

# On Maximizing the Information from an Indirect Experiment

G. L. Hofacker \*

Lehrstuhl für Theoretische Chemie, Technische Universität München

R. D. Levine \*\*

Department of Physical Chemistry, The Hebrew University Jerusalem, Israel

Z. Naturforsch. **35a**, 490–492 (1980); received February 16, 1980

The information about an observable  $A$  provided by an indirect experiment in which observable  $B$  is measured is used as a criterion for comparing the performance of different observables  $B$  and for selecting an optimal initial state.

## I. Introduction

In all but the simplest systems, the observables of interest cannot be directly measured. One therefore conducts indirect experiments where an observable  $B$  is measured in order to infer the results of a measurement of  $A$ . Examples include determining the conformation of molecules by resonance techniques or characterising the density of states of condensed systems by structure functions derived from scattering experiments.

Obviously, indirect experiments differ in their resolving power according to the observable  $B$  chosen and in the preparation of the initial state, the latter being itself limited by those observables which are subject to experimental control. It is the purpose of this paper to propose a quantitative formulation on the basis of which the amount of information provided by different experiments can be compared. In particular, we show that an optimal preparation of the initial state can be made for any choice of the observable  $B$ .

In his pioneering study Shannon [1, 2] defined the amount of information given by  $X$  about  $Y$ , where  $X$  and  $Y$  are random variables. Here we are rather concerned with quantum mechanical observables which, moreover, may fail to commute. Even so, by explicitly recognising the role of the initial state of the system it is possible to provide a definition suitable for the quantum measurement process [3]. For a wide class of initial states this

definition leads to an explicit algebraic expression which is quite similar to the classical one, though care must be taken since in quantum mechanics it is amplitudes that are superposed so that Bayes' theorem should not be used in algebraic manipulations.

## 2. Uncertainty and Information

Consider an observable  $A$  with a discrete spectrum. The amount of information obtained by measuring  $A$  for a system in the state  $W$  is [3–6]

$$H[A; W] = - \sum_{\alpha} p_{\alpha} \ln p_{\alpha} . \quad (2.1)$$

Here  $p_{\alpha}$  is the probability of observing the value  $\alpha$  when  $A$  is measured. In terms of the spectral resolution of  $A$ ,

$$A = \sum_{\alpha} \alpha P_{\alpha} \quad (2.2)$$

we have that

$$p_{\alpha} = \text{Tr}(P_{\alpha} W) . \quad (2.3)$$

Prior to the measurement of  $A$ ,  $H[A; W]$  is the uncertainty regarding the results of a measurement of  $A$ . Some of this uncertainty can, however, be removed by a measurement of some other observable  $B$ . Say the value of  $\beta$  has been obtained for  $B$ . The state of the system following this measurement is  $W_{\beta}$

$$W_{\beta} = P_{\beta} W P_{\beta} / \text{Tr}(P_{\beta} W) . \quad (2.4)$$

The uncertainty regarding  $A$  after the value  $\beta$  has been obtained for  $B$  is now  $H[A; W_{\beta}]$

$$H[A; W_{\beta}] = - \sum_{\alpha} p_{\alpha\beta} \ln p_{\alpha\beta} \quad (2.5)$$

\* Work supported by the Stiftung Volkswagenwerk and Fonds der Chemischen Industrie.

\*\* Work supported by the Office of Naval Research.

Reprint requests to Prof. Dr. G. L. Hofacker, Lichtenbergstraße 4, D-8046 Garching.

0340-4811 / 80 / 0500-0490 \$ 01.00/0. — Please order a reprint rather than making your own copy.



Dieses Werk wurde im Jahr 2013 vom Verlag Zeitschrift für Naturforschung in Zusammenarbeit mit der Max-Planck-Gesellschaft zur Förderung der Wissenschaften e.V. digitalisiert und unter folgender Lizenz veröffentlicht: Creative Commons Namensnennung-Keine Bearbeitung 3.0 Deutschland Lizenz.

Zum 01.01.2015 ist eine Anpassung der Lizenzbedingungen (Entfall der Creative Commons Lizenzbedingung „Keine Bearbeitung“) beabsichtigt, um eine Nachnutzung auch im Rahmen zukünftiger wissenschaftlicher Nutzungsformen zu ermöglichen.

This work has been digitalized and published in 2013 by Verlag Zeitschrift für Naturforschung in cooperation with the Max Planck Society for the Advancement of Science under a Creative Commons Attribution-NoDerivs 3.0 Germany License.

On 01.01.2015 it is planned to change the License Conditions (the removal of the Creative Commons License condition “no derivative works”). This is to allow reuse in the area of future scientific usage.

where

$$p_{\alpha\beta} = \text{Tr}(P_\alpha W_\beta). \quad (2.6)$$

Since, however, any value from the spectrum of  $B$  may, in principle, be possible one needs to average  $H[A; W_\beta]$  over the distribution of outcomes for  $B$  in the state  $W$ . Thus, the uncertainty regarding  $A$  which remains after  $B$  has been measured is  $H[A|B, W]$

$$H[A|B, W] = \sum_{\beta} p_{\beta} H[A; W_{\beta}], \quad (2.7)$$

$$p_{\beta} = \text{Tr}(P_{\beta} W). \quad (2.8)$$

The amount of information on  $A$  provided by the measurement of  $B$  is the reduction in the uncertainty

$$I[A; BW] = H[A; W] - H[A|B, W]. \quad (2.9)$$

Note that we have not required that  $A$  and  $B$  commute. As a consequence it is not necessarily the case that the amount of information provided by  $B$  about  $A$  equals the amount of information provided by  $A$  and  $B$ . The result (2.9) does, however, satisfy obvious consistency conditions, e.g.

$$I[A; AW] = H[A; W] \quad (2.10)$$

or the measurement of  $A$  removes all the uncertainty regarding  $A$ ; and

$$I[A; BW_{\beta}] = 0 \quad (2.11)$$

or repeated and successive measurements of the same observable  $B$  (for a stationary system) provides no additional information.

The uncertainty about  $A$  when  $B$  has been measured depends only on the state  $W$  of the system following the measurement of  $B$ ,

$$\bar{W} = \sum_{\beta} p_{\beta} W_{\beta}, \quad (2.12)$$

$$H[A|B, W] = H[A|B, \bar{W}]. \quad (2.13)$$

The amount of information about  $A$  provided by  $B$ , in the state following measurement of  $B$

$$I[A; B\bar{W}] = H[A; \bar{W}] - H[A|B, \bar{W}], \quad (2.14)$$

differs from (2.9) only in the first term. Hence, if the initial state has been prepared by a measurement of  $B$  (without, however, selecting the subensembles  $W_{\beta}$ ), the information provided about  $A$  is given by (2.14).

### 3. Preparation of the Initial State to Achieve Maximal Information

This section shows that there exists a unique preparation of the initial state that achieves the maximal information about  $A$  via a measurement of  $B$ . We furthermore construct such an initial state explicitly.

In terms of  $p_{\alpha}$  (cf. Eq. (2.3)),  $p_{\beta}$  and  $p_{\alpha\beta}$  (cf. Eq. (2.6)), we have

$$I[A; B, \bar{W}] = \sum_{\alpha, \beta} p_{\beta} p_{\alpha\beta} \ln(p_{\alpha\beta}/p_{\alpha}), \quad (3.1)$$

which follows rapidly by noting that for the state  $\bar{W}$  (cf. (2.12)) it is \*

$$\begin{aligned} p_{\alpha} &= \text{Tr}(P_{\alpha} \bar{W}) = \sum_{\beta} p_{\beta} \text{Tr}(P_{\alpha} W_{\beta}) \\ &= \sum_{\beta} p_{\beta} p_{\alpha\beta}. \end{aligned} \quad (3.2)$$

The result (3.1) enables us to conclude that (i)

$$I[A; B, \bar{W}] \geq 0 \quad (3.3)$$

with equality if and only if  $W \equiv W_{\beta}$  for some  $\beta$  and (ii) that  $I[A; B, \bar{W}]$  is a convex function of the  $p_{\beta}$ 's or, in other words, that it is possible to find a unique set of  $p_{\beta}$ 's which maximise the amount of information provided by the measurement.

The proofs are based on the observation that the function  $x \ln x$  is strictly convex for all positive  $x$ . Hence, for any set of non-negative  $p_{\beta}$ 's such that

$$\sum_{\beta} p_{\beta} = 1 \quad (3.4)$$

it follows that for  $p_{\alpha\beta} \geq 0$

$$\sum_{\beta} p_{\beta} p_{\alpha\beta} \ln p_{\alpha\beta} \geq (\sum_{\beta} p_{\beta} p_{\alpha\beta}) \ln (\sum_{\beta} p_{\beta} p_{\alpha\beta}). \quad (3.5)$$

The result (3.3) is obtained upon use of (3.2). Similarly, it follows from the standard discussions in the literature (e.g. theorem 3.3.1 of [2]) that (3.1) is strictly convex in the  $p_{\beta}$ 's.

Taking the partial derivative of (3.1) with respect to  $p_{\beta}$  we have that

$$\partial I[A; B, \bar{W}] / \partial p_{\beta} = \sum_{\alpha} p_{\alpha\beta} \ln(p_{\alpha\beta}/p_{\alpha}) - 1. \quad (3.6)$$

The term  $-1$  is the contribution from the  $p_{\alpha}$  term in the logarithm which (cf. (3.2)) depends on  $p_{\beta}$ . Since  $I[A; B, \bar{W}]$  is strictly convex its unique maximum is the solution of

$$\partial \{I[A; B, \bar{W}] - \lambda \sum_{\beta} p_{\beta}\} / \partial p_{\beta} = 0 \quad (3.7)$$

\* In the derivations below we shall also implicitly use the result that for the choice (2.12), the  $p_{\alpha\beta}$ 's are independent of the  $p_{\beta}$ 's.

where  $\lambda$  is the Lagrange parameter that ensures the normalization of the  $p_\beta$ 's. The maximum is thus the solution of

$$\sum_{\alpha} p_{\alpha\beta} \ln(p_{\alpha\beta}/p_{\alpha}) = \lambda + 1 \quad (3.8)$$

for  $p_{\alpha}$ . The significance of this result becomes clearer if we define

$$I[A; \beta, \bar{W}] = \sum_{\alpha} p_{\alpha\beta} \ln(p_{\alpha\beta}/p_{\alpha}), \quad (3.9)$$

so that

$$I[A; B, \bar{W}] = \sum_{\beta} p_{\beta} I[A; \beta, \bar{W}]. \quad (3.10)$$

One can therefore identify  $I[A; \beta, \bar{W}]$  as the amount of information about  $A$  given by obtaining the particular value  $\beta$  in the measurement of  $B$ .

What we have shown is that for the initial state  $\bar{W}$  that achieves the maximal information, measuring any one of the eigenvalues  $\beta$  of  $B$  provides exactly the same information about  $A$ . Intuitively, this is quite obvious. Say some particular projector  $P_{\bar{\beta}}$  gave more information than all others, then one should try to increase the weight,  $p_{\bar{\beta}}$ , of that subspace in  $W$ . Since  $I$  is convex in these weights, the maximum is achieved when all subspaces yield the same information.

To explicitly construct the initial state that achieves the maximum, we define the matrix  $q$  as the inverse of the matrix  $\{p_{\alpha\beta}\}$

$$\sum_{\beta} p_{\alpha\beta} q_{\beta\alpha'} = \delta_{\alpha\alpha'}. \quad (3.11)$$

Consider then the choice

$$p_{\beta}^c = \sum_{\alpha} q_{\beta\alpha} \exp\left\{-\sum_{\beta'} q_{\beta'\alpha} H[A; W_{\beta'}] - C\right\} \quad (3.12)$$

corresponding to which

$$\begin{aligned} p_{\alpha}^c &= \sum_{\beta} p_{\beta}^c p_{\alpha\beta} \\ &= \exp\left\{-\sum_{\beta'} q_{\beta'\alpha} H[A; W_{\beta'}] - C\right\}. \end{aligned} \quad (3.13)$$

Here  $C$  insures normalisation

$$\exp(C) = \sum_{\alpha} \exp\left\{-\sum_{\beta'} q_{\beta'\alpha} H[A; W_{\beta'}]\right\}. \quad (3.14)$$

$C$  can also be interpreted as the value of  $I$  for the initial state  $\bar{W}^c$  which is of the form (2.12) with (3.12) taken as the values of the  $p_{\beta}$ 's. The proof is

by substitution

$$\begin{aligned} I[A; B, \bar{W}^c] &= \sum_{\alpha, \beta} p_{\beta}^c p_{\alpha\beta} \ln(p_{\alpha\beta}/p_{\alpha}^c) \\ &= C - \sum_{\alpha, \beta} p_{\beta}^c p_{\alpha\beta} \sum_{\beta'} q_{\beta'\alpha} H[A; W_{\beta'}] \\ &\quad + \sum_{\alpha, \beta} p_{\beta}^c p_{\alpha\beta} \ln p_{\alpha\beta} = C. \end{aligned} \quad (3.15)$$

At this point it is easy to show that  $C$  is the unique maximal value of the information and hence that the initial state  $\bar{W}^c$  achieves that maximum. To prove this consider the non-negative quantity  $DI$ ,

$$DI = \sum_{\alpha} p_{\alpha} \ln(p_{\alpha}/p_{\alpha}^c) \geq 0, \quad (3.16)$$

that vanishes if and only if  $p_{\alpha} = p_{\alpha}^c$  for all  $\alpha$ . Using (3.13) and (3.1)

$$\begin{aligned} DI &= C - I[A; B, \bar{W}] \\ &= I[A; B, \bar{W}^c] - I[A; B, \bar{W}] \geq 0. \end{aligned} \quad (3.17)$$

Our ansatz (3.12) for  $p_{\beta}^c$  leads therefore to a maximal value of  $I$ . Since the maximum is unique, so is our specification (3.12) of the initial state.

#### 4. Conclusions

We have shown that the experimentalist interested in a quantity  $A$  which he cannot measure directly may judge the efficiency of an indirect experiment by computing  $I[A; B, W]$ . In fact, since  $H[A; W]$  is common to all such computations, all he really need compute is  $H[A|B, W]$  which does not require the a-priori unknown  $p_{\alpha}$ 's. Having once decided which observable  $B$  to use, the  $P_{\alpha}$ 's are computed from (3.2) and the amount of information provided from (3.1). Moreover, for any given  $B$ , the experimentalist can optimise his choice of the initial state of the system. The resulting set of  $p_{\alpha}$ 's are then given by (3.13). In other words, (3.13) is the unique optimal prediction about  $A$  that is possible in an indirect experiment where  $B$  is measured. On this basis, he can compute the optimal expectation value of  $A$  or of any function of  $A$ ,

$$\begin{aligned} \langle f(A) \rangle_c &= \sum_{\alpha} f(\alpha) \exp\left\{-\sum_{\beta} q_{\beta\alpha} H[A; W_{\beta}] - C\right\}, \\ H[A; W_{\beta}] &= -\sum_{\alpha} p_{\alpha\beta} \ln p_{\alpha\beta}. \end{aligned}$$

- [1] C. E. Shannon, Bell System Tech. J. **27**, 379 (1948).
- [2] R. Ash, Information Theory, Interscience, New York 1965.
- [3] J. von Neumann, Mathematical Foundations of Quantum Mechanics, Princeton University Press, New Jersey 1965.

- [4] D. S. Lebedev and L. B. Levitin, Information and Control **9**, 1 (1966).
- [5] G. Lindblad, Comm. Math. Phys. **33**, 305 (1973).
- [6] R. S. Ingarden, Rep. Math. Phys. **10**, 43 (1976).