# Statistical inference, signal analysis, and state distribution

A. Plastino and L. Rebollo Neira
*Departamento de Física, Universidad Nacional de La Plata, Casilla de Correo No. 67, La Plata 1900, Argentina*

A. G. Alvarez
*Centro de Investigación y Desarrollo en Procesos Catalíticos, 47 No. 257 La Plata 1900, Argentina*
(Received 19 July 1988; revised manuscript received 20 March 1989)

A method for finding the state distribution in a physical system is presented, on the basis of a detailed analysis of the experimental response of the system to an external probe. As the study of the signals with which the experimenter probes the system of interest necessarily consists of a finite number of measurements, the maximum entropy principle is employed in order to reconstruct the state distribution on the basis of partial information. The formalism is illustrated by recourse to an elementary problem. A rather complicated x-ray diffraction problem is also discussed as it neatly depicts the power of this new approach.

## I. INTRODUCTION

The word "signal" applies to any preconcerted or intelligible sign conveying information. Here we will deal with signals that convey information with reference to the states of physical systems, and in regard to the possibility of observing them. Indeed, we wish to discuss signals containing information about the state distribution of a given system and will call these type of signals "statistical" ones. The basic idea is that a well-known probe (i.e., electromagnetic radiation) impinges upon the system, interacts with it, and is afterwards analyzed via a convenient detection procedure. As a consequence of this interaction, the signals acquire information about the state distribution of our system.

In order to make our considerations independent of any specific detection procedure we shall adopt the vectorial representation of signals of Shannon,[1] and establish a unique correspondence between "observed" quantities and measurements performed upon the corresponding signal. This is achieved by representing a signal $f$ as a vector ket $|f\rangle$ and a measurement as a mapping that assigns to it a real number. We shall restrict our considerations to measurements that can be represented by linear functionals.

The question that we wish to address is that of determining, for a given system, its state distribution on the basis of a finite number of measurements. To this end we adopt the maximum-entropy postulate (MEP) that characterizes Shannon's information theory[2–4] (IT) and introduce a general procedure that enables one to find an adequate basis for the representation of the kets $|f\rangle$. Both a simple example and a rather sophisticated technical application will illustrate the ideas to be expounded in this work, which is organized as follows: Section II defines the concept of a statistical signal and develops a formalism (based upon the MEP) in order to deal with them. In Sec. III this formalism is applied to a problem involving paramagnetism and, in Sec. IV, to a crystallographic problem by proposing a suitable model that addresses it. Section V deals with a realistic situation that poses a severe test to our approach and some conclusions are drawn in Sec. VI.

## II. FORMALISM

### A. Statistical signals

We take the view that in order to study a given physical system $\mathcal{S}$, one interacts with it by means of an input signal (probe) $|I\rangle$, whose properties are assumed to be well known (i.e., a radio-frequency signal). The signal-$|I\rangle$-system-$\mathcal{S}$ interaction results in a response signal $|f\rangle$ which, after analysis, should provide information concerning the system. The corresponding process is represented according to

$$\hat{W}|I\rangle = |f\rangle \ , \tag{2.1}$$

where the linear operator $\hat{W}$ (which is here associated to the system under study) portrays the effect that $\mathcal{S}$ produces upon the input signal $|I\rangle$, so as to originate the response $|f\rangle$.

If one were to know the properties of the system $\mathcal{S}$ in all possible detail, $|f\rangle$ could be predicted and no experiment would be needed. Of course, this is almost never the case and we must study $|f\rangle$ in order to learn something about $\mathcal{S}$ (at the risk of being redundant we must insist that we study here not the system itself (that is represented by $\hat{W}$), but its action upon $|I\rangle$ as represented by $|f\rangle$).

Our goal is that of finding the state distribution of $\mathcal{S}$ by recourse to a finite number of measurements performed upon $|f\rangle$. We may wish to investigate systems that can be found in a number of possible states (governed by a probability distribution) or systems that consist of a number of identical subsystems, each of them being in a different state. No distinction will be made between these two instances, as they are characterized by a common factor: our ignorance concerning the relevant state distribution.

We then start our considerations by assuming that we know not only the set of all possible state in which $\mathcal{S}$ may be found (we shall label them $\alpha$), but also the response $|\alpha\rangle$ evoked by $\alpha$ when inpinged upon by the input signal $|I\rangle$. We shall denote by $\hat{W}_\alpha$ the operator representing $\mathcal{S}$ in the state "$\alpha$," so that one describes its action upon $|I\rangle$ as

$$\hat{W}_\alpha|I\rangle=|\alpha\rangle \ . \tag{2.2}$$

Let $p_\alpha$ stand for the probability of finding the system in the particular state $\alpha$. We can write then $\hat{W}$, the operator representing the action of $\mathcal{S}$ upon $|I\rangle$ as

$$\hat{W}=\sum_\alpha p_\alpha \hat{W}_\alpha, \quad \sum_\alpha p_\alpha=1 \ . \tag{2.3}$$

The response $|f\rangle$ originated by a system whose action upon $|I\rangle$ is represented by an operator $\hat{W}$ of the form (2.3) will be referred to as a statistical signal

$$\hat{W}|I\rangle=\sum_\alpha p_\alpha \hat{W}_\alpha|I\rangle=\sum_\alpha p_\alpha|\alpha\rangle=|f\rangle \ . \tag{2.4}$$

It is clear from (2.4) that the signal $|f\rangle$ carries information about the state distribution in $\mathcal{S}$. Underlying this description is the assumption that $|\alpha\rangle$ belongs to an orthogonal basis. It is precisely by representing the statistical signal in this basis that we hope to determine the state distribution. In order to accomplish such a goal one needs to perform observations upon $|f\rangle$. The corresponding measurement procedures provide numbers out of $|f\rangle$ and can consequently be thought of as mappings of a vector space upon the real line, i.e., as functionals.

### B. Measurements as observables

We assume that measurements performed over $|f\rangle$ can be represented by linear functionals $\mathcal{L}$. Then

$$\mathcal{L}|f\rangle=\sum_\alpha p_\alpha \mathcal{L}|\alpha\rangle \ . \tag{2.5}$$

We define a "statistical operator" $\hat{\rho}$ by

$$\hat{\rho}=\sum_\alpha |\alpha\rangle p_\alpha\langle\alpha| \tag{2.6}$$

which allows one to write

$$\mathcal{L}|f\rangle=\mathrm{Tr}(\hat{\rho}\hat{L})\equiv\langle\hat{L}\rangle \tag{2.7}$$

with

$$\hat{L}=\sum_\alpha |\alpha\rangle l_\alpha\langle\alpha| \tag{2.8}$$

and

$$l_\alpha=\mathcal{L}|\alpha\rangle \ . \tag{2.9}$$

The eigenkets $|\alpha\rangle$ are eigenvectors of $\hat{L}$. If the situation is such that $\{|\alpha\rangle\}$ is a complete set, we can, by means of a quantum-mechanical analogy, call $\hat{L}$ an observable.[5] Measurements upon $\hat{L}$ can only yield one of its eigenvalues and vice versa (an eigenvalue of $\hat{L}$ is the result of a measurement performed upon a signal). As in quantum mechanics, only $\langle\hat{L}\rangle$ is generally available.

If the input signal is kept fixed, a set $\{\mathcal{L}_i\}$ of $M$-independent measurements performed upon $|f\rangle$ will yield the expectation values of $M$ mutually commuting observables. In practical situations, such measurements will, in general, be proportional to these expectation values

$$f_i=\mathcal{L}_i|f\rangle=A\langle\hat{L}_i\rangle, \quad i=1,\ldots,M \tag{2.10}$$

where $A$ is a proportionality constant.

On the other hand, if the input signal changes, the corresponding measurements will be compatible only when the response signals associated with them belong to a common basis set.

### C. An inversion method

Our objective should be clear by now: We wish to find the statistical operator on the basis of a finite number, $M$, of (linear) measurements performed upon a given signal. We face then the set of equations,

$$f_i=A\,\mathrm{Tr}(\hat{\rho}\hat{L}_i), \quad i=1,\ldots,M \tag{2.11}$$

where

$$f_i=\mathcal{L}_i|f\rangle \ . \tag{2.12}$$

As $A$ is an (in general) unknown constant, we need one of the Equations (2.11) so we can fix it. If we choose the first of these for this task,

$$A=f_1/\mathrm{Tr}(\hat{\rho}\hat{L}_1) \ , \tag{2.13}$$

we can recast (2.11) as

$$\mathrm{Tr}(\hat{\rho}\hat{O}_i)=\langle\hat{O}_i\rangle=0, \quad i=2,\ldots,M \tag{2.14}$$

with

$$\hat{O}_i=f_i\hat{L}_1-f_1\hat{L}_i, \quad i=2,\ldots,M \ . \tag{2.15}$$

Now, an infinite set of statistical operators is compatible, in general, with a finite set of equations of the type (2.14). In order to determine $\hat{\rho}$ in an unique fashion, we resort here to the MEP, by defining first the entropy $S$ as

$$S=-\mathrm{Tr}(\hat{\rho}\ln\hat{\rho}) \tag{2.16}$$

and employing afterwards the IT method. This allows one to write[4,6]

$$\hat{\rho}=\exp\left[\lambda_0-\sum_{i=2}^{M}\lambda_i\hat{O}_i\right] \ , \tag{2.17}$$

where the $\lambda_i$ are Lagrange multipliers that are connected with the set (2.14). On the other hand, the multiplier $\lambda_0$ is associated to the normalization condition

$$\mathrm{Tr}\hat{\rho}=1 \ , \tag{2.18}$$

which yields

$$\lambda_0=-\ln\left[\mathrm{Tr}\exp\left[-\sum_{i=2}^{M}\lambda_i\hat{O}_i\right]\right] \ . \tag{2.19}$$

Once we have $\hat{\rho}$ (built upon the basis of a set of $M$ measurements) we can predict the result of additional mea-

surements (not already employed to construct $\hat{\rho}$)

$$f_{M+1}=f_1\mathrm{Tr}(\hat{\rho}\hat{L}_{M+1})/\mathrm{Tr}(\hat{\rho}\hat{L}_1)\ ,$$
$$f_{M+j}=f_1\mathrm{Tr}(\hat{\rho}\hat{L}_{M+j})/\mathrm{Tr}(\hat{\rho}\hat{L}_1)\ . \tag{2.20}$$

If the predictions turns out to be right, that is, if they are experimentally verified, then we may conclude that we have been able to obtain, by recourse to this statistical inference approach, a good representation of $\hat{\rho}$. On the other hand, the same measurement that proves our predictions (2.20) wrong also supplies new useful information, and with this information we may try again (for a better $\hat{\rho}$ and so on). If for a given selection of the initial input data we finally achieve a good representation of the statistical operator, we will be in a position to determine expectation values corresponding to any additional operator $\hat{R}$,

$$\langle\hat{R}\rangle=\mathrm{Tr}(\hat{\rho}\hat{R})\ . \tag{2.21}$$

### D. Convergence questions

Assume we have at our disposal an ordered set $\{f_i\}$ of, say, $K$ experimental measurements. With just two of these we can build up, as explained above, a MEP statistical operator $\hat{\rho}^{(1)}$. If we add a third piece of data to the original two we are in a position to construct still another MEP statistical operator $\hat{\rho}^{(2)}$, etc. A whole series $\hat{\rho}^{(n)}$ can be manufactured in this fashion. We wish to establish a convergence test for this process. Let $f_i^q$ stand for the prediction for $f_i$ made with the statistical operator $\hat{\rho}^{(q)}$. We will assert that $\{\hat{\rho}^n\}$ converges to a "true" $\hat{\rho}$ (for the available experimental information) if, given $\epsilon_i>0$, there exists an integer $M_0$ such that

$$|f_i-f_i^M|<\epsilon_i \quad \text{for all } M\geq M_0,\quad i=2,\ldots,K\ . \tag{2.22}$$

This kind of convergence can always be achieved if we deal with a set $\{f_i\}$ of linearly independent measurements, as in such a situation IT asserts that the input information is not self-contradictory, which guarantees both the existence and the unicity of the statistical operator.[6] This entails that (2.22) must be fulfilled, at least for $M=M_0=K$. However, we will show in Secs. III B and V that $M_0\ll K$ in practice.

## III. A WEIGHTING METHOD FOR RARE-EARTH MIXTURES

Ions of rare-earth elements possess similar chemical properties, which entails that isolating one of them from a mixture does not constitute an easy task. As a simple illustration of the foregoing formalism, we propose here a weighting method for rare-earth mixtures, based on the quantum theory of paramagnetism.[7,8]

### A. Brief review of elementary concepts

The total angular momentum $h\mathbf{J}$ of an atom or ion is conventionally written in the form

$$\boldsymbol{\mu}=-g\mu_B\mathbf{J}\ , \tag{3.1}$$

where $\mu_B$ is a standard unit of magnetic moment (here the Bohr magneton). The spectral factor $g$ is given by Lande's equation

$$g=1+\frac{J(J+1)+S(S+1)-L(L+1)}{2J(J+1)} \tag{3.2}$$

for atoms or ions of orbital angular momentum $L$ and spin $S$. $(2J+1)$ is the degeneration of the ground state. If we have $N$ noninteracting atoms placed in an external magnetic field $H$ pointing along the $z$ direction, the mean magnetic moment per unit volume, or magnetization $\mathcal{M}$, is conveniently expressed in terms of the Brillouin function

$$B_J(x)=\frac{2J+1}{2J}\coth\left[\frac{(2J+1)x}{2J}\right]-\frac{1}{2J}\coth\left[\frac{x}{2J}\right]\ , \tag{3.3}$$

with

$$x=gJ\mu_BH/k_BT\ , \tag{3.4}$$

where $k_B$ stands for Boltzmann's constant and $T$ is the temperature. We have

$$\mathcal{M}=NgJ\mu_BB_J(x)\ . \tag{3.5}$$

The magnetic properties of rare-earth ions are quite interesting. In particular, triply ionized elements possess almost identical chemical properties as their external shells become identical: a $5s^2,5p^6$ configuration (that of neutral $X_e$). In La, just before the lanthanide series begin, the $4f$ shell is empty. It acquires one electron in Ce and the number of $4f$ electrons steadily augments as we run along the series. We attain 13 electrons for Yb and completely fill the shell for the last rare-earth element (Lu). The corresponding radii decrease in regular fashion, starting at a value of 1.11 Å (Ce) up to 0.94 Å (Yb), the celebrated "lanthanide contraction". The magnetic behavior of these ions is governed by the filling up of the $4f$ shell. No other group of the Periodic Table exhibits such a peculiar magnetic behavior.[8]

### B. Application

We apply here the formalism previously introduced in order to propose a method for rare-earth weighting, based upon the idea of measuring the magnetization $\mathcal{M}$ as a function of the external magnetic field $H$. To this end, we shall run a simple numerical experiment.

Consider that we have a mixture of 11 different rare-earth elements, their respective proportions in the mixture being denoted by $p_\alpha$ $(\alpha=1,\ldots,11)$. For each $\alpha$ we list the corresponding quantum numbers $S_\alpha$, $L_\alpha$, and $J_\alpha$ in Table I. Our goal is to show that the ideas introduced in Sec. II can be employed to experimentally determine the figures $p_\alpha$, starting from appropriate measurements of the magnetization $\mathcal{M}$ (at the temperature $T$) for a series of values $H_i$.

We deal thus with two relevant subindexes, namely $\alpha$, which runs along the actinides and $i$, which indicate a

TABLE I. For each rare-earth ion the pertinent quantum numbers are given. The relative properties in the rare-earth mixture are denoted by $p_\alpha$ (system $\mathcal{S}$) and $p'_\alpha$ (system $\mathcal{S}'$). The diagonal elements of the statistical operator constitute the theoretical predictions for these weights.

| $\alpha$ | Ion | $S_\alpha$ | $L_\alpha$ | $J_\alpha$ | $p_\alpha$ | $\langle\alpha\vert\hat{\rho}^3\vert\alpha\rangle$ | $p'_\alpha$ | $\langle\alpha\vert(\hat{\rho}')^5\vert\alpha\rangle$ |
|---|---|---|---|---|---|---|---|---|
| 1 | $Ce^{3+}$ | $\frac{1}{2}$ | 3 | $\frac{5}{2}$ | 0.003 | 0.007 | 0.549 | 0.549 |
| 2 | $Pr^{3+}$ | 1 | 5 | 4 | 0.004 | 0.004 | 0.329 | 0.330 |
| 3 | $Nd^{3+}$ | $\frac{3}{2}$ | 6 | $\frac{4}{2}$ | 0.005 | 0.003 | 0.076 | 0.076 |
| 4 | $Pm^{3+}$ | 2 | 6 | 4 | 0.006 | 0.005 | 0.042 | 0.043 |
| 5 | $Gd^{3+}$ | $\frac{7}{2}$ | 0 | $\frac{7}{2}$ | 0.055 | 0.057 | 0.000 | 0.000 |
| 6 | $Tb^{3+}$ | 3 | 3 | 6 | 0.156 | 0.157 | 0.000 | 0.000 |
| 7 | $Dy^{3+}$ | $\frac{5}{2}$ | 5 | $\frac{15}{2}$ | 0.338 | 0.339 | 0.000 | 0.000 |
| 8 | $Ho^{3+}$ | 2 | 6 | 8 | 0.301 | 0.299 | 0.000 | 0.000 |
| 9 | $Er^{3+}$ | $\frac{3}{2}$ | 6 | $\frac{15}{2}$ | 0.101 | 0.104 | 0.000 | 0.000 |
| 10 | $Tm^{3+}$ | 1 | 5 | 6 | 0.019 | 0.018 | 0.000 | 0.000 |
| 11 | $Yb^{3+}$ | $\frac{1}{2}$ | 3 | $\frac{7}{2}$ | 0.005 | 0.004 | 0.000 | 0.000 |

given measurement of $\mathcal{M}$ (that for which the external field is $H_i$).

In order to apply the techniques of Sec. II we first of all need an expression for the operator $\hat{L}$ of Eq. (2.8). We set for each $i$ ($\vert\alpha\rangle \equiv \vert S_\alpha L_\alpha J_\alpha\rangle$)

$$\hat{L}_i = \sum_{\alpha=1}^{11} \vert\alpha\rangle l_{i\alpha}\langle\alpha\vert \ , \tag{3.6}$$

$$l_{i\alpha} = \mu_B g_\alpha J_\alpha B_{i\alpha} \ , \tag{3.7}$$

where $g_\alpha$ is the spectral factor for the ion $\alpha$, $J_\alpha$ is the corresponding angular momentum, and $B_{i\alpha}$ is the appropriate Brillouin function. Notice that here the identity operator (that we need in order to evaluate traces) reads

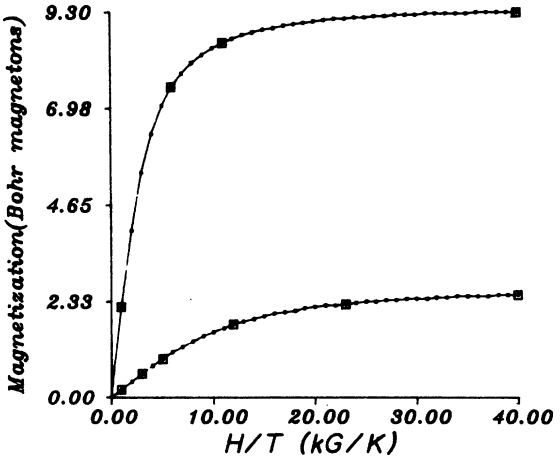


FIG. 1. Magnetization vs external applied field at the temperature $T$. The dots denote numerically simulated values and the squares indicate the figures actually employed in building up $\hat{\rho}^{M-1}$. Theoretical values evaluated with the statistical operator are given by the solid line. Curve (a) correspond to system $\mathcal{S}$ and curve (b) to system $\mathcal{S}'$.

$$\hat{I} = \sum_{\alpha=1}^{11} \vert\alpha\rangle\langle\alpha\vert \ . \tag{3.8}$$

Assume $i$ runs from 1 up to $M$ ($M$ different measurements, i.e., $M$ different values of $M$). Our inversion method (Sec. II) entails dealing with the set of equations [cf. Eq. (2.10)]

$$f_i = A \ \mathrm{Tr}(\hat{\rho}\hat{L}_i), \quad i = 1, \ldots, M \ . \tag{3.9}$$

Two sets of "weights" $\{p_\alpha\}$ and $\{p'_\alpha\}$ are listed in Table I which correspond to two hypothetical mixtures $\mathcal{S}$ and $\mathcal{S}'$. By recourse to these sets we have numerically simulated a series of measurements of the magnetization [dots in Fig. 1, where (a) corresponds to $\mathcal{S}$ and (b) to $\mathcal{S}'$]. These series, in turn, constitute the input information in order to build up, via the MEP, our (approximate) statistical operator $\hat{\rho}^{M-1}$ (on the basis of a finite number, $M$, of measurements). As discussed in Sec. II C, we can check the accuracy of the predictions arising from $\hat{\rho}^{M-1}$ by comparing magnetization values not employed in constructing $\hat{\rho}^{M-1}$ within the appropriate theoretical figures that can be calculated via $\hat{\rho}^{M-1}$ [solid line in Figs. 1(a) and 1(b)].

A glance at the pertinent drawings allows one to conclude that the quantity of these predictions is indeed excellent if we take $M=4$ for $\mathcal{S}$ and $M=6$ for $\mathcal{S}'$ [the small squares in Figs. 1(a) and 1(b) indicate the magnetization values employed as input in our calculations]. Finally, the diagonal elements $\langle\alpha\vert\hat{\rho}^{M-1}\vert\alpha\rangle$ yield the desired values of the relative weight of each ion in our hypothetical mixture. Table I shows that these diagonal elements provide one with a good approximation to $\{p_\alpha\}$ (or $\{p'_\alpha\}$). Obviously, the quality of our figures can be improved by appropriately increasing $M$.

We conclude, by recourse to this simple situation, that our formalism does indeed provide one with reliable results in connection with this particular weighting problem.

## IV. A FANCIER APPLICATION

In order to provide the reader with a more realistic illustration of our formalism, we have chosen a rather in-

volved crystallographic problem, that of the description of mixed layering. We introduce the problem in Sec. IV A and propose a model for addressing it in Secs. IV B and IV C. A specific situation is discussed in Sec. V.

### A. Mixed layering

The terms interlayering, mixed layering, and interstratification describe phyllosilicate structures in which two or more types of layer occur in a vertical stacking sequence, that is a line normal to the (001) plane. Phyllosilicate layers are strongly bonded internally, but rather weakly to each other. Thus, each layer approximates a one-dimensional molecule in the stacking direction, and a two-dimensional one in the $a$ and $b$ directions. The basal surfaces of different kinds of layers are geometrically very similar, consisting of sheets of oxygen or hydroxyl ions disposed in a quasihexagonal array.[9]

Although interstratificates of more than two components have been reported,[10–12] we will deal here with just two-component systems (layer 1 and layer 2). Layer 1 is characterized by an interplanar spacing $d_1$, while $d_2$ does this for layer 2.

X-ray diffraction patterns are quite sensitive to the way these sequences of layers are arranged. Thus, these patterns provide detailed experimental information concerning the internal structure of the interstratificate. Input x radiation of wavelength $\lambda$ of the order of $d_1$ ($d_2$) induces a response signal from the sample that fits nicely into the considerations expounded above and pose a severe experimental test that they must pass.

### B. Response signals from interstratificates in second quantization language

Given $N$ layers, $2^N$ different sequences exist. However, $N$ need not be the same for all plaques in the sample. For this reason we shall resort now to a Fock-Hilbert space, for example,

$$|i_1,i_2,i_3,\ldots,i_j,\ldots\rangle \tag{4.1}$$

where the subindex $j$ along the row indicates the layer one is concerned with and $i_j$ indicates whether its type is 1 ($i_j=1$) or 2 ($i_j=2$). However, and in keeping with the philosophy of the present work, we shall think of (4.1) as the response signal arising from such a configuration. Number operators appear in a natural fashion[13] as

$$\hat{n}_j^{\,1}|i_1,i_2,\ldots,i_j,\ldots\rangle=(2-i_j)|i_1,i_2,\ldots,i_j,\ldots\rangle\ , \tag{4.2}$$

$$\hat{n}_j^{\,2}|i_1,i_2,\ldots,i_j,\ldots\rangle=(i_j-1)|i_1,i_2,\ldots,i_j,\ldots\rangle\ . \tag{4.3}$$

Out of these we construct the number-of-layers operators,

$$\hat{N}_1=\sum_{i=1}^{N}\hat{n}_i^{\,1}\ , \tag{4.4}$$

$$\hat{N}_2=\sum_{i=1}^{N}\hat{n}_i^{\,2}\ . \tag{4.5}$$

### C. Diffraction intensity

The available experimental information consists of an intensity sample given as a function of the diffraction angle $\theta$. Thus, (2.10) adopts the appearance

$$f_k=\langle\theta_k|f\rangle=A\langle\hat{L}_k\rangle,\quad k=1,\ldots,M \tag{4.6}$$

where the observables $\hat{L}_k$ represent diffraction intensities for the angle $\theta_k$. These can be defined by recourse to elementary semiclassical considerations,[14]

$$\hat{L}_k=\sum_{i=1}^{N}\hat{F}_k^{\,2}(i)+2\sum_{\substack{i,j\\ i<j}}^{N}\hat{F}_k(i)\hat{F}_k(j)\times\cos\left[\hat{\phi}_k(j,i)+\sum_{s=i}^{j-1}\hat{\delta}(s+1,s)\right]\ , \tag{4.7}$$

with

$$\hat{F}_k(i)=F_{1,k}\hat{n}_i^{\,1}+F_{2,k}\hat{n}_i^{\,2}\ , \tag{4.8}$$

$$\hat{\phi}_k(j,i)=\phi_{1,k}\hat{n}_j^{\,1}+\phi_{2,k}\hat{n}_j^{\,2}-(\phi_{1,k}\hat{n}_i^{\,1}+\phi_{2,k}\hat{n}_i^{\,2})\ , \tag{4.9}$$

$$\hat{\delta}_k(s+1,s)=\delta_{11,k}\hat{n}_{s+1}^{\,1}\hat{n}_s^{\,1}+\delta_{22,k}\hat{n}_{s+1}^{\,2}\hat{n}_s^{\,2}+\delta_{21,k}\hat{n}_{s+1}^{\,2}\hat{n}_s^{\,1}+\delta_{12,k}\hat{n}_{s+1}^{\,1}\hat{n}_s^{\,2}\ . \tag{4.10}$$

$F_{1,k}$ and $F_{2,k}$ are the (absolute value) structure factors at the angle $\theta_k$ for the 1 and 2 components, respectively, while $\phi_{1,k}$ and $\phi_{2,k}$ are the corresponding associated phases. Additionally,

$$\delta_{11,k}=4\pi d_1\sin\theta_k/\lambda\ , \tag{4.11}$$

$$\delta_{22,k}=4\pi d_2\sin\theta_k/\lambda\ , \tag{4.12}$$

$$\delta_{12,k}=\delta_{21,k}=2\pi(d_1+d_2)\sin\theta_k/\lambda\ . \tag{4.13}$$

As the operators $\hat{L}_k$ are given by (4.7), the eigenstates of the $\hat{n}_i^{\,1}$ and $\hat{n}_i^{\,2}$ operators are also eigenstates of $\hat{L}_k$. For the Fock-Hilbert space we are dealing with the identity operator which is

$$\hat{I}=\sum_{N=1}^{\infty}\sum_{\nu=1}^{2^N}|\nu\rangle\langle\nu|\ , \tag{4.14}$$

where $\nu$ stands for the set of indices needed in order to satisfactorily label an $N$-layer configuration. Traces are thus evaluated according to [cf. Eq. (4.6)]

$$f_k=A\sum_{N=1}^{\infty}\sum_{\nu=1}^{2^N}\langle\nu|\hat{\rho}\hat{L}_k|\nu\rangle,\quad k=1,\ldots,M\ . \tag{4.15}$$

## V. A REALISTIC SITUATION: ILLITE-MONTMORILLONITE SAMPLES

Montmorillonite is a smectite clay that easily interchanges cations with other substances.[9,15] We shall analyze here results concerning sodic montmorillonite samples in which Na cations compensate for lattice (charge) defects.

Illite clays originate from montmorillonite clays and are not able to interchange cations. They do not appear in pure states. Some degree of interestratification is always present and there exists great interest, both in geology and in industry, to learn details about this interestratification. We shall apply the preceding considerations to this situation and assign label 1 to montmorillonite and label 2 to illite ($d_1=15.4$ Å, $d_2=10$ Å). We take the montmorillonite structure factors from Cole and Lancucki,[16] and those of the illite have been computed from the structure reported by Wever and Pollard.[17]

Figure 2 displays (punctual) x-ray diffraction measurements $\langle \theta_k|f\rangle$ obtained from an illite-montmorillonite polycrystalline sample. These data are already corrected for the Lorentz polarization factor $L_p(\theta)=(1+\cos^2 2\theta)/\sin 2\theta$. In all we deal with 63 pieces of data.

The statistical operator associated to these punctual measurements of the diffraction intensity is

$$\hat{\rho}=\exp\left[-\sum_{i=3}^{M}\lambda_i\hat{O}_i\right]\Big/\mathrm{Tr}\left[\exp\left[-\sum_{i=3}^{M}\lambda_i\hat{O}_i\right]\right], \tag{5.1}$$

with

$$\hat{O}_i=(\langle\theta_i|f\rangle-\langle\theta_1|f\rangle)\hat{L}_2-(\langle\theta_i|f\rangle-\langle\theta_2|f\rangle)\hat{L}_1 -(\langle\theta_2|f\rangle-\langle\theta_1|f\rangle)\hat{L}_i\ , \tag{5.2}$$

a peculiar form that is the consequence of having utilized one of the pieces of data in order to adjust the background constant $B$. The proportionality constant $A$ of the previous Sec. IV C has to be calculated, accordingly, as

$$A=(\langle\theta_2|f\rangle-\langle\theta_1|f\rangle)/(\langle\hat{L}_2\rangle-\langle\hat{L}_1\rangle)\ , \tag{5.3}$$

which in turn yields $B$,

$$B=\langle\theta_2|f\rangle-A\langle\hat{L}_2\rangle\ . \tag{5.4}$$

Due to the properties of the montmorillonite clay,[18,19] it is permissible to truncate our Fock-Hilbert space at $N=10$, which entails dealing with 2046 configurations. In order to attain convergence (within the framework of our 63 pieces of data) with an $\epsilon$ [cf. Eq. (2.22)] determined by the error bars, it was found to be necessary to construct $\hat{\rho}$ out of 10 data points (8 associated Lagrange multipliers). This suffices for an excellent prediction of the remaining 53 data, according to the relations

$$f_j^{(8)}=A\langle\hat{L}_j\rangle+B\ . \tag{5.5}$$

As illustrated by the solid curve in Fig. 2, these predictions converge to the data within the experimental error ($\sim 3\%$). Thus, $\hat{\rho}^{(8)}$ can be confidently employed so as to determine the relative quantities of illite and montmorillonite in the sample

$$\langle\hat{N}_1\rangle=\mathrm{Tr}(\hat{\rho}^{(8)}\hat{N}_1)\ , \tag{5.6}$$

$$\langle\hat{N}_2\rangle=\mathrm{Tr}(\hat{\rho}^{(8)}\hat{N}_2)\ , \tag{5.7}$$

and one finds 30% for montmorillonite and 70% for illite. Of this percentage, a 22% that exists in a pure state is responsible for the small peak that appears at 4.4° (Fig. 2). The percentage of montmorillonite in a pure state is negligible, even if the main maximum in Fig. 2 appears at 3.5°, in the vicinity of that place at which pure-state montmorillonite exhibits a diffraction maximum ($\sim 2.85°$). This is due to the fact that the montmorillonite structure factor at this last angle is larger (by a factor of 4) than the corresponding illite structure factor at its characteristic angle of $\sim 4.4°$.

Finally, we point out that calculations performed with a maximum $N=12$ ($\sim 8000$ configurations) differ from the ones reported above by less than 1%.

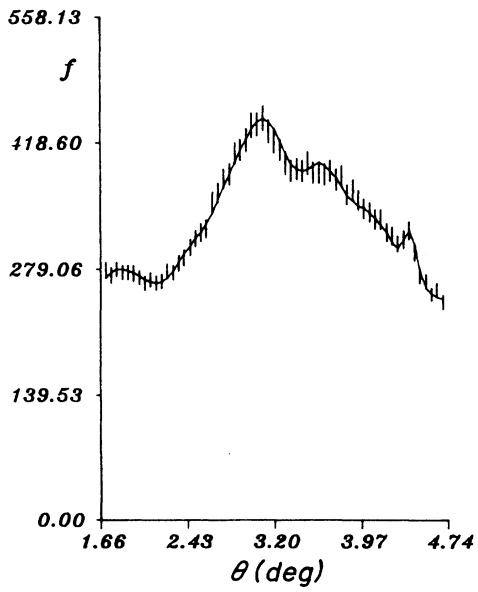


FIG. 2. X-ray diffraction intensities vs diffraction angle $\theta$. The error bars amount to a 3% uncertainty. The solid curve corresponds to the $\hat{\rho}^8$ theoretical prediction. An illite-montmorillonite polycrystalline sample has been employed and the data are corrected for the Lorentz-polarization factor.

## VI. CONCLUSIONS

A very general method for investigating the state distribution in a physical system, on the basis of a detailed analysis of its experimental response to an external probe, has been presented.

The main idea is that of extracting the statistical information carried by the response signal by representing the latter in an appropriate, orthogonal basis. Out of this representation, and by recourse to a finite number of measurements, the state distribution is obtained.

That a finite number of measurements suffices for accomplishing our goal is due to the maximum-entropy principle (MEP). Indeed, we are facing here, from a mathematical standpoint, an inversion problem whose solution is not, in general, unique. The (MEP) provides the crucial criterium that allows one to select just one out

of the multiple solutions the inversion problem may possess.

The quite simple weighting method proposed in Sec. III provides the reader with a rather neat illustration of the ideas introduced in Sec. II.

The example discussed in Secs. IV and V as an application of our formalism provides a rather striking illustration of how powerful the MEP really is. The dimension of the concomitant vector space is quite large indeed, although just a few measurements permit us to construct a rather good solution. A mere 16% of the available data sample is needed in this respect. The remaining 84% has served here just to exhibit the predictive power of the present method.

[1]C. E. Shannon, Proc. IRE **37**, 1 (1949).
[2]C. E. Shannon, Bell. Syst. Tech. J. **27**, 379 (1948); **27**, 623 (1948).
[3]E. T. Jaynes, Phys. Rev. **106**, 620 (1957).
[4]E. T. Jaynes, *The Maximum Entropy Formalism*, edited by R. D. Levine and M. Tribus (MIT, Boston, 1979).
[5]P. A. M. Dirac, *The Principles of Quantum Mechanics* (Oxford University Press, Oxford, 1958).
[6]A. Katz, *Principles of Statistical Mechanics* (Freeman, San Francisco, 1967).
[7]F. Reif, *Fundamentals of Statistical and Thermal Physics* (McGraw-Hill, New York, 1965).
[8]C. Kittel, *Introduction to Solid State Physics* (Wiley, New York, 1976).
[9]G. W. Brindley and G. Brown, *Crystal Structures of Clay Minerals and Their X-Ray Identification* (Mineralogical Society, London, 1980).
[10]C. E. Weaver, Am. Min. **41**, 202 (1956).
[11]E. C. Jonas and T. E. Brown, I. Sedim. Petrol. **29**, 77 (1959).
[12]A. E. Foscolos and H. Kodama, Clays Clay Miner. **22**, 319 (1974).
[13]A. Fetter and J. Walecka, *Quantum Theory of Many Particle Systems* (McGraw-Hill, New York, 1971).
[14]L. Rebollo Neira, Ph.D. thesis, National University La Plata, 1986.
[15]T. Pinavia, Science **220**, 365 (1983).
[16]W. F. Cole and C. J. Lancucki, Acta. Crystallogr. **21**, 836 (1966).
[17]C. E. Weaver and L. D. Pollard, *The Chemistry of Clay Minerals* (Elsevier, Amsterdam, 1973).
[18]D. M. A. Guerin, G. Alvarez, R. Bonetto, A. Plastino, and L. Rebollo Neira, Acta Crystallogr. **A42**, 30 (1986).
[19]G. Alvarez, R. Bonetto, D. Guerin, A. Plastino, and L. Rebollo Neira, Powder Diffraction **2**, 220 (1987).