theory and the use of Shannon entropy and related concepts to characterize interference and entanglement. The quest for a causal model is left open.

Here, knowledge and information are meant to be the same thing. There is much work in information in quantum physics, e.g., [14–17] and references, but they use von Neumann mixed states entropy as a starting point which attributes zero information content for all pure states. In contrast, the starting point and focus here is pure states, the core of quantum physics theory. Mixed states do follow from pure states. The emphasis in using the term knowledge or degree of belief for the probabilities instead of information is that it is the language used in Bayesian theory and it stresses that it is subjective. Yet, the Bayesian view provides all the predictions quantum theory can make today and perhaps one can expand it as discussed next.

### 1.3. Phase Space

From a statistical perspective, randomness associated with a quantum state cannot be fully captured by one operator, afterwards an eigenstate of any one operator will seem to contain no randomness, a measurement by this operator will result with certainty in the eigenvalue of this eigenstate.

A unique aspect of quantum physics, expressed by the uncertainty principle, is that eigenvalues of two non-commuting operators cannot be measured simultaneously. Events cannot occur for such pair of variables simultaneously. These are mutually exclusive quantum variables and so, randomness associated with a state cannot be reduced to zero. To characterize the randomness associated with a state, non-commuting operators are then needed. A special operator is  $x$  with eigenstates  $|x\rangle$ . Together with the unitary evolution

operator, they define the space-time properties of states. In quantum mechanics, the uncertainty principle associated with the operator  $x$  is derived from the non-commuting property  $[x, p] = i\hbar$ , where  $p$  is the momentum operator. The variables  $x$  and  $k = \frac{p}{\hbar}$  are Fourier of each other, where  $k$  is the spatial frequency variable. The randomness of a state  $|\Psi_t\rangle$  is then captured in phase space, the space formed by the pair of Fourier variables  $x, k$ .

In quantum field theory, a relativistic theory, space becomes the domain variable, not an operator, and the spatial frequency  $k$  becomes the Fourier domain variable [18]. The phase space is then the space of the two domains of the quantum fields, one being the Fourier of the other. In this case, quantum field operators are rooted on the creation and annihilation operators defined in these two domains. A creation operator can then create a particle at some position or with some spatial frequency. The coefficients in front of these operators replace the role of the wave functions establishing the distribution in phase space. Phase space becomes the space where all randomness of a state is captured in quantum field theory.

In probability theory, any statistics of interest is derived from the joint distribution of all variables, so the joint distribution in phase space is the quantity to be derived. A constraint to events in phase space occurs from the fact that  $\psi(x, t) = \langle x | \Psi_t \rangle$  and  $\phi(k, t) = \langle k | \Psi_t \rangle$  are the Fourier transform of each other. If one acquires full knowledge of one of the variables at time  $t$ , expressed by a Dirac delta distribution, then the other phase space variable must be described by a uniform distribution, indicating maximum entropy in this other variable. It is this constraint of our knowledge about the pair of variables  $x$  and  $k$  at time  $t$  that yields the uncertainty principle as clearly formulated by Robertson [19]. The uncertainty principle suggests that events do occur in a volume element in phase space of size  $\Delta x \Delta k \geq 1$ , forming a coarse representation of the phase space. Thus, with respect to the statistical mechanical view of Gibbs [20], quantum mechanics already have a coarse mechanism built in to describe events in phase space.

Early attempts to create a quantum probability distribution in phase space by Wigner [21] and by Husimi [22] ended up with pseudo-distributions that fail Kolmogorov probability axioms and also have consistency difficulties with special relativity. Thus, the need for a pursue of a new approach.

With a new approach to create a phase space probability density, to be developed in Section 2 capturing knowledge about the phase space variables, we further exploit how such knowledge can be used to characterize quantum states entanglement and interference.

#### 1.4. Entropy, Interference, and Entanglement

The previous work [23,24] shows how Shannon entropy of a quantum state in phase space captures the loss of knowledge or the loss of information such a state describes. That work explores the hypothesis that knowledge (or information) cannot be gained in a closed quantum system to account for the time arrow. Here, the objective is to quantify interference and entanglement in terms of information loss or gained.

The Kullback–Liebler divergence (reviewed in Appendix B) is employed to define interference as a loss of information if one replaces the state probability density in phase space by a “classical probability density” in phase space. The Kullback–Liebler divergence is also employed to define entanglement as a loss of information if one replaces a state probability density in phase space by a product of state probability density in phase space. Such quantification of interference and entanglement could help our understanding of physical system evolution, for example by restricting which physical phenomena are allowed according to the gain or loss of interference or entanglement.

Position and spin are degrees of freedom (DoFs) required to specify a quantum state. This paper addresses how knowledge in phase space is quantified for position and spin DOFs.

### 1.5. Paper Organization

The remainder of the paper is organized as follows. In Section 2, the Bayesian formulation of probability density in position-momentum phase space is developed. In Section 3, the quantification of interference in phase space is proposed and compared to the entropy in phase space. In Section 4, the quantification of entanglement in phase space is proposed and compared to the entropy in phase space. Section 5 expands the concept of phase space to spin systems and proposes the quantification of entanglement for spin systems. Such quantification can also be expanded to Qbit technology. We also briefly show the approach to mixed states and compare it to von Neumann entropy. Lastly, concluding remarks are provided in Section 6.

## 2. A Bayesian View of Quantum Phase Space

One question immediately arises, what is the meaning of a joint distribution in phase space given that events in phase space cannot occur?

For quantum physics to be a statistical theory, a joint distribution for two variables of non-commuting operators must not require an event and a conditional distribution for the same variables must not require an event to be given. Instead, the joint distribution will describe the degree of belief about the joint variables and the conditional distribution will describe the degree of belief about a variable given the degree of belief about the other variable. An event associated with any one, but not both, of the variables of the non-commuting operators is still possible.

Quantum physics, as a statistical theory, is best described through the states so that classical logical manipulations such as **or** and **and** become additions and products of states analogous to operations in classical probability theory. For example, in the double slit experiment, according to classical logic an electron can pass through slit 1 **or** slit 2, so the quantum state at slit 1 will add with the quantum state at slit 2 to form the final state. To represent a particle A **and** a particle B, one takes the product of the two quantum states (a state in a product of Hilbert Spaces). After the operations with the state occur, quantum probabilities are then associated with the final state  $|\Psi_t\rangle$  via the probability density matrix  $\rho_t = |\Psi_t\rangle\langle\Psi_t|$ .

With this view the following proposition and theorem follows

**Lemma 1** (Conditional Probability Density). *Given the projection of a state to the position basis,  $\psi_t(x) = \langle x|\Psi_t\rangle$ , then the conditional probability density function in the spatial frequency domain is  $\rho(k|\psi_t(x)) = |\langle k|\Psi_t\rangle|^2$ . Also, given  $\phi_t(k) = \langle k|\Psi_t\rangle$ , then the conditional probability density function in the position domain is  $\rho(x|\phi_t(k)) = |\langle x|\Psi_t\rangle|^2$ .*

**Proof.** Given the wave function  $\psi_t(x) = \langle x|\Psi_t\rangle$ , we can derive via the inverse Fourier transform the conditional wave function in spatial frequency

$$\phi_t(k|\psi_t(x)) = \int dx \frac{1}{\sqrt{2\pi}} e^{-ikx} \psi_t(x) \quad (1)$$

Clearly,  $\phi_t(k|\psi_t(x)) = \int dx \langle k|x\rangle \langle x|\Psi_t\rangle = \langle k|\Psi_t\rangle$ , where  $\langle k|x\rangle = \frac{1}{\sqrt{2\pi}} e^{-ikx}$ . Thus,  $\rho(k|\psi_t(x)) = |\phi_t(k|\psi_t(x))|^2 = |\langle k|\Psi_t\rangle|^2$ . Similarly, starting from  $\phi_t(k) = \langle k|\Psi_t\rangle$  and applying the Fourier transform followed by the magnitude square one obtain  $\rho(x|\phi_t(k)) = |\langle x|\Psi_t\rangle|^2$ .  $\square$

The conditional distribution  $\rho(k|\psi_t(x))$ , in general, depends on the entire function  $\psi(x)$  and not on any one event in  $x$ . Note that conditional probabilities do not necessarily describe a cause and consequence relation but rather they are knowledge or information relation.

**Theorem 1** (Joint Distribution in Phase Space). *Given a state  $|\Psi_t\rangle$ , evolving in time according to some Hamiltonian. Then, the joint distribution in phase space is  $\rho_t(x, k) = |\psi_t(x)|^2 |\phi_t(k)|^2$ , where  $\psi_t(x) = \langle x|\Psi_t\rangle$  and  $\phi_t(k) = \langle k|\Psi_t\rangle$  are the projection in position basis and spatial frequency basis.*

**Proof.** Considering the state projected in the position domain to be a prior known function  $\psi_t(x) = \langle x|\Psi_t\rangle$  we obtain the conditional projection of the state in the spatial frequency domain from Lemma 1 to be  $\phi_t(k|\psi_t(x)) = \langle k|\Psi_t\rangle$ . We then have the conditional density as  $\rho_t(k|\psi_t(x)) = |\langle k|\Psi_t\rangle|^2$ .

From Bayes' theorem applied to the density we have  $\rho_t(x, k) = \rho_t(k|\psi_t(x))\rho_t^x(x) = |\langle k|\Psi_t\rangle|^2 |\langle x|\Psi_t\rangle|^2$ .

Clearly, we could have started with the state prior  $\phi_t(k) = \langle k|\Psi_t\rangle$  and obtained via the Fourier transform the (conditional) projection of the state in position  $\psi_t(x|\phi_t(k)) = \langle x|\Psi_t\rangle$  and then obtained the same joint distribution.  $\square$

In quantum statistics, the joint distribution  $\rho_t(x, k) = \rho_t^x(x)\rho_t^k(k)$  is not to be interpreted as the product of two independent random variables since the two distributions  $\rho_t^x$  and  $\rho_t^k$  are not independent from each other, and independent events cannot occur simultaneously.

*Entropy in Phase Space*

Geiger and Kedem [23] proposed a quantification of knowledge of a quantum state through Shannon entropy associated with the phase space distribution, namely

$$S_t = - \int dx dk \rho_t(x, k) \log \rho_t(x, k). \tag{2}$$

They showed various desired properties of this entropy, including it to be invariant to canonical transformations, special relativity and to CPT transformations. We adopt this entropy here for the study of interference and entanglement.

**3. Interference**

Given two states  $|\Psi^A\rangle$  and  $|\Psi^B\rangle$  and a general superposition of these two states  $|\Psi_{A \text{ or } B}\rangle$  as in Equation (A2). The projections of  $|\Psi_{A \text{ or } B}\rangle$  written in polar representation are

$$\psi(x) = \langle x|\Psi_{A \text{ or } B}\rangle = \frac{e^{i\nu}}{\sqrt{Z_P}} \left( \cos \theta_1 |\psi^A(x)| e^{i\xi_A(x)} + \sin \theta_1 |\psi^B(x)| e^{i(\xi_B(x) - \varphi_1)} \right) \tag{3}$$

$$\phi(k) = \langle k|\Psi_{A \text{ or } B}\rangle = \frac{e^{i\nu}}{\sqrt{Z_Q}} \left( \cos \theta_1 |\phi^A(k)| e^{i\chi_A(k)} + \sin \theta_1 |\phi^B(k)| e^{i(\chi_B(k) + \varphi_1)} \right) \tag{4}$$

where  $|\cdot|$  is the magnitude value,  $\xi_{A,B}(x)$  are the complex phases associated to the wave functions  $\psi_{A,B}(x)$  and similarly  $\chi_{A,B}(k)$  are the complex phases associated to the wave functions  $\phi_{A,B}(k)$ .

The probability densities are then

$$\begin{aligned} p(x) &= |\psi(x)|^2 = \frac{1}{Z_P} \left( p^c(x) + |\psi^A(x)| |\psi^B(x)| \sin 2\theta_1 \cos(\xi_A(x) - \xi_B(x) - \varphi_1) \right), \\ q(k) &= |\phi(k)|^2 = \frac{1}{Z_Q} \left( q^c(k) + |\phi^A(k)| |\phi^B(k)| \sin 2\theta_1 \cos(\chi_A(k) - \chi_B(k) - \varphi_1) \right) \end{aligned} \tag{5}$$

where the normalization constants are  $Z_P = 1 + \sin 2\theta_1 \int dx (|\psi^A(x)| |\psi^B(x)| \cos(\xi_A(x) - \xi_B(x) - \varphi_1))$  and  $Z_Q = 1 + \sin 2\theta_1 \int dk (|\phi^A(k)| |\phi^B(k)| \cos(\chi_A(k) - \chi_B(k) - \varphi_1))$ , and

$$\begin{aligned} p^c(x) &= \cos^2 \theta_1 |\psi^A(x)|^2 + \sin^2 \theta_1 |\psi^B(x)|^2 \\ q^c(k) &= \cos^2 \theta_1 |\phi^A(k)|^2 + \sin^2 \theta_1 |\phi^B(k)|^2 \end{aligned} \tag{6}$$

are probability densities without the interference terms. The upper-index  $c$  refers to these probability densities also representing classical statistical combination (weighted average) of probability densities associated to the quantum states A and B.

**Definition 1 (Interference):** Given two states  $|\Psi^A\rangle, |\Psi^B\rangle$  and their linear superposition  $|\Psi_{A \text{ or } B}\rangle$  as in Equation (A2). Interference,  $\mathcal{I}$ , is the amount of information lost when  $\rho^c(x, k) = p^c(x)q^c(k)$  is used to approximate  $\rho(x, k) = p(x)q(k)$ . It is calculated via the Kullback–Liebler divergence between the phase space probability densities  $\rho(x, k)$  and  $\rho^c(x, k)$ , i.e.,

$$\begin{aligned} \mathcal{I}(\theta_1, \varphi_1, |\Psi^A\rangle, |\Psi^B\rangle) &= D_{KL}(p(x)q(k) || p^c(x)q^c(k)) \\ &= \text{CrossS}(p(x), p^c(x)) - S(p(x)) + \text{CrossS}(q(k), q^c(k)) - S(q(k)) \\ &= \int dx |\psi(x)|^2 \log\left(1 + \frac{|\psi^A(x)| |\psi^B(x)| \sin 2\theta_1 \cos(\xi_A(x) - \xi_B(x) - \varphi_1)}{\cos^2 \theta_1 |\psi^A(x)|^2 + \sin^2 \theta_1 |\psi^B(x)|^2}\right) \\ &\quad + \int dk |\phi(k)|^2 \log\left(1 + \frac{|\phi^A(k)| |\phi^B(k)| \sin 2\theta_1 \cos(\chi_A(k) - \chi_B(k) - \varphi_1)}{\cos^2 \theta_1 |\phi^A(k)|^2 + \sin^2 \theta_1 |\phi^B(k)|^2}\right), \end{aligned} \tag{7}$$

where  $\text{CrossS}(p, q)$  is the cross entropy between probability distributions  $p$  and  $q$  (see Equation (A8)) and the phase space entropy Equation (2) for the phase space distributions Equation (5) is given by

$$S(|\Psi_{A \text{ or } B}\rangle) = - \int \int dx dk p(x)q(k) \log(p(x)q(k)) = - \int dx p(x) \log p(x) - \int dk q(k) \log q(k) \tag{8}$$

As one varies the combination of the two states, the larger  $\mathcal{I}$  is, the larger the interference contribution to the distribution  $\rho(x, k) = p(x)q(k)$ .

There is no interference, i.e.,  $\mathcal{I} = 0$ , when

1. the functions' support in phase space do not overlap, i.e.,

$$|\psi^A(x)| |\psi^B(x)| = 0; \quad \forall x, \quad \text{and} \quad |\phi^A(k)| |\phi^B(k)| = 0; \quad \forall k, \tag{9}$$

2. the complex phases are aligned up to a constant multiple of  $\frac{\pi}{2}$ , i.e.,

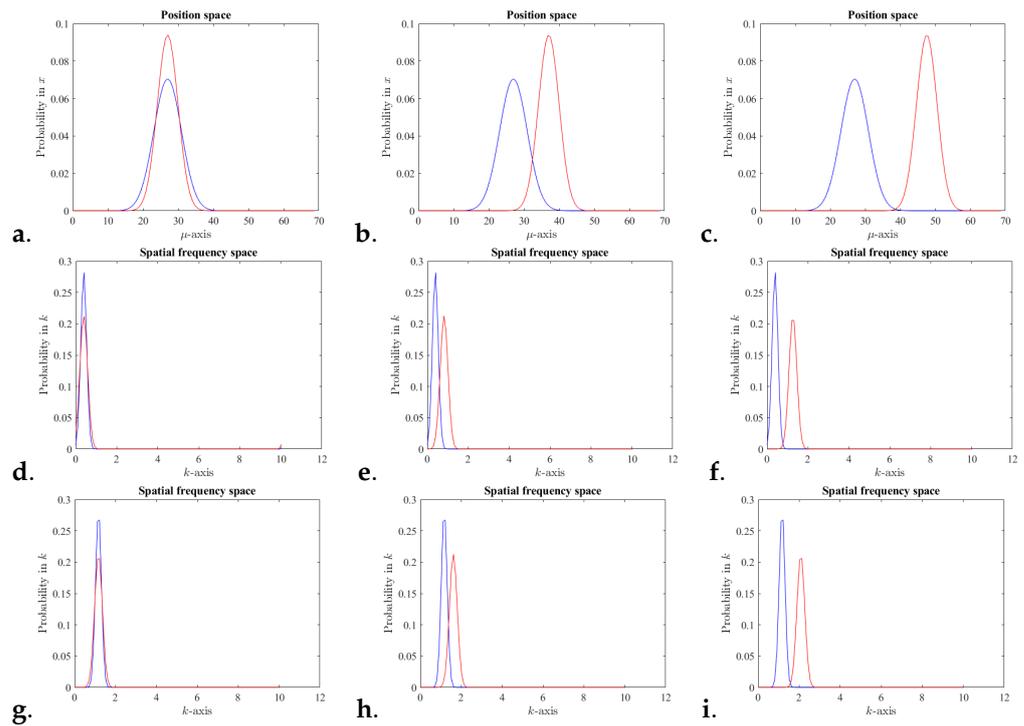
$$\xi_A(x) - \xi_B(x) - \varphi_1 = n \frac{\pi}{2}; \quad n \in \mathbb{Z}^+, \quad \text{and} \quad \chi_A(k) - \chi_B(k) - \varphi_1 = m \frac{\pi}{2}; \quad m \in \mathbb{Z}^+, \tag{10}$$

3. either  $\theta_1 = 0, \frac{\pi}{2}$ , since then there is no superposition of states. This will effectively occur when  $\psi^A(x) = \psi^B(x)$ .

Also, IFF there is no interference,  $P(x) = p^c(x)$  and  $Q(x) = q^c(x)$ . Figure 1c,f,i, illustrates scenarios with each state being a coherent state and not overlapping in neither position nor spatial frequency.

Clearly, one can consider the interference just in position representation or just in spatial frequency representation. However, here, the quantification of the interference in phase space distinguish the case (a) when a projection of superposition of two states in position space does not interfere but the same superposition projection in spatial frequency does interfere, from the case (b) a superposition of two states that do not interfere neither in position nor in spatial frequency.

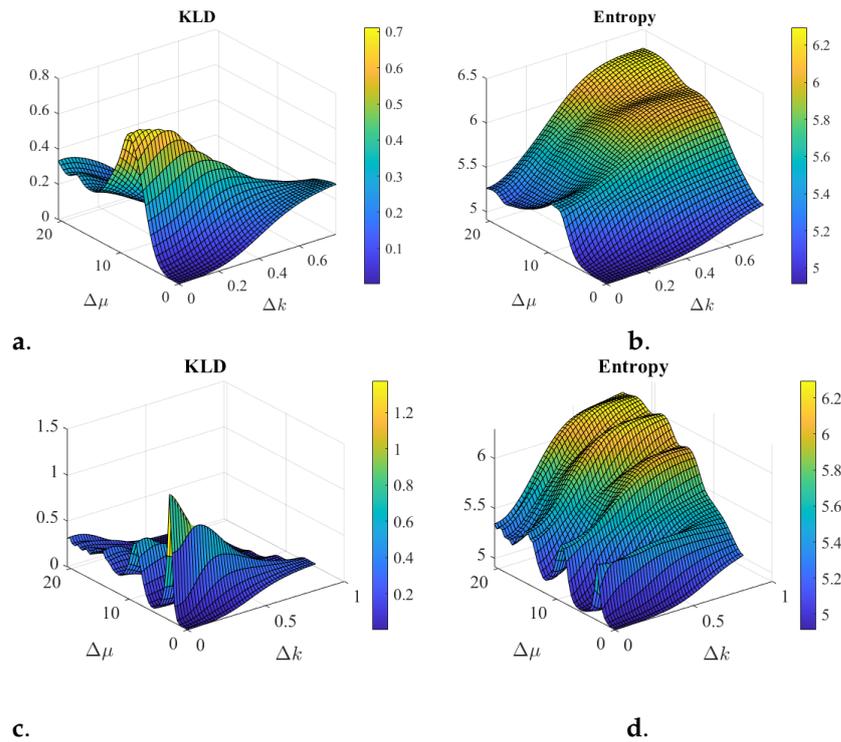
Figure 1 illustrate some scenarios of two coherent states that can be superposed to investigate inference as discussed next. Figure 2 illustrates some scenarios comparing the Kullback–Liebler divergence (KLD) (7) and the entropy (8). The entropy captures the notion of *overlap* of the superposition of states. For example, when both states are similar and superposed, the classical addition of probability densities and the quantum superposition do occur. Then, the KLD and the entropy will be small. However, for cases where there is no overlap between the two states, the classical weighted average distribution is a good approximation to the quantum one, the KLD will be small, but the entropy will be large.



**Figure 1.** Normal distribution in phase space for two coherent states, **A** and **B** in 1D with centers and variances as follows. For all (a–c) position space probabilities, with  $\mu_A = 27$  with  $\sigma_A = 5.6$  and for (a)  $\mu_B = 27$ , (b)  $\mu_B = 37$ , (c)  $\mu_B \approx 48$ , all with  $\sigma_B = 4.2$ . Note that for each coherent state, the spatial frequency value is the phase of the coherent state in position space. For (d–f), spatial frequency space probabilities, with  $k_A \approx 0.35$  and (d)  $k_B \approx 0.35$ , (e)  $k_B \approx 0.75$ , (f)  $k_B \approx 1.15$ . For (g–i), spatial frequency space probabilities, with  $k_A \approx 1.04$  and (g)  $k_B \approx 1.04$ , (h)  $k_B \approx 1.55$ , (i)  $k_B \approx 2.05$ .

Both concepts may be helpful to characterize the knowledge one has about the superposition of states. The KLD captures the distinction between classical probability and quantum probability, while the entropy captures the concept of lack of overlap of two states in quantum phase space. One advantage of the entropy over the KLD is that one does not need to know the components of the superposition of states to evaluate the entropy.

The role of the phase differences is noticeable, in position and in spatial frequency, as per (10). In addition, for coherent states, used in the simulations, the difference in phase of the state projection in position basis is the differences in the centers of the state projection in the spatial frequency basis. Also vice versa, the phase difference in spatial frequency is the center difference in the position domain. The periodic range for  $\Delta\mu$  is reduced for Figure 2c,d. creating the oscillations in the KLD and entropy. Entropy seems to be a good estimation for the interference behavior when the two states overlap either in spatial frequency or in position. However, the more the overlap in both spaces is reduced the more the two quantities differ in behavior.



**Figure 2.** Interference Simulations for a superposition of two coherent states as shown in Figure 1. The coherent state  $|\Psi^A\rangle$  has  $\mu_A = 27$ ,  $\sigma_A = 5.6$ , and for (a,b) the phase is  $k_A = 0.35$  while for (c,d), the phase is  $k_A = 1.04$ . The other coherent state  $|\Psi^B\rangle$  in position has fixed  $\sigma_B = 4.2$  and the position center and phase vary in 48 increments each, as follows:  $\mu_B = [27, \dots, 48]$ , and for (a,b), the phase varies as  $k_B = [0.35, \dots, 1.15]$ , while for (c,d), the phase varies as  $k_B = [1.04, \dots, 2.05]$ . The plots axis are all with  $\Delta\mu = \mu_B - \mu_A$  vs.  $\Delta k = k_B - k_A$ . The KLD and the entropy become small as the two states closely overlap, i.e., where  $\delta\mu \approx \delta k \approx 0$ . However, the KLD becomes small as the states do not overlap while the entropy gets to be larger. As the phase increases from (a,b) to (c,d) oscillation increases for both (KLD and Entropy) as periods reduce. Entropy seems to be a good estimation for the interference behavior when the two states overlap either in spatial frequency or in position. However, the more the overlap in both spaces is reduced the more the two quantities differ in behavior.

#### 4. Entanglement

Given two states  $|\Psi^A\rangle$  and  $|\Psi^B\rangle$  consider the entangled state  $|\Psi_{A \text{ and } B}\rangle$  as described in Equation (A3). Extending the language of probability to the quantum states, each state represent a random object. A quantum Bayesian interpretation of this state is that it reflects a joint state of  $|\Psi^A\rangle$  and  $|\Psi^B\rangle$ , thus the sub-index  $A$  and  $B$ . Considering the prior to be the state  $|\Psi^A\rangle$ , then the conditional state-operator, which is the product of a state in one Hilbert space times an operator acting in another Hilbert space, is

$$|\Psi^{\text{cond-B/A}}\rangle = |\Psi^B|\Psi^A\rangle = \frac{e^{i\varphi_2}}{\sqrt{Z_2}} \left( \cos \theta_2 |\Psi^A\rangle \otimes |\Psi^B\rangle \langle \Psi^A| + e^{i\varphi_2} \sin \theta_2 |\Psi^B\rangle \otimes \mathbf{I} \right), \quad (11)$$

where  $\mathbf{I}$  is the identity operator acting in the other Hilbert space. The conditional state-operator, with an abuse of state notation, reflects the impacts/change to a state  $|\Psi^B\rangle$  given the knowledge of the state  $|\Psi^A\rangle$ . This conditional state-operator leads to the joint state by acting on the prior state  $|\Psi^A\rangle$  as follows

$$|\Psi_{A \text{ and } B}\rangle = |\Psi^{\text{cond-B/A}}\rangle |\Psi^A\rangle = \frac{e^{i\varphi_2}}{\sqrt{Z_2}} \left( \cos \theta_2 |\Psi^A\rangle |\Psi^B\rangle + e^{i\varphi_2} \sin \theta_2 |\Psi^B\rangle |\Psi^A\rangle \right). \quad (12)$$

associated with joint density operator

$$\begin{aligned}
 \rho^{A \text{ and } B} &= |\Psi_{A \text{ and } B}\rangle \langle \Psi_{A \text{ and } B}| \\
 &= \frac{1}{Z_2} \left( \cos \theta_2 |\Psi^A\rangle |\Psi^B\rangle + e^{i\varphi_2} \sin \theta_2 |\Psi^B\rangle |\Psi^A\rangle \right) \left( \cos \theta_2 \langle \Psi^B| \langle \Psi^A| + e^{-i\varphi_2} \sin \theta_2 \langle \Psi^A| \langle \Psi^B| \right) \\
 &= \frac{1}{Z_2} \left( \cos^2 \theta_2 |\Psi^A\rangle |\Psi^B\rangle \langle \Psi^B| \langle \Psi^A| + \sin^2 \theta_2 |\Psi^B\rangle |\Psi^A\rangle \langle \Psi^A| \langle \Psi^B| \right. \\
 &\quad \left. + \frac{1}{2} \sin 2\theta_2 \left( e^{-i\varphi_2} |\Psi^A\rangle |\Psi^B\rangle \langle \Psi^A| \langle \Psi^B| + e^{i\varphi_2} |\Psi^B\rangle |\Psi^A\rangle \langle \Psi^B| \langle \Psi^A| \right) \right). \tag{13}
 \end{aligned}$$

Clearly, we would have obtained the same joint state had we considered the prior to be  $|\Psi^B\rangle$  and the conditional state-operator to be  $|\Psi^{\text{cond-A/B}}\rangle = |\Psi^A|\Psi^B\rangle = \frac{e^{i\varphi_2}}{\sqrt{Z_2}} (\cos \theta_2 |\Psi^A\rangle \otimes \mathbf{I} + e^{i\varphi_2} \sin \theta_2 |\Psi^B\rangle \otimes |\Psi^A\rangle \langle \Psi^B|)$ .

The probability density in position space of the joint state is then

$$\begin{aligned}
 p(x, y) &= \langle x| \langle y| \rho^{A \text{ and } B} |y\rangle |x\rangle = \frac{1}{Z_{2,P}} \left( \cos^2 \theta_2 |\Psi^A(x)|^2 |\Psi^B(y)|^2 + \sin^2 \theta_2 |\Psi^B(x)|^2 |\Psi^A(y)|^2 \right. \\
 &\quad \left. + \sin 2\theta_2 \cos(\zeta_A(x) - \zeta_B(x) + \zeta_B(y) - \zeta_A(y) - \varphi_2) |\Psi^A(x)| |\Psi^B(y)| |\Psi^A(y)| |\Psi^B(x)| \right). \tag{14}
 \end{aligned}$$

Similarly, the probability density  $q(k_x, k_y)$  in spatial frequency space is

$$\begin{aligned}
 q(k_x, k_y) &= \langle k_x| \langle k_y| \rho^{A \text{ and } B} |k_y\rangle |k_x\rangle = \frac{1}{Z_{2,Q}} \left( \cos^2 \theta_2 |\Phi^A(k_x)|^2 |\Phi^B(k_y)|^2 + \sin^2 \theta_2 |\Phi^B(k_x)|^2 |\Phi^A(k_y)|^2 \right. \\
 &\quad \left. + \sin 2\theta_2 \cos(\chi_A(k_x) - \chi_B(k_x) + \chi_B(k_y) - \chi_A(k_y) - \varphi_2) |\Phi^A(k_x)| |\Phi^B(k_y)| |\Phi^A(k_y)| |\Phi^B(k_x)| \right), \tag{15}
 \end{aligned}$$

where  $\theta_2 = \frac{\pi}{4}$  and  $\varphi_2 = 0, \pi$  describe bosons and fermions, respectively. However, in various empirical works, specially related to quantum computers, one can trap fermions on different locations so that they do not occupy the same state, and still entangle their spin with any general set of parameters. Thus, in theory, one could prepare two fermions to have different spin states and allow them to combine in phase space freely, described by a general set of parameters above. Expanding this formalism to Qbits there are no restrictions on the set of parameters used.

The phase space entropy (2) for this joint state becomes

$$S(|\Psi\rangle_2) = - \int \int \int dx dk_x dy dk_y p(x, y) q(k_x, k_y) \log(p(x, y), q(k_x, k_y)). \tag{16}$$

The product of states, or the disentangled states, are described by the two-state (12) with  $\theta_2 = 0, \frac{\pi}{2}$ , i.e.,

$$\begin{aligned}
 \theta_2 = 0 & & \theta_2 = \frac{\pi}{2} \\
 p^{\text{D1}}(x, y) &= |\Psi^A(x)|^2 |\Psi^B(y)|^2 & p^{\text{D2}}(x, y) &= |\Psi^B(x)|^2 |\Psi^A(y)|^2 \\
 q^{\text{D1}}(k_x, k_y) &= |\Phi^A(k_x)|^2 |\Phi^B(k_y)|^2 & q^{\text{D2}}(k_x, k_y) &= |\Phi^B(k_x)|^2 |\Phi^A(k_y)|^2 \tag{17}
 \end{aligned}$$

where the upper index D1 and D2 indicate two different disentangle states.

**Proposition 1.** Given two states  $|\Psi^A\rangle, |\Psi^B\rangle$  and the two-state  $|\Psi_{A \text{ and } B}\rangle$ , shown in (12). Refer by  $p^D(x, y)q^D(k_x, k_y)$  to any of the disentangled states (17). The Kullback–Liebler divergence  $D_{KL}(p(x, y)q(k_x, k_y) || p^D(x, y)q^D(k_x, k_y))$  is invariant to any choice of disentangled states for bosons or fermions, where  $\theta_2 = \frac{\pi}{4}$  and  $\varphi_2 = 0, \pi$ , respectively.

**Proof.** The proposition follows from performing the decomposition of the logarithm of products into the sum of logarithms and then using the symmetric properties of fermions and bosons. More precisely,

$$D_{KL}^D = -S(p, q) + \int dx dy dk_x dk_y p(x, y)q(k_x, k_y) \log(p^D(x, y)q^D(k_x, k_y)) \tag{18}$$

where  $D_{KL}^D$  is short for  $D_{KL}^D(p(x, y)q(k_x, k_y) || p^D(x, y)q^D(k_x, k_y))$ . Then, for D1 we get

$$\begin{aligned} D_{KL}^{D1} &= -S(p, q) + \int dx dy p(x, y) \log(|\Psi^A(x)|^2 |\Psi^B(y)|^2) + \int dk_x dk_y q(k_x, k_y) \log(|\Phi^A(k_x)|^2 |\Phi^B(k_y)|^2) \\ &= -S(p, q) + \int dx p(x) \log |\Psi^A(x)|^2 + \int dy p(y) \log |\Psi^B(y)|^2 \\ &\quad + \int dk_x q(k_x) \log |\Phi^A(k_x)|^2 + \int dk_y q(k_y) \log |\Phi^B(k_y)|^2 \end{aligned} \tag{19}$$

where  $p(x) = \int dy p(x, y)$ ,  $p(y) = \int dx p(x, y)$ ,  $q(k_x) = \int dk_y q(k_x, k_y)$ ,  $q(k_y) = \int dk_x q(k_x, k_y)$ . Note that the integrals yield the same functions  $p(\cdot)$  and  $q(\cdot)$  due to the symmetric properties for bosons and fermions.

Similarly, for D2 we get

$$\begin{aligned} D_{KL}^{D2} &= -S(p, q) + \int dx p(x) \log |\Psi^B(x)|^2 + \int dy p(y) \log |\Psi^A(y)|^2 \\ &\quad + \int dk_x q(k_x) \log |\Phi^B(k_x)|^2 + \int dk_y q(k_y) \log |\Phi^A(k_y)|^2 \end{aligned} \tag{20}$$

and clearly every term here has a perfect match in (19), e.g.,  $\int dx p(x) \log |\Psi^B(x)|^2 = \int dy p(y) \log |\Psi^B(y)|^2$  and so  $D_{KL}^{D2} = D_{KL}^{D1}$ .  $\square$

**Definition 2 (Entanglement):** Given two states  $|\Psi^A\rangle, |\Psi^B\rangle$  and the two-state  $|\Psi_{A \text{ and } B}\rangle$ , shown in (12), that when projected in phase space yields the probability density distribution  $\rho(x, y, k_x, k_y) = p(x, y)q(k_x, k_y)$  given by (14) and (15). Entanglement,  $\mathcal{E}$ , is the amount of information lost when the product of states is used to approximate  $\rho(x, y, k_x, k_y)$ . More formally, for bosons or fermions, where  $\theta_2 = \frac{\pi}{4}$  and  $\varphi_2 = 0, \pi$  we have

$$\begin{aligned} \mathcal{E}(\varphi_2 = 0, \pi, |\Psi^A\rangle, |\Psi^B\rangle) &= D_{KL}(p(x, y)q(k_x, k_y) || p^D(x, y)q^D(k_x, k_y)) \\ &= \text{CrossS}(p(x, y)q(k_x, k_y), p^D(x, y)q^D(k_x, k_y)) - S(p(x, y)q(k_x, k_y)) \end{aligned} \tag{21}$$

and when the parameters  $\theta_2$  and  $\varphi_2$  are free to vary (as in Qbits) then

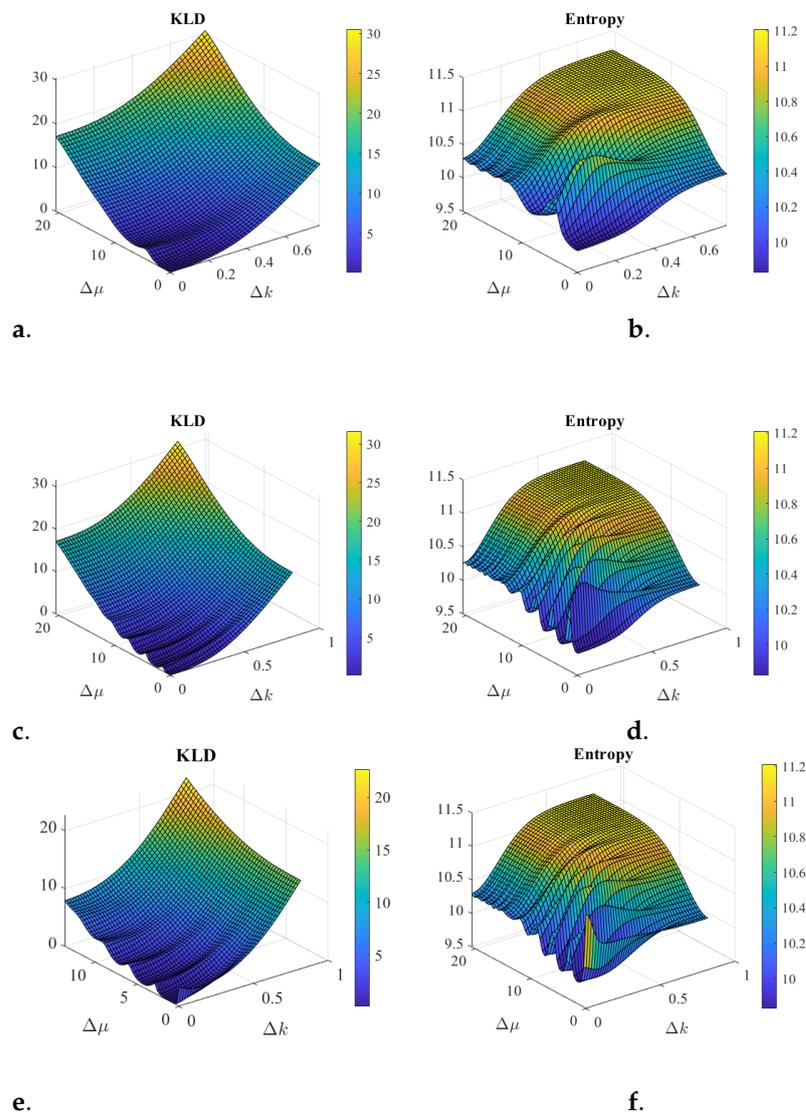
$$\begin{aligned} \mathcal{E}(\theta_2, \varphi_2, |\Psi^A\rangle, |\Psi^B\rangle) &= \arg \min_{D1, D2} [D_{KL}(p(x, y)q(k_x, k_y) || p^{D1}(x, y)q^{D1}(k_x, k_y)) \\ &\quad D_{KL}(p(x, y)q(k_x, k_y) || p^{D2}(x, y)q^{D2}(k_x, k_y))] \end{aligned} \tag{22}$$

The entanglement vanishes when

1.  $\theta_2 = 0, \frac{\pi}{2}$
2.  $|\Psi^A\rangle = |\Psi^B\rangle$ .

One comparison of interest is between the entanglement (21) and the entropy (16). Figure 3 illustrates some scenarios combining two coherent states where these two quantities are evaluated and a comparisons is made. While the definition of entanglement is through the KLD, the entropy captures a similar behavior and it can be evaluated from the state itself, without having to know what the product of the states would be. Entropy behavior seems to be a good estimation for entanglement. One reason is that the basis functions for a product of  $N$  Hilbert spaces of the total operators  $X = x_1 \times x_2 \dots \times x_N$  and  $K$  is given by the product of states  $|x_1\rangle \dots |x_N\rangle$  and  $|k_1\rangle \dots |k_N\rangle$ . These are the eigenstates of the operators  $x_1 \dots x_N$  and  $k_1 \dots k_N$ . Then the product of coherent quantum states, where

each one minimizes the phase space entropy, projected in these bases have lower entropy than the entangled states, which are linear superposition of these products.



**Figure 3.** Entanglement simulations from two coherent states shown in Figure 1, with normal probability distributions. Note that the phase of the coherent state projected to position space is the center of the coherent state projected in the spatial frequency space, and vice versa. The coherent state  $|\Psi^A\rangle$  has a fixed set of parameters,  $\mu_A = 27$ ,  $\sigma_A = 5.6$  in position space, and for (a,b) the phase is  $k_A = 0.35$  while for (c–f) the phase is  $k_A = 1.04$ . The coherent state  $|\Psi^B\rangle$  in position has fixed  $\sigma_B = 4.2$  and the center and phase vary in 48 increments each, as follows:  $\mu_B \in [27, \dots, 48]$ , and for (a,b)  $k_B \in [0.35, \dots, 1.15]$ , while for (c–f) the phase varies as  $k_B \in [1.04, \dots, 2.05]$ . The parameter  $\theta_2 = \frac{\pi}{4}$  is fixed when entangling the two states. Cases (a–d) show KLD and Entropy, respectively, for a symmetric entanglement where phase  $\varphi_2 = 0$ . Cases (e,f) show KLD and Entropy, respectively, for an anti-symmetric entanglement where phase  $\varphi_2 = \pi$ . The effect of the phase  $\varphi_2$  is only noticeable when the two states are very similar to each other and then both, KLD and entropy, yield large values for the anti-symmetric case (after all anti-symmetric functions must vanish in these cases, while the product of states does not). While the KLD has a smoother behavior, both increase as the separation of the two coherent state parameters increases. The larger values of the phase parameters in (c–f) clearly cause a periodic behavior. Entropy behavior seems to be a good estimation for entanglement.

The study of the more general case of entanglement (22), which is applicable to Qbits, is left as future work.

### 5. Entanglement for Spin or Qbit Phase Space

The degrees of freedom (DoFs) in quantum physics specify the wave function (projection of the state in position space) and the spin. Thus, when quantifying our knowledge of a quantum system one must also quantify our knowledge of a spin state. Qbits are like spin in formalism, but with less constraints as the Pauli exclusion principle is no longer required (since other aspects of the complete state may already be identified or are already anti-symmetric). What follows also applies to Qbits.

Let us consider two spin states  $|\zeta_A\rangle$  and  $|\zeta_B\rangle$ , each formed with  $N^A$  and  $N^B$  spin  $s$  particles, respectively. A spin state formed from these two states may be in a superposition of any of the total spin magnitudes  $\mathbf{s} \in [s_N, s_N - 1, \dots, \frac{1}{2} \bmod (2s_N, 2)]$ , where  $s_N = N^A + N^B$ , and is written as

$$|\zeta\rangle = \frac{e^{i\varphi/2}}{\sqrt{Z_2}} \left( \cos \theta_2 |\zeta^A\rangle |\zeta^B\rangle + e^{i\varphi_2} \sin \theta_2 |\zeta^B\rangle |\zeta^A\rangle \right). \tag{23}$$

The operator  $S_z$  associated with  $|\zeta\rangle$  is given by

$$S_z = S_z^A \otimes I_{2N^B} + I_{2N^A} \otimes S_z^B \tag{24}$$

where  $I_N$  is the identity of dimension  $N$  and  $\otimes$  is the exterior product.

As discussed earlier, the position and spatial frequency operators for a product of  $N$  Hilbert spaces is the product of position operators in each Hilbert space. For spin states, the total  $S_z$  operator (24) is not just the product of single particle operator. The set of eigenstates of the operators  $S_z$  and  $S^2$  include entangled spin states. Let us refer to the eigenstates of  $S_z$  and  $S^2$  as  $\{|\zeta_{s,m_s}\rangle; s = s_N, s_N - 1, \dots, s^{\min}, m_s = -s, \dots, s\}$ , where  $s^{\min} = \frac{1}{2} \bmod (2 * s_N, 2)$ . Thus, the state is written in this basis as

$$|\zeta\rangle = \sum_{s=s^{\min}}^{s_N} \sum_{m=-s}^s \alpha_{s,m} |\zeta_{s,m}\rangle, \tag{25}$$

where  $\alpha_{s,m} = \langle \zeta_{s,m} | \zeta \rangle \in \mathbb{C}$  and  $1 = \sum_{s=s^{\min}}^{s_N} \sum_{m=-s}^s |\alpha_{s,m}|^2$ .

The phase space for the spin associated with the operators  $S_z$  and  $S^2$  is derived from quantizing the sphere, the surface of the ball with a radius of the spin magnitude  $\hbar s_N$ , as developed by the Geometric Quantization (GQ) method, e.g., see [25–27]. Geiger and Kedem [24] have proposed this approach to evaluate the entropy of a quantum state which is briefly summarized next.

The conjugate basis to  $\{|\zeta_{s,m}\rangle\}$  is  $\{|\phi\rangle; [0, 2\pi]\}$ , obtained by identifying the angle  $\phi$ , the rotation angle around the  $z$ -axis of  $e^{-iS_z\phi}$ , as the conjugate operator to  $S_z$ . The spin state  $|\zeta\rangle$  in this basis is

$$|\zeta\rangle = \int_0^{2\pi} |\phi\rangle \langle \phi | \zeta \rangle d\phi = \int_0^{2\pi} \lambda(\phi) |\phi\rangle d\phi, \tag{26}$$

where

$$\lambda(\phi) = \langle \phi | \zeta \rangle = \sum_{s=s^{\min}}^{s_N} \sum_{m=-s}^s \alpha_{s,m} \langle \phi | \zeta_{s,m} \rangle = \sum_{s=s^{\min}}^{s_N} \sum_{m=-s}^s \alpha_{s,m} \psi_{s,m}(\phi). \tag{27}$$

and

$$\psi_{s,m}(\phi) = \begin{cases} \frac{1}{\sqrt{2\pi}} e^{i(s+m)\phi}, & m \geq 0 \quad (\text{northern hemisphere}), \\ \frac{1}{\sqrt{2\pi}} e^{i(-s+m)\phi}, & m < 0 \quad (\text{southern hemisphere}). \end{cases} \tag{28}$$

The two solutions in (28) are periodic in  $\phi$  and differ by a phase (gauge) transformation of  $e^{-i2s\phi}$ .

Thus, for a state  $|\zeta\rangle$  with density matrix  $\rho = |\zeta\rangle\langle\zeta|$ , the probabilities of the phase space are the product of the probabilities  $\{\rho_{s,m} = \langle\zeta_{s,m}|\rho|\zeta_{s,m}\rangle = |\alpha_{s,m}|^2\}$  with the probability densities  $\{\rho(\phi) = \langle\phi|\rho|\phi\rangle = |\lambda(\phi)|^2\}$ . Note that given  $\{\alpha_{s,m}\}$  one obtains  $\{\lambda(\phi)\}$  via the predefined set of functions (28), i.e., one can interpret  $\rho(\phi)$  as a conditional probability density  $\rho(\phi/\{\alpha_{s,m}\})$ .

Thus, the entropy (2) of a spin state  $|\zeta\rangle$  in spin phase space is

$$\begin{aligned} S &= S^z + S^\phi = - \sum_{s=s^{\min}}^{s^N} \sum_{m=-s}^s \rho_{s,m} \ln \rho_{s,m} - \int \rho(\phi) \ln \rho(\phi) \, d\phi \\ &= - \sum_{s=s^{\min}}^{s^N} \sum_{m=-s}^s |\alpha_{s,m}|^2 \ln |\alpha_{s,m}|^2 - \int |\lambda(\phi)|^2 \ln |\lambda(\phi)|^2 \, d\phi. \end{aligned} \tag{29}$$

The first term is the Shannon entropy capturing the randomness of the spin value along the z-axis. The second term is differential entropy capturing the randomness of the spin value in the plane perpendicular to the z-axis, i.e., the entropy of the polarization angle  $\phi$ . Geiger and Kedem [24] have shown that this entropy reaches its lowest value zero for the eigenstates of the two operators  $S_z$  and  $S^2$ . Products of states that are described by the superposition of these eigenstates will have higher entropy.

The densities in spin phase space associated with the product state  $|\zeta^A\rangle|\zeta^B\rangle$  are derived from the projections

$$|\zeta^A\rangle|\zeta^B\rangle = \sum_{s=s^{\min}}^{s^N} \sum_{m=-s}^s \alpha_{s,m}^{A,B} |\zeta_{s,m}\rangle \Rightarrow \alpha_{s,m}^{A,B} = \langle\zeta_{s,m}|\zeta^A\rangle|\zeta^B\rangle \Rightarrow \lambda^{A,B}(\phi) = \sum_{s=s^{\min}}^{s^N} \sum_{m=-s}^s \alpha_{s,m}^{A,B} \psi_{s,m}(\phi), \tag{30}$$

Extending the work of [24] to also define the Kullback–Liebler divergence between a joint state as described by (12) and the product of the states  $|\zeta^A\rangle|\zeta^B\rangle$ .

**Definition 3 (Spin Entanglement):** Given two spin states  $|\zeta^A\rangle, |\zeta^B\rangle$  and the joint state  $|\zeta\rangle$ , shown in (23), that when projected in spin phase space yields the probability density distribution  $\rho_{s,m}(\phi) = |\alpha_{s,m}|^2 |\lambda(\phi)|^2$ . Spin entanglement,  $s\mathcal{E}$ , is the amount of information lost when the product of states is used to approximate  $\rho_{s,m}(\phi)$ . More precisely,

$$\begin{aligned} s\mathcal{E}(\theta_2, \varphi_2, |\zeta^A\rangle, |\zeta^B\rangle) &= \arg \min_{(A,B),(B,A)} \left[ D_{KL} \left( |\alpha_{s,m}|^2 |\lambda(\phi)|^2 \middle| |\alpha_{s,m}^{A,B}|^2 |\lambda^{A,B}(\phi)|^2 \right), \right. \\ &\quad \left. D_{KL} \left( |\alpha_{s,m}|^2 |\lambda(\phi)|^2 \middle| |\alpha_{s,m}^{B,A}|^2 |\lambda^{B,A}(\phi)|^2 \right) \right] \end{aligned} \tag{31}$$

One comparison of interest to be made is between spin entanglement (31) and spin entropy (29). In contrast to the position  $\times$  spatial frequency phase space entanglement, the spin entropy will be minimized and attain value zero for some entangled states that are eigenstates of the total spin operators, such as it is the case of Bell entangled states for two particles [24]. The product of states will then be described by the superposition of these entangled states and thus, will not have a well-defined spin magnitude. Their entropy will be larger. The KLD will then anti-correlate with the entropy. A more detailed study is left for the future.

#### Expansion to Mixed States

One extension of this approach to mixed states starts from the density matrix derived from the general state (23), namely

$$\begin{aligned} \rho = |\zeta\rangle\langle\zeta| &= \frac{1}{Z_2} \left( \cos^2 \theta_2 |\zeta^A\rangle|\zeta^B\rangle\langle\zeta^B|\langle\zeta^A| + \sin^2 \theta_2 |\zeta^B\rangle|\zeta^A\rangle\langle\zeta^A|\langle\zeta^B| \right. \\ &\quad \left. + \frac{1}{2} e^{-i\varphi_2} \sin 2\theta_2 |\zeta^A\rangle|\zeta^B\rangle\langle\zeta^A|\langle\zeta^B| + \frac{1}{2} e^{i\varphi_2} \sin 2\theta_2 |\zeta^B\rangle|\zeta^A\rangle\langle\zeta^B|\langle\zeta^A| \right). \end{aligned} \tag{32}$$

Then, tracing out the density matrix (and assuming the states  $|\zeta^A\rangle, |\zeta^B\rangle$  to be in orthogonal Hilbert spaces to each other)

$$\rho_s^{\text{Mixed}} = \langle \zeta^A | \langle \zeta_s | \langle \zeta_s | |\zeta^A\rangle + \langle \zeta^B | \langle \zeta_s | \langle \zeta_s | |\zeta^B\rangle = \cos^2 \theta_2 |\zeta^B\rangle \langle \zeta^B| + \sin^2 \theta_2 |\zeta^A\rangle \langle \zeta^A|. \quad (33)$$

Then, von Neumann entropy is the Shannon entropy of this mixed state. While this is of interest to much research, Shannon entropy of pure states precedes von Neumann entropy. Also, von Neumann entropy obtained by tracing out pure states has some similarity to the entropy of the superposition of two states Equation (A2), but neglecting both quantities the interference and the conjugate variable of the phase space. Geiger and Kedem [23] showed invariant properties of the Shannon entropy in phase space that von Neumann entropy would not have.

## 6. Conclusions

The Bayesian statistic view was developed to construct the phase space probability density. The Bayesian approach to describe quantum physics in phase space considered the conditional probability densities and joint probability densities to be descriptions of the degree of belief about the phase space variables. It was observed that events representing measurements in the phase space do not occur. However, in a coarse description of the phase space, where the elementary volumes satisfy the uncertainty principle, they do occur. This derivation of the phase space density provides further support to the work of Geiger and Kedem [23,24] where they developed the entropy in phase space.

As a Bayesian theory, quantum physics is subjective, even if the prior knowledge is common to different observers, the measurements may not be. The entanglement scenarios where teleportation experiments have been reported, demonstrate the subjectivity of quantum physics as different observers make different predictions about the outcome of a variable. Moreover, the Bayesian view of quantum physics is that the theory is not a complete causal theory. The time evolution of a state via the unitary evolution is causal, as the Hamiltonian causes the state to evolve. However, the Bayesian combination of observation and the prior knowledge is not causal.

The Bayesian theory of quantum physics also considered the probabilistic object of manipulations to be the state and not the probabilities. Addition and multiplication operations usually associated with **or** and **and** logical operations were applied to quantum states and not to quantum probabilities.

With the phase space probability density constructed, the next objective was to quantify interference and entanglement in terms of information loss or gain, using the Kullback–Liebler divergence (rooted on Shannon entropy). A comparison to the Shannon entropy of the state and some similarities between the two quantities were revealed. One advantage of entropy is that it can be inferred from the quantum state, without any reference to the two states that were used to compute the divergence. It was noted that for spin phase space the relations between entropy and the Kullback–Liebler diverges will differ since the eigenstates of the total spin operator become entangled states.

Extrapolating the Bayesian approach to a philosophical interpretation, and revisiting the Bohr vs. Einstein debate, the view put forward here is similar to the epistemological view of Bohr: quantum theory is today the best model for predictions. The view put forward here also resonates with the ontological concern of Einstein adding the question: “Is quantum theory a causal theory?” If not, “is there a causal theory to be discovered?” After all, by adopting the Bayesian view, the question of a complete causal theory was left open.

**Funding:** This research received funding from National Science Foundation under Grant No. DMS-1439786 and the Simons Foundation Institute Grant Award ID 507536 while the author was in residence at the Institute for Computational and Experimental Research in Mathematics in Providence, RI, during the spring 2019 semester “Computer Vision” program.

**Informed Consent Statement:** Not applicable.

**Data Availability Statement:** The only data used here is shown in Figure 1 and all the parameters specification to generate the data is described in the caption.

**Acknowledgments:** Thank you to Carlos Tomei for the attention with the early drafts of this work and to the reviewers feedback.

**Conflicts of Interest:** The author declares no conflict of interest.

### Appendix A. Combining Two States

Consider a quantum state  $|\Psi^A\rangle$  in Hilbert space and its representation in phase space

$$\psi^A(x) = \langle x|\Psi^A\rangle \quad \text{and} \quad \phi^A(k) = \langle k|\Phi^A\rangle \tag{A1}$$

The time parameter is not shown here for simplicity/clarity. Figure 1 shows some scenarios of two coherent states projected in one dimension position and spatial frequency.

Given two quantum states in Hilbert space,  $|\Psi^A\rangle$  and  $|\Psi^B\rangle$ , are one can combine them as follows: (i) a superposition of these two states or (ii) a sum of an exterior product of the two states (product of the two Hilbert spaces). More precisely, a general quantum state formed from these two states can be written as

(i) One-State Superposition: 
$$|\Psi_{A \text{ or } B}\rangle = \frac{e^{i\nu_1}}{\sqrt{Z}} \left( \cos \theta_1 |\Psi^A\rangle + e^{i\varphi_1} \sin \theta_1 |\Psi^B\rangle \right) \tag{A2}$$

(ii) Two-State: 
$$|\Psi_{A \text{ and } B}\rangle = \frac{e^{i\nu_2}}{\sqrt{Z_2}} \left( \cos \theta_2 |\Psi^A\rangle |\Psi^B\rangle + e^{i\varphi_2} \sin \theta_2 |\Psi^B\rangle |\Psi^A\rangle \right) \tag{A3}$$

where  $\varphi_{1,2}, \nu_{1,2} \in [0, \pi]$ ,  $\theta_{1,2} \in [0, \frac{\pi}{2}]$  and  $Z, Z_2$  are normalization constants. The notation of the subindex  $|\Psi_{A \text{ or } B}\rangle$  refers to classical logic associated with the superopostion operation and of the subindex  $|\Psi_{A \text{ and } B}\rangle$  refers to classical logic associated with the product operation. For bosons or fermions there is the constraint  $\theta_2 = \frac{\pi}{4}$  and  $\varphi_2 = 0, \pi$ , respectively, so these states will be either symmetric or anti-symmetric. The special case when two boson states occupy the same state is also captured by  $\theta_2 = 0, \frac{\pi}{2}$ .

The projection to the spatial basis leads to the wave functions

(i) 
$$\psi(x) = \langle x|\psi\rangle = \frac{e^{i\nu}}{\sqrt{Z}} \left( \cos \theta_1 \psi^A(x) + e^{i\varphi_1} \sin \theta_1 \psi^B(x) \right) \tag{A4}$$

(ii) 
$$\psi(x, y) = \langle x|\langle y|\Psi\rangle_2 = \frac{e^{i\nu}}{\sqrt{Z_2}} \left( \cos \theta_2 \psi^A(x) \psi^B(y) + e^{i\varphi_2} \sin \theta_2 \psi^B(x) \psi^A(y) \right) \tag{A5}$$

and similarly for scenario (i)  $\phi(k) = \langle k|\psi\rangle$  and for scenario (ii)  $\phi(k_x, k_y) = \langle k_x|\langle k_y|\Psi\rangle_2$ .

### Appendix B. Entropy Concepts

The Rényi entropy of order  $\alpha$  of a given continuous distribution  $P(x)$

$$H_\alpha(P) = \frac{1}{1-\alpha} \log \left( \int P^\alpha(x) dx \right) \tag{A6}$$

The work shown in this paper can be extended to Rényi entropy.

Shannon entropy can be derived from Rényi entropy as follows

$$S(P) = \lim_{\alpha \rightarrow 1} H_\alpha(P) = - \int P(x) \log P(x) d\mu(x) \tag{A7}$$

where  $\mu$  is some reference measure, usually a Lebesgue measure on a Borel  $\sigma$ -algebra.

The cross entropy of two probability distributions  $p(x)$  and  $q(x)$  on the variable  $x$  is given by

$$\text{CrossS}(p, q) = - \int p(x) \log q(x) d\mu(x). \tag{A8}$$

The Kullback–Liebler divergence between two probability distributions  $p(x)$  and  $q(x)$  is given by

$$D_{KL}(p||q) = - \int p(x) \log \frac{q(x)}{p(x)} d\mu(x) = \text{CrossS}(p, q) - S(p). \quad (\text{A9})$$

## References

- Shannon, C.E. A mathematical theory of communication. *Bell Syst. Tech. J.* **1948**, *27*, 379–423. [[CrossRef](#)]
- Avery, J. *Information Theory and Evolution*; World Scientific: Singapore, 2003; ISBN 981-238-400-6.
- Bayes, T. An Essay toward solving a Problem in the Doctrine of Chances. *Philos. Trans. R. Soc. Lond.* **1764**, *53*, 370–418. [[CrossRef](#)]
- Cournot, A.A. *Exposition de la Théorie des Chances et des Probabilités*; L. Hachette: Paris, France, 1843.
- Venn, J. *The Logic of Chance: An Essay on the Foundations and Province of the Theory of Probability, with Especial Reference to Its Logical Bearings and Its Application to Moral and Social Science, and to Statistics*; Macmillan & Co, London: London, UK, 1888.
- Einstein, A.; Podolsky, B.; Rosen, N. Can Quantum-Mechanical Description of Physical Reality be Considered Complete? *Phys. Rev.* **1935**, *47*, 777–780. [[CrossRef](#)]
- Bennett, C.H.; Brassard, G.; Crépeau, C.; Jozsa, R.; Peres, A.; Wootters, W.K. Teleporting an unknown quantum state via dual classical and Einstein–Podolsky–Rosen channels. *Phys. Rev. Lett.* **1993**, *70*, 1895–1899. [[CrossRef](#)] [[PubMed](#)]
- Bouwmeester, D.; Mattle, K.; Eibl, M.; Weinfurter, H.; Zeilinger, A. Experimental quantum teleportation. *Nature* **2006**, *390*, 575–579. [[CrossRef](#)]
- Boschi, D.; Branca, S.; De Martini, F.; Hardy, L.; Popescu, S. Experimental Realization of Teleporting an Unknown Pure Quantum State via Dual Classical and Einstein–Podolsky–Rosen Channels. *Phys. Rev. Lett.* **1998**, *80*, 1121–1125. [[CrossRef](#)]
- Boyd, R.W. Slow and fast light: Fundamentals and applications. *J. Mod. Opt.* **2009**, *56*, 1908–1915. [[CrossRef](#)]
- Ferrie, C. *Quantum Entanglement Isn't All That Spooky After All*; Scientific American: New York, NY, USA, 2023.
- Marage, P.; Wallenborn, G. The Debate between Einstein and Bohr, or How to Interpret Quantum Mechanics. In *The Solvay Councils and the Birth of Modern Physics*; Marage, P., Wallenborn, G., Eds.; Birkhäuser Basel: Basel, Switzerland, 1999; pp. 161–174. [[CrossRef](#)]
- Becker, A. *What Is Real?: The Unfinished Quest for the Meaning of Quantum Physics*; Basic Books: New York, NY, USA, 2018.
- Deutsch, D. Quantum theory, the Church–Turing principle and the universal quantum computer. *Proc. R. Soc. Lond. Math. Phys. Sci.* **1985**, *400*, 97–117.
- Nielsen, M.A.; Chuang, I. *Quantum Computation and Quantum Information*; Cambridge University Press: Cambridge, UK, 2010.
- Witten, E. A Mini-Introduction To Information Theory. *arXiv* **2018**, arXiv:1805.11965.
- Casini, H.; Huerta, M.; Magan, J.; Pontello, D. Entanglement entropy and superselection sectors. Part I. Global symmetries. *J. High Energy Phys.* **2020**, *2020*, 14. [[CrossRef](#)]
- Weinberg, S. *The Quantum Theory of Fields: Volume 1, (Foundations)*; Cambridge University Press: Cambridge, UK, 1995.
- Robertson, H.P. The Uncertainty Principle. *Phys. Rev.* **1929**, *34*, 163–164. [[CrossRef](#)]
- Gibbs, J.W. *Elementary Principles in Statistical Mechanics*; Dover Publications: New York, NY, USA, 2014.
- Wigner, E.P. On the quantum correction for thermodynamic equilibrium. *Phys. Rev.* **1932**, *40*, 749–759. [[CrossRef](#)]
- Husimi, K. Some Formal Properties of the Density Matrix. In *Proceedings of the Physico-Mathematical Society of Japan, 3rd Series*; The Physical Society of Japan, The Mathematical Society of Japan: Tokyo, Japan, 1940; Volume 22, pp. 264–314. [[CrossRef](#)]
- Geiger, D.; Kedem, Z.M. On Quantum Entropy. *Entropy* **2022**, *24*, 1341. [[CrossRef](#)] [[PubMed](#)]
- Geiger, D.; Kedem, Z.M. Spin Entropy. *Entropy* **2022**, *24*, 1292. [[CrossRef](#)] [[PubMed](#)]
- Woodhouse, N.M.J. *Geometric Quantization*; Oxford University Press: Oxford, UK, 1997.
- Blau, M. Symplectic Geometry and Geometric Quantization. Lecture Notes. 1992. Available online: <https://ncatlab.org/nlab/files/BlauGeometricQuantization.pdf> (accessed on 14 August 2023).
- Nair, V.P. Elements of geometric quantization and applications to fields and fluids. *arXiv* **2016**, arXiv:1606.06407.

**Disclaimer/Publisher’s Note:** The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.