Deriving the Qubit from Entropy Principles

Adam Brandenburger[*] Pierfrancesco La Mura[†]

Version 03/24/15

The Heisenberg uncertainty principle[1] $\Delta x \cdot \Delta p \geq \hbar/2$ is one of the most famous features of quantum mechanics. It sets a fundamental limit to the extent to which certain pairs of physical variables (such as position x and momentum p) can take definite values in microscopic systems. However, the non-determinism implied by the Heisenberg uncertainty principle — together with other prominent aspects of quantum mechanics such as superposition, entanglement, and nonlocality — poses deep puzzles about the underlying physical reality, even while these same features are at the heart of exciting developments such as quantum cryptography, algorithms, and computing. These puzzles might be resolved if the mathematical structure of quantum mechanics were built up from physically interpretable axioms, but it is not. Conventionally, a formal structure (e.g., unit vectors in Hilbert space) is simply posited and is then seen to work. We contribute to the program to axiomatize quantum mechanics, by proposing two physically-based axioms which together characterize the simplest quantum system, namely the qubit. Our starting point is the class of all no-signaling theories[2], which respect the condition that information cannot be transmitted faster than the speed of light. Each such theory can be regarded as a family of empirical models, where each empirical model specifies a set of possible measurements on the physical system, and, for each measurement, a probability distribution on outcomes. Next, to measure the information in an empirical model, we move to phase space[3] and use Rényi entropy[4], which is a general family of entropy measures that satisfies extensivity (i.e., is additive across statistically independent systems) and includes Shannon entropy[5] as a special case. Within this framework, we take two important features of quantum mechanics and turn them into physically justified axioms. The first axiom is an Uncertainty Principle, stated in terms of entropy. The second axiom is an Unbiasedness Principle, which requires that whenever there is complete certainty about the outcome of a measurement in one of three mutually orthogonal directions, there must be maximal uncertainty about the outcomes in each of the two other directions. We show that the quantum mechanics of a single qubit, as represented by the Bloch sphere, is fully characterized by our two axioms.

[*]Stern School of Business, Polytechnic School of Engineering, Institute for the Interdisciplinary Study of Decision Making, Center for Data Science, New York University, New York, NY 10012, U.S.A., adam.brandenburger@stern.nyu.edu, adambrandenburger.com

[†]HHL – Leipzig Graduate School of Management, 04109 Leipzig, Germany, plamura@hhl.de

The search for underlying principles or axioms yielding quantum mechanics has a long history, including attempts to understand quantum mechanics as a non-classical logic[6] and, in recent years, much work employing information-theoretic principles such as communication complexity[7], information causality[8], information capacity[9], and purification[10].

One goal of this search is to provide insight into why the formal structure of quantum mechanics takes the shape it does. With the rise of the field of quantum information and the discovery of novel information processing algorithms (e.g., for database search[11] and prime factorization[12]), there is another good reason to look for axioms. Finding a set of principles that yield quantum mechanics may offer insight into what kinds of tasks can and cannot be achieved by use of quantum information resources.

An important step in this quest was the formulation of the no-signaling condition[2], which says that the statistics of the outcomes of measurements that Alice makes on her part of a physical system cannot depend on the choices of measurement Bob makes on his part of the system. Any theory that does not satisfy no signaling conflicts with special relativity, which forbids instantaneous transmission of information. But, while quantum mechanics satisfies no signaling, this condition also allows superquantum systems such as the so-called PR boxes[2]. Nevertheless, the family of all no-signaling theories makes a good starting point from which to try to identify quantum theory.

Within the family of no-signaling theories, we want to model the simplest non-trivial system, namely a two-level system such as the spin of a particle. The experimenter can observe a property such as spin in three arbitrarily chosen mutually orthogonal directions x, y, and z. In each direction, the outcome of a measurement is 0 or 1. An empirical model (see Figure 1) gives the frequencies of these outcomes when identical copies of the same two-level system are prepared and one of the three measurements is performed on a given copy. For the x-, y-, and z-directions, the frequencies of the 0 outcome are f_x, f_y, and f_z, respectively.

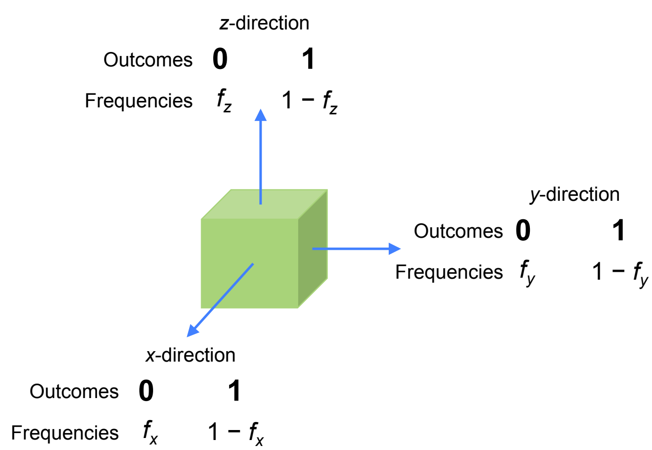

Figure 1: An Empirical Model

Later, we will want to state an Uncertainty Principle in entropic form and, to do this, we need to associate entropies with empirical models. This step is not immediate because entropy is a measure of the information in a single probability distribution, and an empirical model contains three probability distributions (one for each direction). Our solution is to move to phase space, where an empirical model is represented by a single probability distribution. The phase space Ω for a two-level system contains eight points $\omega_{000}, \omega_{001}, \omega_{010}, \ldots, \omega_{111}$, where each point specifies the outcomes (0 or 1) of each of the three possible measurements. We label the points so that if the state is ω_{abc}, the outcomes of the x-, y-, and z- measurements are a, b, and c, respectively. The possibility of a non-deterministic response to measurement — as in quantum mechanics — is allowed for by specifying a probability q_{abc} for each point ω_{abc}. We can now represent a particular empirical model given by frequencies f_x, f_y, f_z as a probability distribution in phase space, where the phase-space probabilities satisfy:

$$q_{000} + q_{001} + q_{010} + q_{011} = f_x, \tag{1}$$

$$q_{000} + q_{001} + q_{100} + q_{101} = f_y, \tag{2}$$

$$q_{000} + q_{010} + q_{100} + q_{110} = f_z. \tag{3}$$

These equations ensure that the phase-space representation reproduces the correct frequencies for outcome 0 (and therefore for 1) in each of the three directions. In general, the phase-space probabilities that reproduce a given empirical model are not unique, but we have achieved the immediate goal of representing each empirical model through a single (if not uniquely defined) probability distribution.

A phase-space representation of an empirical model can be thought of as a particular type of local hidden-variable model[13], where the possible values of the hidden variable are precisely the possible points in phase space. It follows from Bell's Theorem[13] that there are empirical models which arise in quantum mechanics and which cannot be represented in phase space with ordinary non-negative probabilities. We follow Wigner[3], Dirac[4], and Feynman[5] in allowing probabilities to be negative, in which case it can be shown that the family of empirical models which can now be represented is precisely the family of all no-signaling theories[14]. We emphasize that even though the probabilities $q_{000}, q_{001}, q_{010}, \ldots, q_{111}$ are allowed to take negative values, the frequencies of all observable events, namely the sums of probabilities in equations (1)-(3), remain non-negative.

We are now ready to associate entropies with probability distributions on phase space. Within quantum mechanics, the most common entropy measure is the von Neumann entropy[15] $S(\rho) = -\mathrm{Tr}(\rho \log \rho)$ where ρ is the density matrix. This is unsuitable for our purpose because it is defined within the quantum formalism, while we want to derive rather than assume this structure. Shannon entropy[16] is also unsuitable when applied to probabilities in phase space, because the latter can be negative (in which case, Shannon entropy would take complex values). To avoid these shortcomings, we shall work with Rényi entropy[17], which is a general family of entropy measures that includes Shannon entropy as a special case. Rényi entropy satisfies the basic requirement of extensivity, i.e, it is additive across statistically independent systems. In fact, it is defined by this property together with some technical axioms; details are given in the Supplementary Information. The Rényi notion is used in various applications in quantum mechanics[18,19].

To define Rényi entropy, start with a finite probability space, i.e., a finite set $X = \{x_1, \ldots, x_n\}$ together with a probability distribution q on X. Write $q = (q_1, q_2, \ldots, q_n)$ where each $q_i \geq 0$ and $\sum_i q_i = 1$. The Rényi entropy of q is defined as a family of measures indexed by a free parameter $0 < \alpha < \infty$:

$$H_\alpha(q) = \begin{cases} -\frac{1}{\alpha-1} \log_2(\sum_{i=1}^{n} q_i^\alpha) & \text{if} \quad \alpha \neq 1, \\ -\sum_{i=1}^{n} q_i \log_2 q_i & \text{if} \quad \alpha = 1. \end{cases} \tag{4}$$

We see that Shannon entropy is the particular case $\alpha = 1$. In the Supplementary Information, we extend the axioms for Rényi entropy to the domain of signed probability distributions, i.e., to $q = (q_1, q_2, \ldots, q_n)$ where $\sum_i q_i = 1$ but we now allow $q_i < 0$ for some i. This extension identifies a smaller family of Rényi entropies, namely all entropies $H_\alpha(q)$ where α is an even positive integer, i.e. $\alpha = 2k$, for $k = 1, 2, 3, \ldots$. At first sight, it may seem surprising that extending Rényi entropy to a larger domain has the effect of limiting the values of α that yield well-behaved entropies. However, this is correct. When defining an object (here, Rényi entropy) via axioms, if one asks that the axioms hold over a larger vs. smaller set (here, signed vs. unsigned probability distributions), this puts more restrictions on the form the object can take.

With our phase space-Rényi entropy framework in place, we can now choose certain physical properties satisfied by quantum mechanics and express them as axioms in this framework — with the hope that these axioms will then identify the theory. The first property we introduce is an Uncertainty Principle expressed in entropic terms. Differently from other entropic Uncertainty Principles in quantum mechanics[20,21,22,23], our principle is formulated in phase space. Furthermore, we do not derive the principle from quantum mechanics but introduce it as an axiom. Because Rényi entropy is a family of measures of information, indexed by k, we should state our Uncertainty Principle as holding independently of the specific value of k. The principle specifies a lower bound on entropy for each value of k. The values of the lower bounds β_{2k} will be implied by our second axiom.

> **Uncertainty Principle**: An empirical model defined by frequencies f_x, f_y, f_z is allowable if and only if for every $k = 1, 2, 3, \ldots$, there is a phase-space probability distribution q that represents it and that satisfies $H_{2k}(q) \geq \beta_{2k}$.

In quantum mechanics, a set of measurements on a system is called mutually unbiased[24], or complementary, if complete certainty of the measured value of the outcome of one of them implies maximal uncertainty about the outcomes of the others. Mutually unbiased measurements play a central role in the reconstruction of quantum states from observation of outcomes[25] and in public key distribution in quantum cryptography[26]. As we did with the Uncertainty Principle, we now turn this feature of quantum mechanics into an axiom. Specifically, we shall assume that the three measurement directions in the empirical model of Figure 1 form a mutually unbiased set.

> **Unbiasedness Principle**: If there is probability 1 of observing one of the outcomes (0, say) in the x-direction, then there must be probability 1/2 of observing each of the outcomes 0 and 1 in the y-direction, and probability 1/2 of observing each of the outcomes 0 and 1 in the z-direction.

Algebraically, the axiom says that if $f_x = 1$, then $f_y = 1/2$ and $f_z = 1/2$ (see Figure 2). The analogous conditions are required to hold if $f_x = 0$ and when we permute the roles of x, y, and z in the stated condition.

The Unbiasedness Principle implies that the lower bounds β_{2k} in the Uncertainty Principle must be equal to 2 for all k. The proof proceeds by showing that there is a phase-space probability distribution q that represents an empirical model with probabilities $f_x = 1$, $f_y = (1+\epsilon)/2$, $f_z = 1/2$ and that satisfies $H_{2k}(q) > 2 - \delta(\epsilon)$. Here, $\delta(\epsilon) \downarrow 0$ as $\epsilon \downarrow 0$, and therefore, if $\beta_{2k} < 2$, we can choose ϵ sufficiently small (but still positive) so that the Uncertainty Principle holds. But if $\epsilon > 0$, then $f_y \neq 1/2$ and therefore the empirical model violates the Unbiasedness Principle. Details of the proof are in the Supplementary Information.

Summarizing so far, we want to find all empirical probabilities that satisfy the Uncertainty and Unbiasedness Principles. To solve this problem, it will be helpful to re-parameterize the empirical model of Figure 1 by setting $r_j = 2f_j - 1$ for $j = x, y, z$. Since $-1 \leq r_j \leq 1$ for each j, we can now identify each empirical model with a point (r_x, r_y, r_z) in the cube $[-1, 1]^3$. For each Rényi entropy H_{2k}, we define a set $\mathcal{R}_{2k}$ of points in the cube by:

$$\mathcal{R}_{2k} = \{(r_x, r_y, r_z) \in [-1, 1]^3 : \exists\, q \text{ with } H_{2k}(q) \geq 2 \text{ which represents } (r_x, r_y, r_z)\}. \tag{5}$$

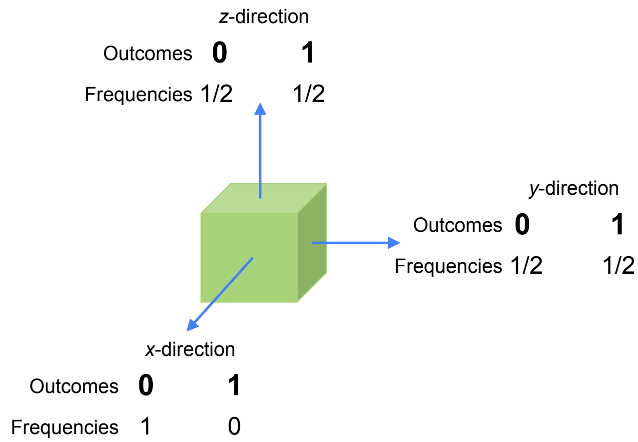

Figure 2: The Unbiasedness Principle

Our problem is to identify the intersection $\cap_k \mathcal{R}_{2k}$ of all these sets. This is the set of empirical models that satisfy our axioms.

The first step is to show that the sequence of sets $\mathcal{R}_2, \mathcal{R}_4, \ldots, \mathcal{R}_{2k}, \mathcal{R}_{2(k+1)}, \ldots$ takes a simple form. If only non-negative phase-space probability distributions q were allowed, then Rényi entropy $H_\alpha(q)$ would be monotonically decreasing in the free parameter α and we would get $H_2(q) \geq H_4(q) \geq \cdots \geq H_{2k}(q) \geq H_{2(k+1)}(q) \geq \cdots$. Because of this inclusion and equation (5), the $\mathcal{R}_{2k}$ sets would then be decreasing: $\mathcal{R}_2 \supseteq \mathcal{R}_4 \supseteq \cdots \supseteq \mathcal{R}_{2k} \supseteq \mathcal{R}_{2(k+1)} \supseteq \cdots$. But we know that in order to represent quantum mechanics and other non-classical theories, we must admit probability distributions with negative entries, and in this case Rényi entropy $H_\alpha(q)$ is no longer monotonic in α (neither decreasing nor increasing). The main mathematical step in the paper (found in the Supplementary Information) uses convex optimization and duality to prove that, once signed probability distributions are allowed, the ordering of the $\mathcal{R}_{2k}$ sets is reversed: $\mathcal{R}_2 \subseteq \mathcal{R}_4 \subseteq \cdots \subseteq \mathcal{R}_{2k} \subseteq \mathcal{R}_{2(k+1)} \subseteq \cdots$. Figure 3 depicts five nested convex regions (starting with blue and ending with red) that correspond to the projections onto the x-y plane of the first five sets in this sequence.

This step tells us that the set $\cap_k \mathcal{R}_{2k}$ of empirical models that satisfy our three axioms is equal to the first set in the sequence, namely $\mathcal{R}_2$. The remaining step (also found in the Supplementary Information) is to identify the set $\mathcal{R}_2$. This step can be formulated as a simple constrained optimization problem and solved in closed form to yield:

$$\mathcal{R}_2 = \{(r_x, r_y, r_z) \in [-1, 1]^3 : r_x^2 + r_y^2 + r_z^2 \leq 1\}. \tag{6}$$

Equation (6) says that the set of empirical models, parametrized by the three numbers r_x, r_y, r_z, coincides with the unit ball in Euclidean ($\mathbb{R}^3$) space. (The projection of the unit ball onto the x-y plane is the blue innermost region in Figure 3.) Call a point in the unit ball a state of our two-level system. Then our result can be stated in the following form: For any choice of directions x, y, z, and choice of frequencies $(1+r_x)/2$, $(1+r_y)/2$, $(1+r_z)/2$ of the 0 outcomes, the system can be prepared in a state such that these frequencies coincide with those obtained from measurements along the three directions. This is exactly the Bloch sphere[27] representation of a qubit, when defined with respect to the same three directions. Observe that the states with minimum entropy (namely, entropy of 2) lie on the surface of the unit ball. These are the pure states in the Bloch sphere

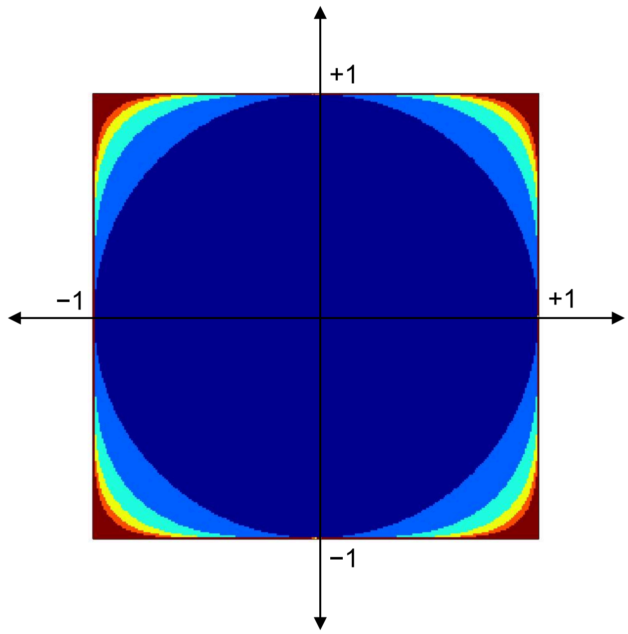

Figure 3: Empirical Models Projected onto the x-y Plane

representation. All interior points in the unit ball have strictly higher entropy and correspond to the mixed states in the Bloch sphere.

A description of the qubit which is equivalent to the Bloch sphere, but cast in terms of signed probabilities, is the discrete Wigner function[28]. In the Supplementary Information we relate the two descriptions and show that our axioms equally identify the single-qubit Wigner function.

In this paper we succeed in axiomatizing only the single qubit. Nevertheless, we believe that our approach can be extended to multiple qubits. In particular, our axioms remain meaningful in the case of n-qubit systems and we hope that, in conjunction with some additional axioms, they will offer a general characterization. A number of related papers[29,30,31,32] treat the n-qubit case, but at the cost of offering partial rather than full characterizations.

We see our axiomatization of the qubit as in line with the program enunciated by Fuchs[33] to find "deep *physical* principles" that yield quantum mechanics. But, in the end, the value of any axiomatization lies in the potential to bring new insights to our understanding of quantum mechanics. We emphasize that there is no claim that our axioms are self-evident. In relativity theory, the principle of light speed invariance is not an intuitive axiom — the point is that it is physically intelligible. Our interest in the Uncertainty Principle — which refers to a limit to definiteness in microscopic systems — as an axiom is that it passes this same test. The same is true of the Unbiasedness Principle. Both axioms are very mysterious at the everyday macroscopic level, but they are physically intelligible — if famously surprising — and evidently true of microscopic systems.

Notes

[1]Heisenberg, W., “Uber den anschulichen Inhalt der quantentheoretischen Kinematik und Mechanik,” *Zeitschrift fur Physik*, 43, 1927, 172-198.

[2]Popescu, S., and D. Rohrlich, “Quantum Nonlocality as an Axiom,” *Foundations of Physics*, 24, 1994, 379-385.

[3]Wigner, E., “On the Quantum Correction For Thermodynamic Equilibrium,” *Physical Review*, 40, 1932, 749-759.

[4]Rényi, A., “On Measures of Information and Entropy,” in Neymann, J., (ed.), *Proceedings of the 4th Berkeley Symposium on Mathematical Statistics and Probability*, University of California Press, 1961, 547-561.

[5]Shannon, C., “A Mathematical Theory of Communication,” *Bell System Technical Journal*, 27, 1948, 379-423 and 623-656.

[6]Birkhoff, G., and J. von Neumann, “The Logic of Quantum Mechanics,” *Annals of Mathematics*, 37, 1936, 823-843.

[7]Van Dam, W., “Implausible Consequences of Superstrong Nonlocality,” 2005, at http://arxiv.org/abs/quant-ph/0501159.

[8]Pawlowski M., T. Paterek, D. Kaszlikowski, V. Scarani, A. Winter, and M. Zukowski, “Information Causality as a Physical Principle,” *Nature*, 461, 2009, 1101-1104.

[9]Dakić, B., and C. Brukner, “Quantum Theory and Beyond: Is Entanglement Special?” in Halvorson, H. (ed.), *Deep Beauty: Understanding the Quantum World through Mathematical Innovation*, Cambridge University Press, 2011, 365-392.

[10]Chiribella, G., G. D'Ariano, and P. Perinotti, “Informational Derivation of Quantum Theory,” *Physical Review A*, 84, 2011, 012311.

[11]Grover, L., “A Fast Quantum Mechanical Algorithm for Database Search,” *Proceedings of the Twenty-Eighth Annual ACM Symposium on Theory of Computing*, 1996, 212-219.

[12]Shor, P., “Polynomial-Time Algorithms for Prime Factorization and Discrete Logarithms on a Quantum Computer,” *SIAM Journal of Computing*, 26, 1997, 1484-1509.

[13]Bell, J., “On the Einstein-Podolsky-Rosen Paradox,” *Physics*, 1, 1964, 195-200.

[14]Abramsky, S., and A. Brandenburger, “The Sheaf-Theoretic Structure of Non-Locality and Contextuality,” *New Journal of Physics*, 13, 2011, 113036.

[15]Von Neumann, J., *Mathematische Grundlagen der Quantenmechanik*, Springer, Berlin, 1932.

[16]Shannon, C., “A Mathematical Theory of Communication,” *Bell System Technical Journal*, 27, 1948, 379-423 and 623-656.

[17]Rényi, A., “On Measures of Information and Entropy,” in Neymann, J., (ed.), *Proceedings of the 4th Berkeley Symposium on Mathematical Statistics and Probability*, University of California Press, 1961, 547-561.

[18]Wehner, S., and A. Winter, “Entropic Uncertainty Relations — A Survey,” *New Journal of Physics*, 12, 2010, 025009.

[19]Bialynicki-Birula, I., and Ł. Rudnicki, “Entropic Uncertainty Relations in Quantum Physics,” in K. Sen (ed.), *Statistical Complexity: Applications in Electronic Structure*, Springer, 2011, 1-34.

[20]Everett, H., *On the Foundations of Quantum Mechanics*, doctoral dissertation, Princeton University, 1957.

[21]Hirschman, I., “A Note on Entropy,” *American Journal of Mathematics*, 79, 1957, 152-156.

[22]Beckner, W., “Inequalities in Fourier Analysis,” *Annals of Mathematics*, 102, 1975, 159-182.

[23]Bialynicki-Birula, I., and J. Mycielski, “Uncertainty Relations for Information Entropy in Wave Mechanics,” *Communications in Mathematical Physics*, 44, 1975, 129-132.

[24]Schwinger, J., “Unitary Operator Bases,” *Proceedings of the National Academy Sciences*, 46, 1960, 570-579.

[25]Ivonovic, I., “Geometrical Description of Quantal State Determination,” *Journal of Physics A*, 14, 1981, 3241-3245.

[26]Bennett, C., and G. Brassard, “Quantum Cryptography: Public Key Distribution and Coin Tossing,” *Proceedings of the IEEE International Conference on Computers, Systems, and Signal Processing*, 1984, 175-179.

[27]Bloch, F., “Nuclear Induction,” *Physical Review*, 70, 1946, 460-474.

[28]Wootters, W., “A Wigner-Function Formulation of Finite-State Quantum Mechanics,” *Annals of Physics*, 176, 1987, 1-21.

[29]Wigner, E., “Quantum Mechanical Distribution Functions Revisited,” in Yourgrau, W., and A. van der Merwe (eds.), *Perspectives in Quantum Theory*, MIT Press, 1971, 25-36.

[30]Hall, M.J.W., and M. Reginatto, “Schrödinger Equation From an Exact Uncertainty Principle,” *Journal of Physics A: Mathematical and General*, 35 2002, 3289.

[31]Brukner, C., and A. Zeilinger, “Information Invariance and Quantum Probabilities,” *Foundations of Physics*, 39, 2009, 677-689.

[32]Oppenheim, J., and S. Wehner, “The Uncertainty Principle Determines the Nonlocality of Quantum Mechanics,” *Science*, 330, 2010, 1072-1074.

[33]Fuchs, C., “Quantum Mechanics as Quantum Information (and Only a Little More),” in Khrennikov, A., (ed.), *Quantum Theory: Reconsideration of Foundations*, Växjö University Press, 2002, 463-543.

Supplementary Information can be found at adambrandenburger.com/articles/papers.

Acknowledgements We are grateful to Matthew Deutsch for excellent research assistance, to Samson Abramsky, Lucy Brandenburger, Časlav Brukner, Eugenia Cerda, Christopher Fuchs, Leslie Greengard, Michael J.W. Hall, H. Jerome Keisler, Elliot Lipnowski, David Pine, Rui Soares Barbosa, Yuri Tschinkel, Noson Yanofsky, and audiences at the Max Planck Institute for Mathematics in the Sciences, Leipzig, the Workshop on Information and Processes, CIAPA, Costa Rica, December 2013, the Workshop on Correlated Information Change, University of Amsterdam, November 2014, and the conference on Quantum Probability and the Mathematical Modelling of Decision Making, Fields Institute, University of Toronto, March 2015, for extremely valuable comments, and to the NYU Stern School of Business, NYU Abu Dhabi, and the HHL – Leipzig Graduate School of Management for financial support.