Open Science Framework DOI:10.31219/osf.io/bsnh7

The rule of conditional probability is valid in quantum theory

P.G.L. Porta Mana
Kavli Institute, Trondheim <pgl@portamana.org>

13 July 2020; updated 15 July 2020

In a recent manuscript, Gelman & Yao (2020) claim that "the usual rules of conditional probability fail in the quantum realm", and purport to support that statement with the example of a double-slit experiment. Their statement is false. In fact, an opposite statement can be made, from two different perspectives:

- The example given in that manuscript confirms, rather than invalidates, the probability rules. The probability calculus shows that a particular relation between probabilities, to be discussed below, *cannot a priori* be assumed to be an equality or an inequality. In the quantum example it turns out to be an inequality, thus confirming what the probability calculus says.
- But actually the same inequality can be shown to appear in very non-quantum examples, such as drawing from an urn. Thus there is nothing peculiar to quantum theory in this matter.

In the present note I will prove the two points above, recalling some relevant literature in quantum theory. I shall also correct a couple of wrong or imprecise statements that Gelman & Yao make about quantum physics in their example.

Let me point out at the outset that the rule of conditional probability (and the other two rules, sum and negation) are in fact routinely used in quantum theory, with full validity, especially in problems of state "retrodiction" and measurement reconstruction (Jones 1991; Slater 1995; de Muynck 2002 chs 7, 8; Barnett et al. 2003; Ziman et al. 2006; D'Ariano et al. 2004; see Månsson et al. 2006 § 1 and the rest of the present note for many further references). An example is the inference of the state of a quantum laser given its output through different optical apparatus (Leonhardt 1997).

Similar incorrect claims with similar examples have appeared before in the quantum literature. Bernard O. Koopman (of the Pitman-Koopman theorem for sufficient statistics, Koopman 1936) discussed the falsity of such claims already in 1957. The Introduction in his work is very clear:

> Ever since the advent of modern quantum mechanics in the late 1920's, the idea has been prevalent that the classical laws of probability cease, in some sense, to be valid in the new theory. More or less explicit statements to this effect have been made in large number and by many of the most eminent workers in the new physics [. . .]. Some authors have even gone farther and stated that the formal structure of logic must be altered to conform to the terms of reference of quantum physics [. . .].
>
> Such a thesis is surprising, to say the least, to anyone holding more or less conventional views regarding the positions of logic, probability, and experimental science: many of us have been apt – perhaps too naively – to assume that experiments can lead to conclusions only when worked up by means of logic and probability, whose laws seem to be on a different level from those of physical science.
>
> The primary object of this presentation is to show that the thesis in question is entirely without validity and is the product of a confused view of the laws of probability.

A more recent claim, somewhat similar to Gelman & Yao's and with a similar supporting example, was made in a work by Brukner & Zeilinger (2001) and disproved by Porta Mana (2004) through a step-by-step analysis and calculation. The fallacy in this kind of examples rests in the neglect of the experimental setup, leading either to an incorrect calculation of conditional probabilities, or to the incorrect claim that the probability calculus yields an equality, where it actually does not. The same incorrect claims can be obtained *with completely non-quantum systems*, such as drawing from an urn, if the setup is neglected (Porta Mana 2004 § IV).

Let us start with such a non-quantum counter-example.

A non-quantum counter-example Consider an urn with one *B*lue and one *R*ed ball. Two possible drawing setups are given:

- D_a: With replacement for blue, without replacement for red. That is, if blue is drawn, it is put back before the next draw (and the urn is shaken); if red is drawn, it is thrown away before the next draw.
- D_b: With replacement for red, without replacement for blue.

These two setups are obviously mutually exclusive.

We can easily find the unconditional probability for blue at the second draw in the setup D_a:

$$\mathrm{P}(B_2 \mid D_\mathrm{a}) = \tfrac{3}{4} \;. \tag{1}$$

Note that this probability can be intuitively found by simple enumeration, à la Boole, with a "possible worlds" diagram. Out of four possible worlds, half of which has blue at the first draw, and the other half has red, we can count that three worlds have blue at the second draw.

The conditional probabilities for blue at the second draw, conditional on the first draw, are also easily found:

$$\mathrm{P}(B_2 \mid B_1 \land D_\mathrm{a}) = \tfrac{1}{2} \qquad \mathrm{P}(B_2 \mid R_1 \land D_\mathrm{a}) = 1 \;. \tag{2}$$

We find that

$$\mathrm{P}(B_2 \mid D_\mathrm{a}) = \mathrm{P}(B_2 \mid B_1 \land D_\mathrm{a})\; \mathrm{P}(B_1 \mid D_\mathrm{a}) + \mathrm{P}(B_2 \mid R_1 \land D_\mathrm{a})\; \mathrm{P}(R_1 \mid D_\mathrm{a}) \;, \tag{3}$$

which is just the rule of conditional probability. It is in fact just the systematization and generalization of the intuitive "possible worlds" reasoning done above.

Now consider the setup D_b. We easily find

$$\mathrm{P}(B_2 \mid D_\mathrm{b}) = \tfrac{1}{4} \;, \tag{4}$$

$$\mathrm{P}(B_2 \mid B_1 \land D_\mathrm{b}) = 0 \qquad \mathrm{P}(B_2 \mid R_1 \land D_\mathrm{b}) = \tfrac{1}{2} \;, \tag{5}$$

$$\mathrm{P}(B_2 \mid D_\mathrm{b}) = \mathrm{P}(B_2 \mid B_1 \land D_\mathrm{b})\; \mathrm{P}(B_1 \mid D_\mathrm{b}) + \mathrm{P}(B_2 \mid R_1 \land D_\mathrm{b})\; \mathrm{P}(R_1 \mid D_\mathrm{b}) \;. \tag{6}$$

Now compare the unconditional probability for blue at the second draw in the setup D_a, with the conditional probabilities for blue at the second draw given the first draw in the setup D_b:

$$\mathrm{P}(B_2 \mid D_\mathrm{a}) \neq \mathrm{P}(B_2 \mid B_1 \land D_\mathrm{b})\; \mathrm{P}(B_1 \mid D_\mathrm{b}) + \mathrm{P}(B_2 \mid R_1 \land D_\mathrm{b})\; \mathrm{P}(R_1 \mid D_\mathrm{b}) \;. \tag{7}$$

This inequality is not surprising – we are comparing different setups. It is *not* an instance of the conditional-probability rule. In fact the probability calculus has nothing to say, a priori, about the relation between the left side and right side, which are conditional on different statements or, if you like, pertain to two different sample spaces.

You can call the inequality above "interference" if you want; for further examples see Kirkpatrick (2003a,b) and Porta Mana (2004 § IV).

Now consider another pair of drawing setups: setup D_{c}, with replacement for both colours; and setup D_{d}, without replacement for either colour. You can easily find that

$$\mathrm{P}(B_2 \mid D_{\mathrm{c}}) = \mathrm{P}(B_2 \mid B_1 \land D_{\mathrm{c}})\, \mathrm{P}(B_1 \mid D_{\mathrm{c}}) + \mathrm{P}(B_2 \mid R_1 \land D_{\mathrm{c}})\, \mathrm{P}(R_1 \mid D_{\mathrm{c}}) \;, \tag{8}$$

$$\mathrm{P}(B_2 \mid D_{\mathrm{d}}) = \mathrm{P}(B_2 \mid B_1 \land D_{\mathrm{d}})\, \mathrm{P}(B_1 \mid D_{\mathrm{d}}) + \mathrm{P}(B_2 \mid R_1 \land D_{\mathrm{d}})\, \mathrm{P}(R_1 \mid D_{\mathrm{d}}) \;, \tag{9}$$

$$\mathrm{P}(B_2 \mid D_{\mathrm{c}}) \boldsymbol{=} \mathrm{P}(B_2 \mid B_1 \land D_{\mathrm{d}})\, \mathrm{P}(B_1 \mid D_{\mathrm{d}}) + \mathrm{P}(B_2 \mid R_1 \land D_{\mathrm{d}})\, \mathrm{P}(R_1 \mid D_{\mathrm{d}}) \;. \tag{10}$$

The first two equalities above are expressions of the conditional-probability rule. The third is *not*, however. It is simply a peculiar equality contingent on the two specific setups.

The probability calculus therefore correctly handles situations leading to inequalities such as (7), and to equalities such as (10).

Strictly speaking it is wrong to use the expression 'B_2' for all these setups, because 'B_2' in each setup denotes a different statement (or random variable) than in the others. Just like "it rains (on 14 July 2020 in Trondheim)" is different from "it rains (on 20 November 2019 in Rome)". I should have used different symbols. The explicit presence of '$D_{\dots}$', which represents given information, luckily avoided any ambiguities. But if in our formulae we omit the notation of the setup *and* we use the same notation for actually different statements or random variables, then we're in for trouble and for incorrect applications of the probability rules.

The inequality (7) is what Gelman & Yao (2020 p. 2) complain about, but in the context of a pair of quantum setups. I do not see how one can complain about it, or claim inconsistencies. It is obviously correct even from an intuitive analysis of the two setups. And the probability calculus correctly leads to it, too. The probability calculus correctly leads also to the equality (10). As already said, given two mutually exclusive setups, the probability calculus a priori neither commits to an equality nor to an inequality.

I will now show that the simple example above is in fact conceptually quite close to the quantum experiment mentioned by Gelman & Yao. The closeness is especially clear from the experimental and mathematical developments of quantum theory of the past 40 years (at the very least), as the literature cited below shows.

The quantum two-slit experiments The basic argument of Gelman & Yao is that, in a given setup of the quantum two-slit experiment, we have a specific probability distribution for the appearance of an emulsion or excitation on some point of the screen. We can call this a "screen detection", but please keep in mind that in so doing we are adding an extra interpretation that modern quantum theory does not actually commit to (see discussion and references below). In a different experimental setup we have conditional probabilities for screen detection conditional on slit detection. Now, the probability of the first setup is not equal to the combination of the conditional probabilities of the second setup.

But this is exactly what happened in our urn example above, eq. (7). In the present quantum case we do *not* have a violation of the conditional probability rule either – if anything it is a confirmation.

To see the analogy more clearly, let me present some additional facts from quantum theory.

The experimental setup without detectors at the slits and the setup with slit detectors are actually limit cases of a continuum of experimental setups (Wootters & Zurek 1979; for a recent review and further references see Banaszek et al. 2013). In the general case, such a setup has "noisy" slit detectors, with a degree of noise $q \in \left[\frac{1}{2}, 1\right]$ that can be chosen in the setup. The possible noise degrees are of course mutually exclusive, so these setups are mutually exclusive.

The slit detector is called "noisy" in the following sense:

Let us call y the detection position on the screen; X_1 is the statement that detection occurs at slit #1, and X_2 at slit #2 (you can translate to random-variable jargon if you prefer). We can prepare the electromagnetic field (quantum state S_1) in such a way that, in the setup with non-noisy slit detectors, detector #1 always fires; that is, statement X_1 has probability 1 and statement X_2 has probability 0 in this setup. We can also prepare a state S_2 such that the opposite holds in the same non-noisy setup: X_1 has zero probability, and X_2 unit probability.

If we use the setup with noisy detectors of degree q – denote this setup by D_q – then we have probability q that slit detector #1 fires, when state S_1 is prepared; and probability q that slit detector #2 fires, when state S_2 is prepared; the probabilities for the other slit are $1 - q$. That is,

$$\mathrm{p}(X_1 \mid D_q, S_1) = q\ , \qquad \mathrm{p}(X_2 \mid D_q, S_2) = q\ . \tag{11}$$

The setup with non-noisy detectors is the limit case $q = 1$. In the limit of the fully noisy case, $q = 1/2$, there is basically no relation between the light states and the firing of the slit detectors; that is, we are always fully uncertain as to what detector would fire, no matter how the light state is prepared.

In each setup D_q (and given the light state S) we also have the conditional probability distribution $\mathrm{p}(y \mid D_q, S)$ for detection at y on the screen, and the conditional probability distributions $\mathrm{p}(y \mid X, D_q, S)$ for detection at y on the screen, given detection X at the slits. We have

$$\begin{aligned}&\mathrm{p}(y \mid D_q, S) = \\ &\qquad \mathrm{p}(y \mid X_1, D_q, S)\ \mathrm{p}(X_1 \mid D_q, S) + \mathrm{p}(y \mid X_2, D_q, S)\ \mathrm{p}(X_2 \mid D_q, S)\ .\end{aligned} \tag{12}$$

This is an instance of the conditional-probability rule, which is of course valid. This equality also holds for long-run frequencies (see point (iii) below). Note that such conditional and unconditional frequencies are experimentally observed. I would like you to convince yourself, though, that the *equality* above (not the specific values of the frequencies) is not an experimental fact, since it rests on the very definition of conditional frequency.

The conditional and unconditional distributions above will of course be different depending on the setup D_q and the light state S. But in each instance the rule of conditional probability holds.

Now let me discuss a couple of very interesting experimental facts about this collection of setups:

First, *both the conditional and unconditional* probability (or long-run frequency) distributions for the screen detection y generally *have an oscillatory profile, typical of interference* (Wootters & Zurek 1979; Banaszek et al. 2013; see also Chiao et al. 1995 for further variations). The oscillatory character is maximal for the fully-noisy setup $q = 1/2$ and decreases as $q \longrightarrow 1$; for the non-noisy setup $q = 1$ there is no interference. But we can have quite a lot of interference even when the detection noise is quite low, so that we are almost certain about slit detection; see references above. (The profile depends on the specific light state, of course, which we are assuming fixed.)

Second, the unconditional (frequency) distribution that we observe in the setup without slit detectors is experimentally equal to that observed in the setup $D_{1/2}$ with fully noisy slit detectors.

Third, one conditional distribution observed in the setup with one slit closed is experimentally equal to one in the setup D_1 with perfect slit detectors. (Here we must be careful, because there is no slit detection in the second setup; rather, we speak of appearance or non-appearance at the screen, and in the latter case no conditional distribution is defined.)

The equalities in the last two cases should a priori not be expected, because the setups are physically different. Of course one can look for physical, "hidden variables" explanations of such equalities. Experimental quantum optics simply acknowledges the fact that two setups are equivalent for such detection purposes, and incorporates this information into its mathematical formalism (see below).

Note the statistical analogy between the cases above and the cases with the setups of the urn examples previously discussed. In each setup, the rule of conditional probability holds (and in the quantum case we can have distributions, conditional and unconditional, with oscillatory profiles). Across different setups, probability theory says that such a rule cannot be applied; and indeed we find inequalities across some setups and equalities across others, both in the quantum and non-quantum case, eqs (7), (10).

It is also possible to consider situations in which we are uncertain about which measurement setup applies. For example we may not know whether there were slit detectors, or the value q of the detector noisiness. In such situations we introduce probabilities $\mathrm{p}(D_{\dots})$ for the possible setups and the conditional-probability rule applies, yielding for example

$$\mathrm{p}(y \mid S) = \sum_q \mathrm{p}(y \mid D_q, S)\, \mathrm{p}(D_q) \tag{13}$$

(here our knowledge of the state was assumed to be irrelevant to our inference about the setup). Then, given the measurement outcome, we can make inferences about the setup – for example whether a slit detector was present or not – again using the conditional-probability rule in the guise of Bayes's theorem. This kind of inference is especially important in quantum cryptography and key distribution (Nielsen & Chuang 2010), where we try to infer whether a third party was eavesdropping through covert measurements. Again no violations of the probability rules in the quantum realm; quite the opposite, those rules allow us to make important inferences.

Further remarks and curiosities about quantum two-slits experiments I would like to mention a couple more experimental facts – which are, besides, statistically very interesting – to correct some statements by Gelman & Yao in relation to the two-slit experiment.

(i) It *does* matter whether many photons are sent at once, or one at a time (cf. Gelman & Yao 2020 § 2 point 1); as well as their wavelength, temporal spread, and so on. These details are part of the specification of the light state S mentioned above, and lead to different probabilities distributions of screen detection.

For example, in some setups and for some states we can have a detection probability density $\mathrm{p}(y_1)$ for the first photon, and a *different* density for the second photon $\mathrm{p}(y_2 \mid y_1)$, conditional on the detection of the first – both being different from the long-run joint density of detections $f(y)$. See e.g. the phenomena of higher-order coherence, bunching, anti-bunching, and so on[1]. Interference phenomena can also be observed in time, not only in space. The rules of the probability calculus also apply in all such situations. We can infer, for example, the position of the first photon detection given the second from the conditional probability rule $\mathrm{p}(y_1 \mid y_2) \propto \mathrm{p}(y_2 \mid y_1)\ \mathrm{p}(y_1)$.

(ii) The details about the light source and the setup are not "latent variables": they specify the quantum state of light and the measurement performed on it. They are like the initial and boundary conditions necessary for the specification of the behaviour of any physical system.

(iii) In view of point (i) above, it is important not to conflate the *probabilities* for single-photon detections and the *frequency* distribution of a long-run of such detections (Gelman & Yao 2020 § 2, seem

[1] e.g. Mandel & Wolf 1965; Morgan & Mandel 1966; Paul 1982; Jacobson et al. 1995; and textbooks such as Loudon 2000; Mandel & Wolf 2008; Scully & Zubairy 2001; Bachor & Ralph 2004; Walls & Milburn 1994.

to conflate the two). Such distinction is always important from a Bayesian point of view.

I may add that the idea and parlance of "photons passing through slits" are used today only out of tradition; maybe a little poetically. The technical parlance, as routinely used in quantum-optics labs for example (Leonhardt 1997; Bachor & Ralph 2004), has a different underlying picture. The 'system' in a quantum-optics experiment is not photons, but the modes of the field-configuration operator[1] (note that this is not yet Quantum ElectroDynamics). "Photon numbers" denote the discrete outcomes of a specific energy-measurement operator; "photon states" denote specific states of the field operators. As another example, "entanglement" is strictly speaking not among photons, but among modes of the field operator (van Enk 2003). Several quantum physicists indeed oppose the idea and parlance of "photons", owing to the confusion it leads to. Lamb (of the Lamb shift, Lamb & Retherford 1947) wrote in 1995:

> the author does not like the use of the word "photon", which dates from 1926. In his view, there is no such thing as a photon. Only a comedy of errors and historical accidents led to its popularity among physicists and optical scientists.

Wald (1994) warns:

> standard treatments of quantum field theory in flat spacetime rely heavily on Poincaré symmetry (usually entering the analysis implicitly via plane-wave expansions) and interpret the theory primarily in terms of a notion of "particles". Neither Poincaré (or other) symmetry nor a useful notion of "particles" exists in a general, curved spacetime, so a number of the familiar tools and concepts of field theory must be "unlearned" in order to have a clear grasp of quantum field theory in curved spacetime. [p. ix] [. . .] the notion of "particles" plays no fundamental role either in the formulation or interpretation of the theory. [p. 2]

See also Davies's *Particles do not exist* (1984).

A summary of the modern formalism of quantum theory It may be useful to give a summary of how probability enters the modern formalism of quantum theory. See textbooks such as Holevo (2011), Busch et al. (1995), Peres (1995 especially ch. 12), de Muynck (2002 especially ch. 3), and the excellent text by Bengtsson & Życzkowski (2017).

A quantum system is defined by its sets of possible states and possible measurements. A state ρ is represented by an Hermitean, unit-trace matrix $\boldsymbol{\rho}$ (which satisfies additional mathematical properties), called 'density matrix'. States traditionally represented by kets $|\psi\rangle$ are just special cases of density matrices. A measurement setup M is represented by a set of Hermitean matrices $\{\boldsymbol{M_r}\}$ (of the same order as the density matrices) adding up to the identity matrix. They are called 'positive-operator-valued measures', usually abbreviated POVMs. Traditional von Neumann projection operators $\{|\phi_r\rangle\langle\phi_r|\}$ are just special cases of POVMs. Each matrix $\boldsymbol{M_r}$ is associated with an outcome $\boldsymbol{r}$ of the measurement. These outcomes are mutually exclusive. An outcome can actually represent a combination of simpler outcomes, $\boldsymbol{r} \equiv (x, y, z, \dots)$, such as the intensities at two detectors.

The probability of observing outcome $\boldsymbol{r} \equiv (x, y, \dots)$ given the measurement setup M and the state S is encoded in the trace-product of the respective matrices:

$$\mathrm{p}(x, y, \dots \mid M \land S) \equiv \mathrm{tr}(\boldsymbol{M}_{x,y,\dots}\, \boldsymbol{\rho}) \,, \tag{14}$$

forming a probability distribution. The traditional expression '$|\langle\phi_r|\psi\rangle|^2$' is just a special case of the above formula. The probabilities in the formula come from repeated measurement experiences in the same experimental conditions – we can invoke de Finetti's theorem, and some quantum physicists indeed do (Caves et al. 2002; van Enk & Fuchs 2002; Fuchs et al. 2004). The trace-product above is just a scalar product in a particular space. How a set of probability or frequency distributions can be encoded in scalar products is explained in a down-to-earth way in Porta Mana (2003).

Once the probability distribution above is given we can use the full-fledged probability calculus. We can for example sum (or integrate) over detector outcomes $y, \dots$, obtaining the marginal probabilities for detector outcomes x; or calculate the probability of outcome y conditional on x; or make inferences about the measurement setup or the state. (Again, there are no violations of the probability rules.) The formalism (14) is neat in this respect because it allows us to represent such situations through new POVMs or density matrices. You can easily check, for example, that the marginal probability for x from eq. (14) can be encoded in the POVM $\{\boldsymbol{M}'_x\} \equiv \{\sum_{y,\dots} \boldsymbol{M}_{x,y,\dots}\}$. A situation of uncertainty between setups M'

and M'', as in eq. (13), can be encoded in the POVM $\mathrm{p}(M')\ \boldsymbol{M}'_r + \mathrm{p}(M'')\ \boldsymbol{M}''_r$. And so on, and similarly for states and their density operators.

For systems with infinite degrees of freedom such as electromagnetic fields or electrons (Fermionic fields), the matrices above are actually operators defined in particular algebras; for example, a POVM element can be a space-time-valued operator. The computational details can become quite complicated, but the same basic ideas apply.

This formalism obviously also includes the specification of post-measurement states (if the system still exist afterwards), transformations, and so on. I shall not discuss these; see the textbooks cited above.

Conclusions I hope that the above discussion and bibliography clearly show that:

- the rules of probability theory, including the conditional-probability one, are fully valid in quantum theory;
- some peculiar equalities or inequalities across different experimental conditions do not contradict the conditional-probability rule, and they appear just as well in quantum as in non-quantum systems, such as drawing from an urn.

Quantum theory already has its physically conceptual difficulties and in some cases computational difficulties, as should be clear from the portrait given in the present note. I do not see the point in making it seem even more difficult with false claims of non-validity of probability theory or with distorted pictures of its experimental content.

Bibliography

("de X" is listed under D, "van X" under V, and so on, regardless of national conventions.)

Bachor, H.-A., Ralph, T. C. (2004): *A Guide to Experiments in Quantum Optics*, 2nd rev. and enl. ed. (Wiley-VCH, Weinheim). First publ. 1998.

Banaszek, K., Horodecki, P., Karpiński, M., Radzewicz, C. (2013): *Quantum mechanical which-way experiment with an internal degree of freedom*. Nat. Comm. **4**, 2594.

Barnett, S. M., Jeffers, J., Pegg, D. T. (2003): *Retrodictive quantum optics*. In: Bigelow, Eberly, Stroud, Walmsley (2003), 87–94.

Bengtsson, I., Życzkowski, K. (2017): *Geometry of Quantum States: An Introduction to Quantum Entanglement*, 2nd ed. (Cambridge University Press, Cambridge). First publ. 2006.

Bigelow, N. P., Eberly, J. H., Stroud, C. R., Walmsley, I. A., eds. (2003): *Coherence and Quantum Optics VIII*. (Springer, New York).

Brukner, Č., Zeilinger, A. (2001): *Conceptual inadequacy of the Shannon information in quantum measurements*. Phys. Rev. A **63**, 022113. See also Porta Mana (2004).

Busch, P., Grabowski, M., Lahti, P. J. (1995): *Operational Quantum Physics*. (Springer, Berlin).

Caves, C. M., Fuchs, C. A., Schack, R. (2002): *Unknown quantum states: the quantum de Finetti representation*. J. Math. Phys. $\mathbf{43}^{9}$, 4537–4559.

Chiao, R. Y., Kwia, P. G., Steinberg, A. M. (1995): *Quantum non-locality in two-photon experiments at Berkeley*. Quant. Semiclass. Opt. $\mathbf{7}^{3}$, 259–278.

Christensen, S. M., ed. (1984): *Quantum Theory of Gravity: Essays in honor of the 60th birthday of Bryce S DeWitt*. (Adam Hilger, Bristol).

D'Ariano, G. M., Maccone, L., Lo Presti, P. (2004): *Quantum calibration of measurement instrumentation*. Phys. Rev. Lett. **93**, 250407.

Davies, P. C. W. (1984): *Particles do not exist*. In: Christensen (1984), 66–77.

de Muynck, W. M. (2002): *Foundations of Quantum Mechanics, an Empiricist Approach*. (Kluwer, Dordrecht).

Fuchs, C. A., Schack, R., Scudo, P. F. (2004): *De Finetti representation theorem for quantum-process tomography*. Phys. Rev. A **69**, 062305.

Gelman, A., Yao, Y. (2020): *Holes in Bayesian statistics*. arXiv:2002.06467.

Holevo, A. S. (2011): *Probabilistic and Statistical Aspects of Quantum Theory*, 2nd ed. (Edizioni della Normale, Pisa). First publ. in Russian 1980.

Jacobson, J., Björk, G., Chuang, I., Yamamoto, Y. (1995): *Photonic de Broglie waves*. Phys. Rev. Lett. $\mathbf{74}^{24}$, 4835–4838.

Jones, K. R. W. (1991): *Principles of quantum inference*. Ann. of Phys. $\mathbf{207}^{1}$, 140–170.

Kirkpatrick, K. A. (2003a): *"Quantal" behavior in classical probability*. Found. Phys. Lett. $\mathbf{16}^{3}$, 199–224. First publ. 2001.

— (2003b): *Classical three-box "paradox"*. J. Phys. A $\mathbf{36}^{17}$, 4891–4900. First publ. 2002. See also Ravon, Vaidman (2007) and Kirkpatrick (2007).

— (2007): *Reply to 'The three-box paradox revisited' by T. Ravon and L. Vaidman*. J. Phys. A $\mathbf{40}^{11}$, 2883–2890. See Kirkpatrick (2003b) and Ravon, Vaidman (2007).

Koopman, B. O. (1936): *On distributions admitting a sufficient statistic*. Trans. Am. Math. Soc. $\mathbf{39}^{3}$, 399–409.

— (1957): *Quantum theory and the foundations of probability*. In: MacColl (1957), 97–102.

Lamb Jr., W. E. (1995): *Anti-photon*. Appl. Phys. B $\mathbf{60}^{2\text{–}3}$, 77–84.

Lamb Jr., W. E., Retherford, R. C. (1947): *Fine structure of the hydrogen atom by a microwave method*. Phys. Rev. $\mathbf{72}^{3}$, 241–243.

Leonhardt, U. (1997): *Measuring the Quantum State of Light*. (Cambridge University Press, Cambridge).

Loudon, R. (2000): *The Quantum Theory of Light*, 3rd ed. (Oxford University Press, Oxford). First publ. 1973.

MacColl, L. A., ed. (1957): *Applied Probability*. (McGraw-Hill, New York).

Mandel, L., Wolf, E. (1965): *Coherence properties of optical fields*. Rev. Mod. Phys. $\mathbf{37}^{2}$, 231–287.

— (2008): *Optical coherence and quantum optics*, repr. with corrections. (Cambridge University Press, Cambridge). First publ. 1995.

Månsson, A., Porta Mana, P. G. L., Björk, G. (2006): *Numerical Bayesian state assignment for a three-level quantum system. I. Absolute-frequency data; constant and Gaussian-like priors*. arXiv:quant-ph/0612105.

Morgan, B. L., Mandel, L. (1966): *Measurement of photon bunching in a thermal light beam*. Phys. Rev. Lett. **16**[22], 1012–1015.

Nielsen, M. A., Chuang, I. L. (2010): *Quantum Computation and Quantum Information*, 10th anniversary ed. (Cambridge University Press, Cambridge). First publ. 2000.

Paul, H. (1982): *Photon antibunching*. Rev. Mod. Phys. **54**[4], 1061–1102.

Peres, A. (1995): *Quantum Theory: Concepts and Methods*. (Kluwer, Dordrecht).

Porta Mana, P. G. L. (2003): *Why can states and measurement outcomes be represented as vectors?* `arXiv:quant-ph/0305117`.

— (2004): *Consistency of the Shannon entropy in quantum experiments*. Phys. Rev. A **69**[6], 062108. Rev. version at `arXiv:quant-ph/0302049`.

Ravon, T., Vaidman, L. (2007): *The three-box paradox revisited*. J. Phys. A **40**[11], 2873–2882. See Kirkpatrick (2003b). Unfortunately the arguments of this work are marred by vagueness and contradictions. A reply is given in Kirkpatrick (2007).

Scully, M. O., Zubairy, M. S. (2001): *Quantum Optics*. (Cambridge University Press, Cambridge). First publ. 1997.

Slater, P. B. (1995): *Reformulation for arbitrary mixed states of Jones' Bayes estimation of pure states*. Physica A **214**[4], 584–604.

van Enk, S. J. (2003): *Entanglement of electromagnetic fields*. Phys. Rev. A **67**, 022303.

van Enk, S. J., Fuchs, C. A. (2002): *Quantum state of an ideal propagating laser field*. Phys. Rev. Lett. **88**, 027902.

Wald, R. M. (1994): *Quantum Field Theory in Curved Spacetime and Black Hole Thermodynamics*. (University of Chicago Press, Chicago).

Walls, D. F., Milburn, G. J. (1994): *Quantum Optics*. (Springer, Berlin).

Wootters, W. K., Zurek, W. H. (1979): *Complementarity in the double-slit experiment: Quantum nonseparability and a quantitative statement of Bohr's principle*. Phys. Rev. D **19**[2], 473–484.

Ziman, M., Plesch, M., Bužek, V. (2006): *Reconstruction of superoperators from incomplete measurements*. Found. Phys. **36**[1], 127–156. First publ. 2004.