Discontinuities in the Maximum-Entropy Inference

Stephan Weis

Max Planck Institute for Mathematics in the Sciences,
Inselstrasse 22, D-04103 Leipzig, Germany

Abstract. We revisit the maximum-entropy inference of the state of a finite-level quantum system under linear constraints. The constraints are specified by the expected values of a set of fixed observables. We point out the existence of discontinuities in this inference method. This is a pure quantum phenomenon since the maximum-entropy inference is continuous for mutually commuting observables. The question arises why some sets of observables are distinguished by a discontinuity in an inference method which is still discussed as a universal inference method. In this paper we make an example of a discontinuity and we explain a characterization of the discontinuities in terms of the openness of the (restricted) linear map that assigns expected values to states.

INTRODUCTION

Methods of maximizing entropies have a long tradition. Typically, partial information is available, and a unique inference is desired, in terms of a probability distribution describing the state of a system under observation. The universality of entropy methods is a controversial topic, but from the perspective of the present article, it makes the question interesting why some sets of observables are distinguished through the discontinuity of an entropic inference method. Another important question is whether such observables can be implemented in physical experiments. We have no answers to both questions. In this article we describe a condition of continuity which will certainly be helpful to find more examples theoretically. We stress that this discussion takes place in a finite-dimensional matrix algebra, where entropy functionals are continuous. The Shannon and von Neumann entropies will not be considered in infinite settings [HY, Sh].

Let us dwell upon the idea of universality. The maximization of the Shannon entropy under linear constraints dates back to 1877 with Boltzmann's work about the energy distribution of a particle in a gas. See for example [Uf2, G, L, Ca3]. Jaynes [J] argued that the axioms of the Shannon entropy prove that the method of maximizing entropy is the least biased inference possible if partial information is available. Shore and Johnson [SJ] moved on to axiomatize not the entropy but the inference under partial information, leading to the method, later termed *ME*, of minimizing the Kullback-Leibler divergence under the constraint of partial information, relative to a prior probability distribution describing the state of the system. Later on Skilling [Sk] argued that ME is a universal method of updating a probability distribution (or a positive distribution) given new information in terms of a constraint. The universality comes from the idea of *induction*

in the philosophical sense, that is, finding a theory or a general rule from examples. See also [Ca1, G, Ca3] and see Csiszár [Cs] for axioms of ME under linear constraints, where continuity is postulated (for distributions of full support). Critics of the universality of ME claim that a (relative) *Renyi-entropy* should be minimized in inference [K, Uf1] and that the entropy function depends on a choice of properties that should be preserved during updating the prior [Ca1]. A counterattack against the Renyi-entropies is [CG]. The universality of ME is doubted in particular outside of physics where the wide range of mathematical results about the Shannon entropy, from coding theorems over data compression and much more, *"provide a rationale, as well as several caveats, to the maximum entropy principle"* [L]. The Kullback-Leibler divergence generalizes some mathematical properties of the Shannon entropy, like a game theoretical aspect [T, GD] or a data compression property [BD]. After all, maximum-entropy methods have a wide range of applications, as documented for example in the conference proceedings of *Maximum Entropy and Bayesian Methods*, which make their analysis important.

Von Neumann [vN] has applied for the first time maximum-entropy methods to quantum system. The quantum analogue of ME uses the Umegaki relative entropy [Um]. Although Ochs, Ohya and Petz [O, OP] have given axioms for the von Neumann entropy and the Umegaki relative entropy, no axioms of the inference are known which would lead to the quantum ME. However, several justifications of the quantum ME are known [BB, Ca2, St] and a few references to applications are collected in [AM].

Let us turn to a more detailed characteristics of the discontinuities in question. Given a fixed set of observables and a prior state, the (quantum) ME-inference is, in a suitable restriction, a real-analytic map. This restriction has a continuous extension for mutually commuting observables, as was shown by Barndorff-Nielsen [Ba], p. 154, for the case of probability distributions. So the discontinuities are a pure quantum phenomenon, like for example entanglement. We [WK] have found a pair of non-commutative observables of a three-level quantum system, where this continuous extension is not possible and we [We2] stress that

1. discontinuities are not removable by changing single values of the ME-inference,
2. if the true state of the quantum system has full support then the discontinuities have no consequences in asymptotic state estimation, except possibly about the convergence rate,
3. the continuity of the inference map is characterized by the openness of the restricted linear map assigning expected values to the (fixed) observables,
4. the openness condition has the physical interpretation of tolerance of an ME-inference state for small ambiguity of expected values.

In the following we will explain these statements.

THE ME-INFERENCE

We define the ME-inference. We show why discontinuities of the ME-inference are not removable and we address the asymptotic state estimation.
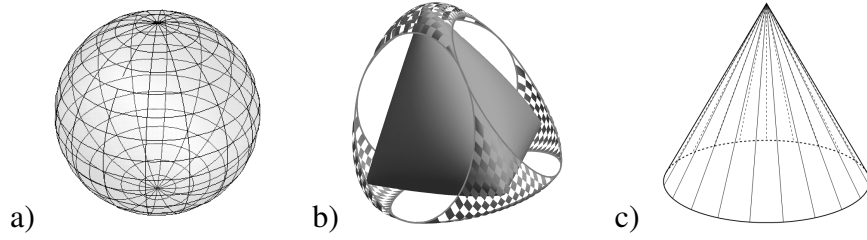
FIGURE 1. a) Bloch ball, b) and c) reductions of $\mathscr{S}(\mathrm{Mat}(3,\mathbb{C}))$.

Mathematically, the state of an *n*-level *quantum system* is described by a linear functional on the C*-algebra $\mathrm{Mat}(n,\mathbb{C})$ of complex $n \times n$-matrices, with identity $\mathbb{1}_n$ and zero 0_n, $n \in \mathbb{N}$. We will use C*-subalgebras $\mathscr{A} \subset \mathrm{Mat}(n,\mathbb{C})$ with identity $\mathbb{1}_n$, for example diagonal matrices $\mathscr{A} \cong \mathbb{C}^n$ describe probability distributions on the sample space $\{1,\dots,n\}$. States on $\mathscr{A}$ are in one-to-one correspondence with *density matrices* in $\mathscr{A}$, that is positive semi-definite matrices of trace one. The *state space* of the algebra $\mathscr{A}$ is the convex body

$$\mathscr{S} = \mathscr{S}(\mathscr{A}) := \{\rho \in \mathscr{A} \mid \rho \text{ is a density matrix}\}.$$

We use the terms of density matrix and state synonymously. The state space of a two-level quantum system is a three-dimensional Euclidean ball, known as *Bloch ball*, see Figure 1 a).

A self-adjoint matrix $a \in \mathscr{A}$ is also known as *observable*, it can be used to make measurements on a quantum system. The real vector space $\mathscr{A}_{\mathrm{sa}} := \{a \in \mathscr{A} \mid a^* = a\}$ of observables is a Euclidean space with scalar product $\langle a,b \rangle := \mathrm{tr}(ab)$, $a,b \in \mathscr{A}_{\mathrm{sa}}$. We insist that the identity of $\mathscr{A}$ is $\mathbb{1}_n$. Then an observable a has spectral decomposition $a = \sum_\lambda \lambda p_\lambda$, where summation runs over all eigenvalues λ of a and $p_\lambda = p_\lambda^* = p_\lambda^2$ is an orthogonal projection such that $p_\lambda p_{\lambda'} = 0$ whenever $\lambda' \neq \lambda$ and $\sum_\lambda p_\lambda = \mathbb{1}_n$. The eigenvalues of a are real. The observable a represents a *simple* or *von Neumann* measurement, where the outcome λ is observed with probability $\langle \rho, p_\lambda \rangle$, given the quantum system is in the state $\rho \in \mathscr{S}$. Therefore, the *expected value* of a is $\langle \rho, a \rangle$, see for example [N]. Given several observables

$$a_1,\dots,a_k, \quad k \in \mathbb{N},$$

we introduce a linear map

$$\mathbb{E} : \mathscr{A}_{\mathrm{sa}} \to \mathbb{R}^k, \quad b \mapsto (\langle b, a_1 \rangle, \dots, \langle b, a_k \rangle).$$

If $\rho \in \mathscr{S}$ is a density matrix, then we call $\mathbb{E}(\rho)$ the *expected value* of ρ (the observables are fixed) and we define the convex body

$$\mathbb{M} := \mathbb{E}(\mathscr{S}).$$

The partial information of an expected value $m \in \mathbb{M}$, may not be sufficient to specify a quantum state. The *(Umegaki) relative entropy* of a state ρ from a state σ is defined by

$$S(\rho,\sigma) := \begin{cases} \mathrm{tr}\,\rho(\log(\rho) - \log(\sigma)) & \text{if } \mathrm{Im}(\rho) \subset \mathrm{Im}(\sigma), \\ +\infty & \text{else,} \end{cases}$$

where Im is the image of a matrix as a linear map in a fixed basis. A self-adjoint matrix $\theta \in \mathscr{A}_{sa}$ defines an invertible *prior state* $\sigma := e^{\theta}/\text{tr}(e^{\theta})$. We assume that observables and prior are fixed. Then the *ME-inference* is defined by

$$\Psi = \Psi_{\theta} : \mathbb{M} \to \mathscr{S}, \quad m \mapsto \text{argmin}\{S(\rho, \sigma) \mid \mathbb{E}(\rho) = m, \rho \in \mathscr{S}\}.$$

Since $S(\rho, \mathbb{1}_n/n) = -H(\rho) + \log(n)$ holds for $\rho \in \mathscr{S}$, where $H(\rho) := \text{tr}(\rho \log(\rho))$ is the *von Neumann entropy*, we call the ME-inference for $\theta = 0$ *maximum-entropy inference*.

We address the irremovability of discontinuities of Ψ. It is well-known that all states

$$R(\lambda) := \frac{\exp(\theta + \lambda_1 a_1 + \cdots + \lambda_k a_k)}{\text{tr} \exp(\theta + \lambda_1 a_1 + \cdots + \lambda_k a_k)}, \quad \lambda = (\lambda_1, \ldots, \lambda_k) \in \mathbb{R}^k$$

are ME-inference states. The coefficients λ_i are Lagrangian multipliers for the expected value constraints and $-\lambda_i$ are considered *generalized inverse temperatures* [I]. What is more, if $\mathbb{1}_n, a_1, \ldots, a_k$ are linearly independent, then

$$\mathbb{R}^k \to \mathbb{M}^{\circ}, \quad \lambda \mapsto \mathbb{E} \circ R(\lambda) \tag{1}$$

is a real-analytic diffeomorphism to the interior $\mathbb{M}^{\circ}$ of $\mathbb{M}$. (If linear independence does not hold then the interior has to be replaced by the interior with respect to the affine hull of $\mathbb{M}$.) For the maximum-entropy inference (1) was proved by Wichmann [Wi] and we [We1] have proved the real analyticity for arbitrary priors. It is known that all elements of the *exponential family* $\mathscr{E} := R(\mathbb{R}^k)$ are ME-inference states and the inclusion $\Psi(\mathbb{M}) \subset \overline{\mathscr{E}}$ of inference states into the norm closure $\overline{\mathscr{E}}$ of the exponential family $\mathscr{E}$ was proved [Wi, We2]. So

$$\mathscr{E} = \Psi(\mathbb{M}^{\circ}) \subset \Psi(\mathbb{M}) \subset \overline{\mathscr{E}} \tag{2}$$

follows and (2) shows that a discontinuity of Ψ, if present, is noticeable by taking the closure of $\mathscr{E}$. To find the discontinuities of Ψ there is no need to evaluate Ψ on the boundary of $\mathbb{M}$, where all its discontinuities lie.

Sample mean values, identified with expected values, can be used to make an inference from a real experiment. We show that discontinuities of the ME-inference do not affect the asymptotical inference. Measuring the observable a on N iid copies of the state ρ of a quantum system, the *sample mean value* of a is defined as the arithmetic mean of the N measurement outcomes, $N \in \mathbb{N}$. By the law of large numbers, the sample mean value of a converges to the expected value of a for $N \to \infty$. If kN iid copies of ρ are available, we can measure N times each of the observables $\{a_i\}_{i=1}^k$. Their joint sample mean value will converge to $\mathbb{E}(\rho) \in \mathbb{R}^k$. Whenever a sample mean value lies outside of the set $\mathbb{M}$ of expected values, we can follow Petz's suggestion [P] and apply the (Lipschitz-continuous) method of least squares to estimate an expected value $m_N \in \mathbb{M}$. (More generally, we can use a continuous estimator which is the identity on $\mathbb{M}$.) If ρ has full support, that is ρ is invertible as a linear map, then the expected value $\mathbb{E}(\rho)$ lies in $\mathbb{M}^{\circ}$, see for example [We2]. Therefore, for large N, the conversion of a sample mean value into an expected value m_N and subsequently into an ME-inference state $\Psi(m_N)$ is, as a composition of the continuous estimation, the inverse diffeomorphism (1) and the parametrization R, a continuous mapping. In particular, the ME-inference states $\Psi(m_N)$

converge to the state on the exponential family $\mathscr{E}$ with expected value $\mathbb{E}(\rho)$ for $N \to \infty$. It is possible that the convergence rate of the sequence $\{\Psi(m_N)\}_{N \in \mathbb{N}}$ becomes worse the closer $\mathbb{E}(\rho)$ lies at a discontinuity of Ψ.

THE OPENNESS CONDITION

We explain the openness condition about continuity of the ME-inference, we make an example within the simplest possible algebra, and we address a physical interpretation of the openness condition.

We call the restricted linear map $\mathbb{E}|_{\mathscr{S}}$ *open at* $\rho \in \mathscr{S}$ if for all neighborhoods $U \subset \mathscr{S}$ of ρ the image $\mathbb{E}(U)$ is a neighborhood of $\mathbb{E}(\rho)$ in $\mathbb{M}$. Here we use the norm topology, restricted to $\mathscr{S}$ respectively to $\mathbb{M}$.

Lemma. If Ψ is continuous at $m \in \mathbb{M}$ then $\mathbb{E}|_{\mathscr{S}}$ is open at $\Psi(m)$.

Proof. Let U be a neighborhood of $\Psi(m)$. Then $\Psi^{-1}(U)$ is a neighborhood of m by continuity of Ψ. By definition of Ψ we have $\mathbb{E}(U) \supset \Psi^{-1}(U)$, completing the proof. $\square$

The converse of the lemma is true, it follows from the optimality of the ME-inference (minimization of the relative entropy) and uniqueness of the minimizer [We2].

We recall the simplest algebra to make an example of a discontinuous ME-inference. We have shown [We2] that the ME-inference is continuous for all balls $\mathscr{S}$ as well as for all polytopes $\mathbb{M}$. Since the state space of a two-level quantum system is the Bloch-ball and the state space of a commutative algebra is a simplex (where $\mathbb{M}$ is a polytope) a discontinuity of Ψ is only possible for a C*-algebra which properly includes the algebra $\mathrm{Mat}(2, \mathbb{C})$ of a two-level quantum system. Before we arrive at a suitable three-dimensional state space of a subalgebra of $\mathscr{A} := \mathrm{Mat}(3, \mathbb{C})$, let us recall that the eight-dimensional state space $\mathscr{S}(\mathscr{A})$ is not a ball. This is demonstrated by the section of $\mathscr{S}(\mathscr{A})$ with the plane $\{\frac{\mathbb{1}_3}{3} + \begin{pmatrix} 0 & x & y \\ x & 0 & z \\ y & z & 0 \end{pmatrix} \mid x, y, z \in \mathbb{R}\}$, depicted[1] in Figure 1 b) inside the projection of $\mathscr{S}(\mathscr{A})$ to the same plane. The section is an inflated tetrahedron, the projection has four disks on its boundary, mutually intersecting in six points [BZ].

The observables we [WK] have found generating a discontinuous ME-inference, are

$$a_1 := \sigma_1 \oplus 0, \qquad a_2 := \sigma_2 \oplus 1, \tag{3}$$

interpreted as 3×3-blockdiagonal matrices with block sizes two and one. Here we have used Pauli matrices $\sigma_1 := \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$, $\sigma_2 := \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}$ and $\sigma_3 := \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}$. The simplest C*-algebra containing $\mathbb{1}_3, a_1, a_2$ is the direct sum of $\mathrm{Mat}(2, \mathbb{C})$ and $\mathbb{C}$ with a state space equal to a four-dimensional cone based on the Bloch ball. The matrices $\mathbb{1}_3, a_1, a_2$ belong to a smaller algebra $\mathscr{B}$, which is the real *-subalgebra spanned by

$$\sigma_1 \oplus 0, \qquad \sigma_2 \oplus 0, \qquad i\sigma_3 \oplus 0, \qquad \mathbb{1}_2 \oplus 0, \qquad \mathbb{1}_3.$$

---

[1] With kind permission of Springer Science+Business Media. Reprinted from [BZ].
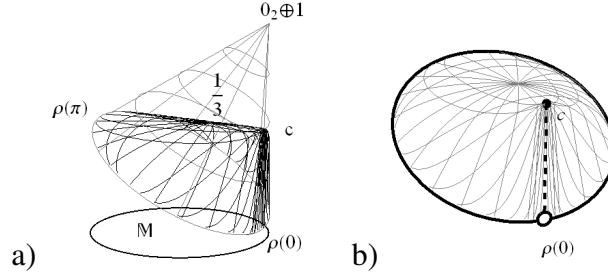
FIGURE 2. The Staffelberg family a) inside the cone $\mathscr{S}(\mathscr{B})$ and b) extended to $\Psi_0(\mathbb{M})$.

The algebra $\mathscr{B}$ is isomorphic to $\mathrm{Mat}(2,\mathbb{R}) \oplus \mathbb{R}$ (by exchanging σ_2 and σ_3). We extend the above definitions from a C*-algebra to a real *-algebra. The state space $\mathscr{S}(\mathscr{B})$ is a three-dimensional cone, depicted in Figure 1 c). Its directrix (base circle) is parametrized for real α by

$$\rho(\alpha) := \tfrac{1}{2}(\mathbb{1}_2 + \sin(\alpha)\sigma_1 + \cos(\alpha)\sigma_2) \oplus 0.$$

Since the exponential family $\mathscr{E}$, defined previously, is included in $\mathscr{S}(\mathscr{B})$ and since $\Psi(\mathbb{M})$, computed in the algebra $\mathrm{Mat}(3,\mathbb{C})$, is a subset of the norm closure $\overline{\mathscr{E}}$ by (2), we have $\Psi(\mathbb{M}) \subset \mathscr{S}(\mathscr{B})$ because $\mathscr{S}(\mathscr{B})$ is norm closed. Therefore the ME-inference with respect to a_1, a_2 and with prior in $\mathscr{B}$ can be studied equivalently in the three-dimensional cone $\mathscr{S}(\mathscr{B})$ or in the eight-dimensional state space of $\mathrm{Mat}(3,\mathbb{C})$.

We compute a full set $\Psi_0(\mathbb{M})$ of maximum-entropy inference states. For the prior $\mathbb{1}_3/3$ and observables a_1, a_2 from (3), the exponential family $\mathscr{E}$ is called *Staffelberg family* in [WK]. The generatrix (surface line) $[\rho(0), 0_2 \oplus 1]$ of the cone $\mathscr{S}(\mathscr{B})$ is perpendicular to $\mathrm{span}_{\mathbb{R}}\{a_1, a_2\}$, its midpoint is denoted c. See[2] Figure 2 a) for a drawing of $\mathscr{E}$ inside the cone $\mathscr{S}(\mathscr{B})$ with the set of expected values underneath, such that $\mathbb{E}$ acts by vertical projection along $[\rho(0), 0_2 \oplus 1]$. Maximum-entropy inference $\Psi_0(m)$ for m in the interior $\mathbb{M}^\circ$ of expected values covers the exponential family $\mathscr{E}$. For $m_0 := (0,1) = \mathbb{E}(\rho(0)) = \mathbb{E}(0_2 \oplus 1)$ we have $\Psi_0(m_0) = c$ and $\Psi_0(m)$ lies on the directrix of $\mathscr{S}(\mathscr{B})$ for boundary points $m \neq m_0$. The maximum-entropy inference $\Psi_0 : \mathbb{M} \to \mathscr{S}$ jumps at m_0 [WK]. This can be seen[2] in Figure 2 b), where $\Psi_0(\mathbb{M})$ is depicted, consisting of $\mathscr{E}$, of the pointed circle $\rho(\alpha)$ for $\alpha \in (0, 2\pi)$ and of c. The segment $[\rho(0), c) := \{(1-\lambda)\rho(0) + \lambda c \mid 0 \leq \lambda < 1\}$, dashed in the figure, belongs to the norm closure $\overline{\mathscr{E}}$ and not to $\Psi_0(\mathbb{M})$.

We arrive at the same conclusion of a discontinuity at $m_0 = (0,1)$ through the openness condition. We have $\Psi_0(m_0) = c$ and the neighborhood $U(c) := \{\rho \in \mathscr{S}(\mathscr{B}) \mid \langle \rho, 0_2 \oplus 1 - \rho(0)\rangle \geq -\tfrac{1}{3}\}$, shown in Figure 3 a), proves that $\mathbb{E}|_{\mathscr{S}(\mathscr{B})}$ is not open at c, because the image $\mathbb{E}(U)$ in Figure 3 b) has, at $\mathbb{E}(c) = m_0$ a larger boundary curvature than $\mathbb{M}$. Hence the above lemma proves that the maximum-entropy inference Ψ_0 is not continuous at m_0.

The openness of $\mathbb{E}|_{\mathscr{S}}$ has a natural interpretation in empirical quantum state estimation. Before we can apply the ME-inference to a sample mean value we have to estimate

---

[2] Reprinted with permission from [WK]. Copyright 2012, American Institute of Physics.
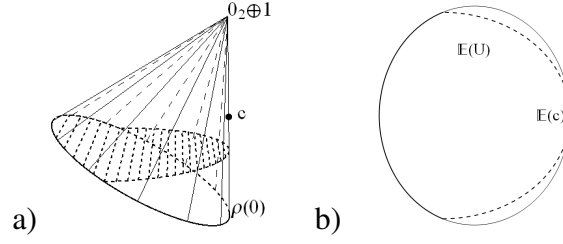
FIGURE 3. a) The upper part U above the dashed separation is a neighborhood of c in the cone $\mathscr{S}(\mathscr{B})$. b) The image $\mathbb{E}(U)$ is not a neighborhood of $\mathbb{E}(c)$.

an expected value in $\mathbb{M}$, see the last section. The estimation provides a finite sequence (indexed by sample size) of expected values $m_1, \ldots, m_N \in \mathbb{M} \subset \mathbb{R}^k$, $N \in \mathbb{N}$ and the (theoretical) asymptotical sequence converges to the expected value $m := \mathbb{E}(\rho)$ of the true state ρ of the given quantum system for $N \to \infty$. The openness of $\mathbb{E}|_{\mathscr{S}}$ at $\Psi(m) \in \mathscr{S}$ means that a small ambiguity of m, resulting for example from the identification of $m = m_N$ for a finite sample length $N \in \mathbb{N}$, can be balanced by a small adjustment of $\Psi(m)$. See [Sh] for this point of view in a different context.

ACKNOWLEDGMENTS

REFERENCES

AM. Ali, S. A., Cafaro, C., Giffin, A., Lupo, C. and Mancini S.: *On a differential geometric viewpoint of Jaynes' MaxEnt method and its quantum extension*, AIP Conf. Proc. **1443** 120–128 (2012)

BB. Balian, R. and Balazs, N. L.: *Equiprobability, Inference and Entropy in Quantum Theory*, Ann. Phys. **179** 97–144 (1987)

Ba. Barndorff-Nielsen, O.: *Information and Exponential Families in Statistical Theory*, John Wiley & Sons, New York (1978)

BD. Benedetto, D., Caglioti, E., Cristadoro, G. and Degli Esposti, M.: *Relative entropy via non-sequential recursive pair substitution* J. Stat. Mech. P09010 (2010)

BZ. Bengtsson, I., Weis, S. and Życzkowski, K.: *Geometry of the set of mixed quantum states: An apophatic approach*, Geometric Methods in Physics. XXX Workshop 2011, Trends in Mathematics 175–197, Springer Basel (2013)

Ca1. Caticha, A.: *Relative Entropy and Inductive Inference*, AIP Conf. Proc. **707** 75–96 (2004)

Ca2. Caticha, A.: *Notes on Quantum Relative Entropy*, unpublished, University of Albany - SUNY (2004)

Ca3. Caticha, A.: *Entropic Inference and the Foundations of Physics*, Brazilian Chapter of the International Society for Bayesian Analysis-ISBrA, Sao Paulo, Brazil (2012)

CG. Caticha, A. and Giffin, A.: *Updating Probabilities*, AIP Conf. Proc. **872** 31–42 (2006)

Cs. Csiszár, I.: *Why Least Squares and Maximum Entropy? An Axiomatic Approach to Inference for Linear Inverse Problems*, Ann. Statist. **19** 2032–2066 (1991)

G. Giffin, A.: *Maximum Entropy: The Universal Method for Inference*, Ph.D. Thesis, SUNY at Albany, Albany (New York - USA), `http://arxiv.org/abs/0901.2987` (2008)

GD. Grünwald, P. D. and Dawid, A. P.: *Game theory, maximum entropy, minimum discrepancy and robust Bayesian decision theory*, Ann. Statist. **32** 1367–1433 (2004)

HY. Ho, S.-W. and Yeung, R. W.: *On the Discontinuity of the Shannon Information Measures*, IEEE Trans. Inf. Theory *55* 5362–5374 (2009)

I. Ingarden, R. S., Kossakowski, A. and Ohya, M.: *Information Dynamics and Open Systems*, Kluwer Academic Publishers Group (1997)

J. Jaynes, E. T.: *Information Theory and Statistical Mechanics*, Phys. Rev. **106** 620–630 (1957)

K. Karbelkar, S. N.: *On the axiomatic approach to the maximum entropy principle of inference*, Pramāna J. Phys. **26** 301–310 (1986)

L. Lesne, A.: *Statistical entropy: At the crossroads between probability, information theory, dynamical systems and statistical physics*, Preprint

N. Nielsen, M. A. and Chuang, I. L.: *Quantum Computation and Quantum Information*, Cambridge University Press (2000)

O. Ochs, W.: *A new axiomatic characterization of the von Neumann entropy*, Rep. Math. Phys. **8** 109–120 (1975)

OP. Ohya, M. and Petz, D.: *Quantum entropy and its use*, Springer-Verlag, New York (1993)

P. Petz, D.: *Quantum Information Theory and Quantum Statistics*, Theoretical and Mathematical Physics, Springer-Verlag, Berlin (2008)

Sh. Shirokov, M. E.: *Continuity of the von Neumann Entropy*, Commun. Math. Phys. **296** 625–654 (2010)

SJ. Shore, J. E. and Johnson, R. W.: *Axiomatic derivation of the principle of maximum entropy and the principle of minimum cross-entropy*, IEEE Trans. Inf. Theory **26** 26–37 (1980); correction **29** 942–943 (1983)

Sk. Skilling, J.: *Classic Maximum Entropy*, in Maximum Entropy and Bayesian Methods, ed. Skilling, J., Kluwer, Dordrecht (1989)

St. Streater, R. F.: *Proof of a Modified Jaynes's Estimation Theory*, Open Syst. Inf. Dyn. **18** 223–233 (2011)

T. Topsøe, F.: *Information Theoretical Optimization Techniques*, Kybernetika **15** 8–27 (1979)

Uf1. Uffink, J.: *Can the maximum entropy principle be explained as a consistency requirement?*, Stud. Hist. Phil. Sci. B **26** 223–261 (1995)

Uf2. Uffink, J.: *Compendium of the foundations of classical statistical physics*, in Handbook for Philsophy of Physics, ed. Butterfield, J. and Earman, J. (2006)

Um. Umegaki, H.: *Conditional expectation in an operator algebra, IV*, Kōdai Math. Sem. Rep. **14** 59–85 (1962)

vN. von Neumann, J.: *Thermodynamik quantenmechanischer Gesamtheiten*, Gött. Nach. 273–291 (1927)

We1. Weis, S.: *Information Topologies in Non-Commutative State Spaces*, Journal of Convex Analysis, in press (2014)

We2. Weis, S.: *Continuity of the maximum-entropy inference*, Preprint available at `http://weis-stephan.de/`

WK. Weis, S. and Knauf, A.: *Entropy Distance: New Quantum Phenomena*, J. Math. Phys. **53** 102206 (2012)

Wi. Wichmann, E. H.: *Density matrices arising from incomplete measurements*, J. Math. Phys. **4** 884–896 (1963)