

## Article

# A Dynamic Probabilistic Model for Heterogeneous Data Fusion: A Pilot Case Study from Computer-Aided Detection of Depression

Federica Vitale <sup>1</sup>, Bruno Carbonaro <sup>1</sup> and Anna Esposito <sup>2,3,\*</sup>

<sup>1</sup> Department of Mathematics and Physics, University of Campania “L. Vanvitelli”, Viale Lincoln 5, 81100 Caserta, Italy; federica.vitale@unicampania.it (F.V.); bruno.carbonaro@unicampania.it (B.C.)

<sup>2</sup> Department of Psychology, Università degli Studi della Campania “L. Vanvitelli”, Viale Ellittico 31, 81100 Caserta, Italy

<sup>3</sup> International Institute for Advanced Scientific Studies (IIASS), 84019 Vietri sul Mare, Italy

\* Correspondence: anna.esposito@unicampania.it or iiass.annaesp@tin.it

**Abstract:** The present paper, in the framework of a search for a computer-aided method to detect depression, deals with experimental data of various types, with their correlation, and with the way relevant information about depression delivered by different sets of data can be fused to build a unique body of knowledge about individuals’ mental states facilitating the diagnosis and its accuracy. To this aim, it suggests the use of a recently introduced «limiting form» of the kinetic-theoretic language, at present widely used to describe complex systems of objects of the most diverse nature. In this connection, the paper mainly aims to show how a wide range of experimental procedures can be described as examples of this «limiting case» and possibly rendered by this description more effective as methods of prediction from experience. In particular, the paper contains a simple, preliminary application of the method to the detection of depression, to show how the consideration of statistical parameters connected with the analysis of speech can modify, at least in a stochastic sense, each diagnosis of depression delivered by the Beck Depression Inventory (BDI-II).

**Keywords:** probability; kinetic theory; mathematical models; speech recognition; statistical parametric speech synthesis



**Citation:** Vitale, F.; Carbonaro, B.; Esposito, A. A Dynamic Probabilistic Model for Heterogeneous Data Fusion: A Pilot Case Study from Computer-Aided Detection of Depression. *Brain Sci.* **2023**, *13*, 1339. <https://doi.org/10.3390/brainsci13091339>

Academic Editor: Mohammad Tanveer

Received: 24 July 2023

Revised: 25 August 2023

Accepted: 4 September 2023

Published: 18 September 2023



**Copyright:** © 2023 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

## 1. Introduction

This paper deals with a special but important and interesting case of the problem of treating experimental data of various types (*heterogeneous* data) together, to provide a detailed description of a subject, or in particular to classify it with a property chosen *a priori*.

The starting point of the work is finding methods and tools for implementing computer-aided devices for the detection of major depressive disorders (MDD), i.e., the automatic classification of any given subject with respect to a set of previously identified possible depressive states. This task requires in general the analysis of a rather large amount of data obtained from experiments of various kinds and settled for investigating whether features describing the collected data can discriminate between healthy and depressed participants. The involved experiments aimed to gather behavioral data (such as speech, handwriting, facial, vocal, and gestural expressions) from healthy/depressed diagnosed subjects by defining behavioral tasks known to be affected by depressive states. The collected data were then appropriately processed to extract features describing changes in the healthy perception of individuals due to depressive states.

Depression is classified as a mood disorder by the Diagnostic and Statistical Manual of Mental Disorders (DSM-5, American Psychiatric Association, 2013). The DSM-5 report as its symptoms the impairment impaired of emotional expressions, cognitive functions, social behavior, social relationships, and physical functioning (changes in appetite or in

weight, headaches, pain). However, not all depressed individuals meet these criteria. Some individuals may show irritable rather than depressed mood and may not be classified as depressed. “The conditions and underlying physiologies associated with the experience of depression may vary widely, and no core composition can be assumed” ([1], p. 1). Given the complexity of the accompanying symptoms, it was shown that the automatic analysis of behavioral data can provide quantitative measures for describing changes affecting healthy/depressed individuals (see [2] for depression symptoms and [3] and reference therein for automatic behavioral analyses) calling for the identification of a multi-dimensional mathematical model serving for the implementation of automated and cost-effective technological systems devoted to the early detection of depressive states.

The mathematical model proposed in this paper has been applied to speech data. More precisely, in this case, for any subject, the experimental procedure involves the collection of speech data obtained from the reading of a narrative text and the free description of daily activities performed during a week, and the answers provided by the participants to data collection to the questions of BDI-II [4]. Further data can be collected on this line of conduct, other than speech, including facial expressions, handwriting and drawing, body movements, hands’ gestures, and any other activity useful for behavior analysis ([5–13], among others). Though all the outcomes of these experiments can be (and actually are) given a quantitative expression, by means of suitably defined measures or scores, nevertheless, this raises a new problem, regarding the comparison of heterogeneous measures.

To solve this problem, a good method seems to be shifting the attention from the measures as such to their joint probabilities and, as a consequence, to their mutual conditional probabilities. This goal has been achieved by the use of a recently introduced «limiting form» of the kinetic-theoretic language [14], a *scheme* of models based on the use of a generalization of Boltzmann equation governing gas dynamics [15,16], which in the last fifty years has been proposed to describe the evolution of many-particle systems in which the interactions between elements go far beyond purely mechanical ones, and can produce unexpected and not deterministically predictable behaviors, both mechanical and non-mechanical. The elements of the system are themselves no longer just particles in the mechanical sense, but individuals of the most varied nature: cells [17–21], cars [22,23], insects (dynamics of swarms) [24], and finally, human beings and possibly care structures [21,25–29].

The key idea the whole scheme relies upon is that the results of pairwise interactions between the «particles» of the system cannot be predicted with certainty, but they still produce a *deterministic* evolution of the (relative) frequency (or, in appropriate contexts, probability) distribution on the set (space) of possible states of particles [15,30].

One of the techniques often employed in the analysis of the dependence of this distribution on the interactions consists in subdividing the system into *homogeneous* subsystems (i.e., containing the same kind of objects), which can also have different spaces of states. No models, however, have been outlined in which at least one functional subsystem consisted of abstract objects, e.g., «procedures» or «experiments» or «diagnostic methods».

In addition, as is quite natural, in all models based on the Kinetic Theory, the time variable is continuous, and as a consequence the evolution of the system turns out to be governed by a system of differential equations, when the space  $\mathcal{D}$  of all possible states of the «particles» is discrete (see Section 2) or of integro-differential equations, when  $\mathcal{D}$  is continuous, and described by either functions of time or functions of state variables and time. In this paper, a formulation of the model introduced in [14] is used, in which the complex system under examination contains such abstract objects as «random variables» and the time variable is allowed to be *discrete*.

The proposed model is applied to speech data, collected from depressed (diagnosed from medical doctors) and healthy subjects. The research assumes that in the acoustic and linguistic content of people’s speeches it is possible to capture information regarding their psychological state, according to the narrative psychological content analysis framework [31].

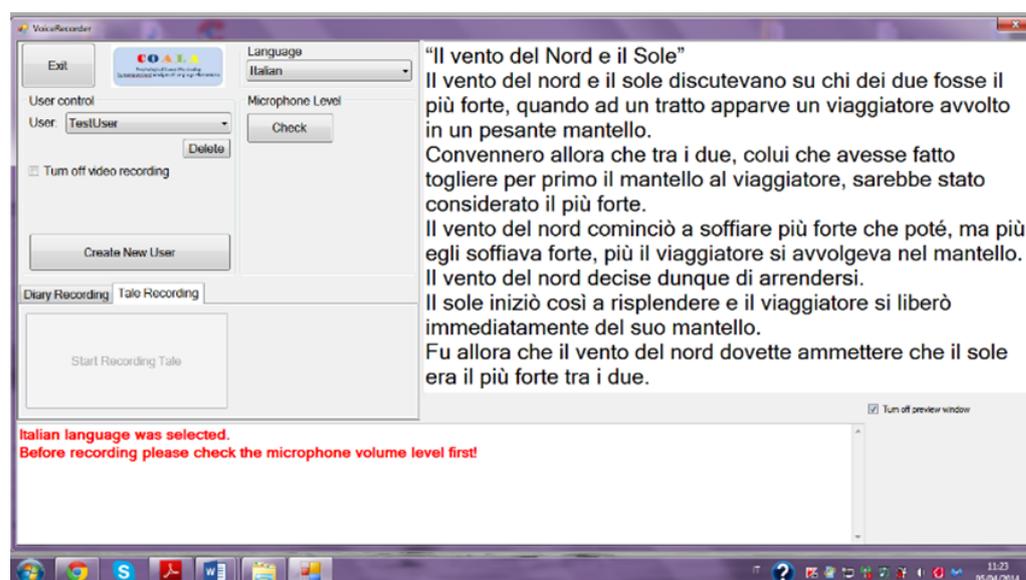
There are several studies showing that there are vocal changes in the speech of subjects with depression. These changes concern, mainly, (a) the fundamental frequency  $f_0$ , which seems to reduce its frequency range and average values in the acute stages of the pathology; (b) the speaking rate which seems to increase as the pathology improves, and finally, that this improvement is related to c) a decrease in the number and duration of speech pauses (for a review, see [3,32]).

The data utilized for the following research included 24 subjects divided into an experimental and control group. The experimental group, included 12 subjects (nine females) with severe (four subjects), moderate (five) and mild (three) depression, aged from 27 to 60 years (mean age = 41.07 years; standard deviation = 12.5). The control group was selected to match the experimental groups for age and gender.

The diagnostic criteria of the Beck Depression Inventory, Second Edition (BDI II, [4]), an instrument for self-assessment of depression severity in already diagnosed patients and for detecting the risk of depression in the normal population, were used to assess depression. The BDI consists of 21 sets of statements to be rated on a four-point scale. A score of >16 indicates the presence of depression in the patient.

Subjects with depression were recruited in Italy, from the mental health department of the ASL of Caserta, the mental hygiene institute of the ASL of Santamaria Capua Vetere, the Psychological Listening Center of Aversa, and a private medical practice.

The task performed by the subjects, divided into two parts, consisted of a reading of the Aesop's tale "the North Wind and the Sun" visualized on a computer screen (see Figure 1).



**Figure 1.** The Aesop's tale "The North Wind and the Sun" as it was presented to the participants involved in the data collection.

In the second part, participants were required to produce a spontaneous narrative (diary), describing salient events which happened during the previous week, highs and lows with friends, relatives, co-workers and also, in many cases, opinions or criticisms of the Italian political situation.

During the reading and diary sections, the subjects were audio and/or video recorded upon consent. The duration of the recordings varied. The diary ranges from a minimum of 2 min to a maximum of 5/6 min; while the reading takes about 50 s. The total time of the task, from meeting with the examiner to the end of the recording, was around 15 min.

The recordings were made using a clip-on microphone (Audio-Technica ATR3350), with an external USB sound card. Speech was sampled at 16 kHz and quantized at 16 bits.

This software/hardware was developed inside the European Space Agency (ESA 2012) funded project (n. HSO-US 2012-108) “Psychological Status Monitoring by Computerised Analysis of Language phenomena (COALA)” of which one of the authors (Professor Anna Esposito) was the coordinator. It utilized standard software/hardware equipped with a sound card to sample the speech data. It can run on any operating system. It was installed on Window 7 for collecting the COALA data. In the following years it has been installed on the operating systems Windows 8, 9 and 10 and can also be installed on Windows 11. The data exploited in this study were collected between years 2016 and 2017, after obtaining the permission of the ethical committee of the department of psychology of the Università della Campania “Luigi Vanvitelli”. The date displayed in Figure 1 and the time in which the data were collected do not affect the originality and innovation of the results presented in this study, which seeks to propose a mathematical model to fuse heterogeneous data.

The recording took place in an isolated and soundproofed room, granted by the health center where the data collection was taking place.

Prior to the recording, subjects signed an informed consent form for data processing and filled out the BDI II [33], while the examiner prepared the instrumentation consisting of a personal computer and recording software.

The research had received the approval from the ethical committee of the department of Psychology, University of Campania “Luigi Vanvitelli” with the protocol code 09/2016.

The model suggests a way to give a sharp definition of states and a precise method to place each individual in a state and to confirm or modify the placement after each «interaction». This point will be discussed in details in the following Sections. More precisely, in Sections 2 and 3, the formulation of the «discretized» model is presented, in Sections 4–6, the language to the detection of depression using the results reported in [34] is applied, and in Section 7 some possible research perspectives are outlined.

## 2. The Standard Model

In order to provide the reader at least some hints about the origin of the methods that will be used in the next Sections to analyze the correlation between different features possibly describing depression, the current one is devoted to recall the main features of the kinetic-theoretic language in its most general form.

Our starting point is a set  $S$  of a very large number of objects of any nature (cells, living individuals, and also—as was shown in [21] and will be discussed in more details in the next Section—care structures, e.g., hospitals and systems of care tools and experimental devices) called *individuals* or (*active*) *particles*. In most cases of interest,  $S$  is assumed to be the joining of a family  $\{S_1, S_2, \dots, S_k\}$  of subsystems, called its *functional subsystems*. The introduction of such subsystems finds its natural application in biology, for instance to describe the fight between tumor cells and immune system, or the competition between different species to model Darwinian selection, or in social sciences, to model interactions between social classes, or in economics, to model the fluxes of wealth. Accordingly, for each subsystem  $S_i$ , the *state* of each individual in  $S_i$  is defined according to the context in which the present description should be developed, and can be expressed as a scalar variable  $u_i$  or a vector variable  $\mathbf{u}_i \equiv (u_i^1, u_i^2, \dots, u_i^m)$  which will be called *state variable*. This variable may describe any property we can find suitable for our research according to the context (biological, economic, social, etc.) in which it is performed. According to the context, the state variable may be assumed to take its values in a *discrete* domain or in a *continuous* domain: in the former case, the domain can be a finite or countable subset of  $\mathbb{Z}$  or  $\mathbb{Z}^m$ , while in the latter case it can be a bounded or unbounded real interval or a bounded or unbounded domain of  $\mathbb{R}^m$ . In both cases, the domain  $D_{u_i}$  (or  $D_{\mathbf{u}_i}$ ) of the state variable is called the *state space* of subsystem  $S_i$  (in some cases, the same variable can be used to identify the state of the members of all the functional subsystems, but in most cases different state variables must be used for different subsystems).

As in Boltzmann’s Kinetic Theory of Gases, the mathematical framework developed in [21] is statistical, i.e., one is interested to describe the evolution in time of the system as a

whole rather than the evolution of the states of single particles. More precisely, it must be acknowledged that any precise description of the state of each particle of  $S$  must be given up, for both theoretical and practical reasons [16,35], so one cannot but decide to study the probability (expressed experimentally in terms of relative frequency) distribution over the state space at each instant. This means that the state variable is conceived as a *random variable* at each instant and the aim of the study is the forecast of its probability density function at any time.

In this Section, with no loss of generality, and only for the sake of simplicity and in view of the application to be discussed in the sequel, it will be described here in some detail only the case of a system  $S$  consisting in two different subsystems  $S_1$  and  $S_2$  such that the state of individuals in each  $S_h$  is described by a different *discrete* scalar variable  $u_h$ . We set

$$D_h = \{u_{h,1}, u_{h,2}, \dots, u_{h,m_h}\} \quad (h = 1, 2).$$

For any  $t$  in a time-interval  $\mathbb{I} \subseteq \mathbb{R}$ , the state of system  $S_h$  at time  $t$  will be identified by a probability distribution (also called *state vector*)  $\mathbf{f}_t^{(h)} \equiv (f_t^{(h)}(u_{h,1}), f_t^{(h)}(u_{h,2}), \dots, f_t^{(h)}(u_{h,m_h}))$  on  $D_h$ , and, for any  $k \in \{1, 2, \dots, m_h\}$ , one can define the function  $f_k^{(h)} : t \in \mathbb{I} \rightarrow f_k^{(h)}(t) \equiv f_t^{(h)}(u_{h,k}) \in [0, 1]$ . According to this definition, one has

$$\sum_{k=1}^{m_h} f_k^{(h)}(t) = \sum_{k=1}^{m_h} f_t^{(h)}(u_{h,k}) = 1 \quad (h = 1, 2).$$

Now, if the evolution of the system  $S$  is envisaged as a time-continuous stochastic process, then the time derivative of each probability function  $f_k^{(h)}$  is expressed, according to the law of alternatives, in terms of transition probabilities. More precisely,

- for any  $(r, s, j) \in \{1, 2, \dots, m_h\}^3$ , the symbol  $F_{sj}^r(h) \equiv F_h(u_{h,s}, u_{h,j}; u_{h,r})$  will denote the probability that a particle of  $S_h$  falls from the state  $u_{h,s}$  to the state  $u_{h,r}$  after an interaction with another particle of  $S_h$  which is in the state  $u_{h,j}$ : accordingly,

$$\sum_{r=1}^{m_h} F_{sj}^r(h) = 1, \quad \forall (s, j) \in \{1, 2, \dots, m_h\}^2 \quad (h = 1, 2).$$

- for any  $(r, s) \in \{1, 2, \dots, m_h\}^2$ , and for any  $j \in \{1, 2, \dots, m_k\}$  (with  $k \neq h$ ), the symbol  $\Phi_{sj}^r(h, k) \equiv \Phi_{hk}(u_{h,r}, u_{k,j}; u_{h,s})$  will denote the probability that a particle of  $S_h$  falls from the state  $u_{h,s}$  to the state  $u_{h,r}$  after an interaction with a particle of  $S_k$  which is in the state  $u_{k,j}$ : accordingly,

$$\sum_{r=1}^{m_h} \Phi_{sj}^r(h, k) = 1, \quad \forall s \in \{1, 2, \dots, m_h\}, \forall j \in \{1, 2, \dots, m_k\},$$

$$(h, k = 1, 2; k \neq h).$$

So, the law of alternatives yields the following system of differential equations:

$$\begin{aligned} \frac{df_h^{(1)}}{dt}(t) &= \sum_{i,j}^{1..m_1} \tau_{ij}^1 F_{ij}^h(1) f_i^{(1)}(t) f_j^{(1)}(t) - \sum_{i,j}^{1..m_1} \tau_{ij}^1 F_{ij}^i(1) f_h^{(1)}(t) f_j^{(1)}(t) + \\ &+ \sum_{i=1}^{m_1} \sum_{j=1}^{m_2} \left[ \eta_{ij}^{12} \Phi_{ij}^h(1, 2) f_i^{(1)}(t) f_j^{(2)}(t) - \eta_{ij}^{12} \Phi_{ij}^i(1, 2) f_h^{(1)}(t) f_j^{(2)}(t) \right], \quad (h = 1, \dots, m_1) \\ \frac{df_h^{(2)}}{dt}(t) &= \sum_{i,j}^{1..m_2} \tau_{ij}^2 F_{ij}^h(2) f_i^{(2)}(t) f_j^{(2)}(t) - \sum_{i,j}^{1..m_2} \tau_{ij}^2 F_{ij}^i(2) f_h^{(2)}(t) f_j^{(2)}(t) + \\ &+ \sum_{i=1}^{m_2} \sum_{j=1}^{m_1} \left[ \eta_{ij}^{21} \Phi_{ij}^h(2, 1) f_i^{(2)}(t) f_j^{(1)}(t) - \eta_{ij}^{21} \Phi_{ij}^i(2, 1) f_h^{(2)}(t) f_j^{(1)}(t) \right], \quad (h = 1, \dots, m_2) \end{aligned} \quad (1)$$

where, for any  $k \in \{1, 2\}$  and any  $(r, s) \in \{1, 2, \dots, m_k\}^2$ ,  $\tau_{rs}^k \equiv \tau(u_{k,r}, u_{k,s})$  is the so-called *encounter rate* of particles of  $S_k$  in the states  $u_{k,r}$  and  $u_{k,s}$ . It is the number of pairwise interactions between particles of  $S_k$ , that are in the states  $u_{k,r}$  and  $u_{k,s}$ , per time unit, and for any sufficiently small  $\Delta t$  the product  $\tau_{rs}^k \Delta t$  is the probability that one such interaction occur in the time interval  $\Delta t$  between individuals belonging to  $S_k$  with states  $u_{k,r}$  and  $u_{k,s}$  respectively. As a consequence,  $\tau_{rs}^k \Delta t$  is a *conditional probability* and  $\tau_{rs}^k$  is the ratio of a conditional probability to time. Its presence in the first two terms at the right-hand side of each equation of system (1) is due to the fact that these terms express two basic properties:

1. the increase in probability of state  $u_{k,h}$  is the probability that some «candidate» particles of  $S_k$  in the state  $u_{k,r}$  interact with some «field» particles of  $S_k$  in the state  $u_{k,s}$ : they fall in the state  $u_{k,h}$  with a positive probability just in consequence of the interaction;
2. the decrease in probability of state  $u_{k,h}$  is the probability that some «test» particles of  $S_k$  in the state  $u_{k,h}$  interact with some «field» particles of  $S_k$  in the state  $u_{k,s}$ ; they leave the state  $u_{k,h}$  with a positive probability only in consequence of the interaction.

The coefficients  $\eta_{rs}^{\ell k}$  (for  $r \in \{1, 2, \dots, m_\ell\}$  and  $s \in \{1, 2, \dots, m_k\}$ ) in the other terms at the right-hand side of Equation (1) have the same meaning: for any  $(\ell, k) \in \{1, 2\}^2$ ,  $\eta_{rs}^{\ell k}$  is the *encounter rate* between any particle of  $S_\ell$  which is in the state  $r$  and any particle of  $S_k$  which is in the state  $s$ .

Notice that, as we shall do from now on to the end of the paper, the above system of equations has been identified by a single numerical label: from now on, each of the equations that compose a system will be distinguished by a subscript corresponding to its position in it: for instance, the third equation of system (X) will be indicated as  $(X)_3$ .

For the sake of completeness, it will be suitable to recall that it is customary in the literature to write system (1) in the form

$$\frac{df_h^{(\ell)}}{dt}(t) = J_h^{(\ell)}[f^{(1)}, f^{(2)}](t) \quad (h = 1, 2, \dots, m_\ell; \ell = 1, 2), \tag{2}$$

where

$$\begin{aligned} J_h^{(\ell)}[f^{(1)}, f^{(2)}](t) &= \\ &= G_h^{(\ell)}[f^{(1)}, f^{(2)}](t) - L_h^{(\ell)}[f^{(1)}, f^{(2)}](t) \quad (\ell = 1, 2), \end{aligned} \tag{3}$$

and

$$\begin{aligned} G_h^{(\ell)}[f^{(1)}, f^{(2)}](t) &= \sum_{i,j}^{1\dots m_\ell} \tau_{ij}^\ell F_{ij}^h(\ell) f_i^{(\ell)}(t) f_j^{(\ell)}(t) + \\ &+ \sum_{i=1}^{m_\ell} \sum_{j=1}^{m_k} \eta_{ij}^{\ell k} \Phi_{ij}^h(\ell, k) f_i^{(\ell)}(t) f_j^{(k)}(t), \quad (\ell = 1, 2; k \neq \ell) \\ L_h^{(\ell)}[f^{(1)}, f^{(2)}](t) &= f_h^{(\ell)}(t) \left[ \sum_{i,j}^{1\dots m_\ell} \tau_{ij}^\ell F_{ij}^i(\ell) f_j^{(\ell)}(t) + \right. \\ &\left. + \sum_{i=1}^{m_\ell} \sum_{j=1}^{m_k} \eta_{ij}^{\ell k} \Phi_{ij}^i(\ell, k) f_j^{(k)}(t) \right] \quad (\ell = 1, 2; k \neq \ell) \end{aligned} \tag{4}$$

For obvious reasons, the term  $G_h^{(\ell)}[f^{(1)}, f^{(2)}](t)$  is called *gain term*, while the term  $L_h^{(\ell)}[f^{(1)}, f^{(2)}](t)$  is called *loss term*. Note that system (2) consists of  $m = m_1 + m_2$  equations.

### 3. Formulation and Interpretation of the Special Model

In this Section, the general model outlined in the previous one will be strongly modified under three aspects. A system  $S$  is considered, split into two subsystems  $S_1$  and  $S_2$ , but

1. subsystem  $S_1$ , which from now on will be also referred to as «the system of subjects», consists of *only one* member, so that interactions between subjects are excluded as meaningless;
2. subsystem  $S_2$ , which from now on will be also referred to as «the system of experiments», consists of  $m$  members (also called the «questions about the subject»), but again the interactions between these members are excluded;
3. the state space of subjects will be denoted by  $D \equiv \{u_0, u_1, \dots, u_s\}$ ;
4. to *each* experiment in  $S_2$  a *different* state space  $D_i$  is associated, which is a set of positive integers (the outcomes of the experiment): we denote by  $y_i$  the state variable in  $D_i$  so that  $D_i = \{y_{i1}, y_{i2}, \dots, y_{im_i}\}$ ;
5. interactions between the subject and experimental devices *do not* modify the states of the latter;
6. the time variable is assumed to be *discrete*, i.e., the process described here is a *stepwise* process.

**Remark 1.** In virtue of condition 1,  $\tau_{ij}^1 = 0$  for any couple  $(u_i, u_j)$  of states in  $D$ .

**Remark 2.** It is necessary to point out, for the sake of the readability of the text, that some symbols introduced in Section 2 and appearing in system (1) as well as in definitions (4) must be modified because of condition 4. Precisely, the interaction rates  $\tau_{hj}^2$  should be written  $\tau_{h,k;j,r}^2$ , where  $h$  and  $j$  are indexes identifying the state spaces  $D_h$  and  $D_j$  (hence the experiments) involved in the interaction, and  $k$  and  $r$  are the states (in  $D_h$  and  $D_j$ , respectively) of interacting «questions». The same is true for the transition probabilities  $F_{hj}^i(2)$  (to be written  $F_{r,h;s,j}^i(2)$ ). However, of course, in virtue of condition 2, this remark can be completely disregarded, since all the terms containing the considered interaction rates and transition probabilities vanish.

**Remark 3.** The situation is quite different for the «mixed» interaction rates  $\eta_{ij}^{12}$  and  $\eta_{ij}^{21}$  and transition probabilities  $\Phi_{ij}^h(1,2)$  and  $\Phi_{ij}^h(2,1)$ , which should be written as  $\eta_{i,k;j}^{12}$ ,  $\eta_{k;i;j}^{21}$ ,  $\Phi_{i,k;j}^h(1,2)$  and  $\Phi_{k;i;j}^{h,k}(2,1)$ , respectively, where  $k$  identifies in all cases the state space  $D_k$  of the experiment involved in the interaction. Notice however that, in virtue of condition 5,  $\Phi_{k;i;j}^{h,k}(2,1) = \delta_i^h$  for any  $k \in \{1, 2, \dots, m\}$  and any  $j \in \{0, 1, \dots, s\}$ .

**Remark 4.** Analogously,  $f_h^{(2)}$  must be written  $f_{kh}^{(2)}$ . This last symbol expresses the probability that the  $k$ -th experiment is in the  $h$ -th state  $y_{kh}$  of  $D_k$ .

**Remark 5.** Finally, condition 6, according to which the evolution of the system will be described as a stepwise process allows us to consider a sequence  $\{t_h\}_{h \in \mathbb{N}}$  and to replace the time derivatives at the left-hand side of Equation (1) by finite differences of the form  $f_h^{(1)}(t_{n+1}) - f_h^{(1)}(t_n)$  and  $f_{k,h}^{(2)}(t_{n+1}) - f_{k,h}^{(2)}(t_n)$  ( $n \in \mathbb{N}$ ), respectively. In particular, we shall take  $n = 0$ .

In view of the above remarks, system (1) simply becomes

$$\begin{aligned}
 f_h^{(1)}(t_1) - f_h^{(1)}(t_0) &= \sum_{i=1}^s \sum_{k=1}^m \left[ \sum_{j=1}^{m_k} \eta_{i,k;j}^{12} \Phi_{i,k;j}^h(1,2) f_i^{(1)}(t_0) f_{kj}^{(2)}(t_0) + \right. \\
 &\quad \left. - \eta_{h,k;j}^{12} \Phi_{h,k;j}^i(1,2) f_h^{(1)}(t_0) f_{kj}^{(2)}(t_0) \right], \quad (h = 1, \dots, s), \\
 f_{kh}^{(2)}(t_1) - f_{kh}^{(2)}(t_0) &= 0, \quad (k = 1, \dots, m; h = 0, 1, \dots, m_k),
 \end{aligned} \tag{5}$$

to which the initial conditions

$$\begin{aligned}
 f_h^{(1)}(t_0) &= f_{h0}^{(1)} & (h = 1, \dots, s). \\
 f_{kh}^{(2)}(t_0) &= f_{kh0}^{(2)} & (k = 1, \dots, m; h = 0, 1, \dots, m_k)
 \end{aligned} \tag{6}$$

must be associated.

An interpretation as complete as possible of the special model obtained starting from conditions 1–6 will be outlined, not only in order to explain the meaning of these hypotheses, but also to turn the attention of the reader to the significant application of the system obtained above to the detection of depression. In this connection, what it will be necessary to stress is that system (5) is meant to describe the results of interactions between *information* rather than between objects. More precisely, the values  $u_i$  of the variable in  $D$  should be interpreted as labels that *define*, in the appropriate context, possible states of the subject; then, the interaction between the subject and the  $i$ -th experience on it places this latter into a *state*, expressed by a real value  $y_{ih}$ , that is the outcome of a measurement of a physical parameter. Now, in virtue of the information carried by the value  $y_{ih}$ , the  $i$ -th experience *interacts* with the initial definition, in the sense that it can either confirm the initial state (by increasing its probability) or raise doubts about it (by lowering its probability). Accordingly, the terms  $f_i^{(1)}(t_0)$  and  $f_{kj}^{(2)}(t_0)$  in Equation (5) must be interpreted as the *absolute* probability that the *definition* of the state of the subject before the experience is the value  $u_i$  and the *absolute* probability that the outcome of the  $h$ -th experience is the state  $y_{kj}$ , respectively; furthermore, the term  $\Phi_{i;kj}^h(1,2)$  is the probability that definition  $u_i$ , after «interaction» with information  $y_{kj}$  becomes definition  $u_h$ .

This outline of the methodological meaning of system (5) will be now explained in somehow more precise terms by referring to a special experimental problem connected to the detection of depression. In the next sections, by means of an example based on actual numerical data, it will be shown how the values of  $f_i^{(1)}(t_0)$ ,  $f_{kj}^{(2)}(t_0)$  and  $\Phi_{i;kj}^h(1,2)$  should be assigned according to a suitable preliminary statistical analysis.

Therefore, the problem which gave rise to the present analysis will be now described, namely, the search for automated procedures to diagnose the possible depressive state of a human subject. As is well known to psychologists, depression is nowadays mainly diagnosed by means of questionnaires. Among them, of particular interest seems to be the one devised by Aaron Beck called BDI (Beck Depression Inventory). This is a list of 21 questions, each admitting four possible answers, with different scores varying from 0 to 3. According to the total score obtained by his/her answers, the subject is placed in a “class of depression”, namely, with a score from 1 to 10, the “normal class” (which will be denoted here as the *state* 0), with a score from 11 to 16, the “mild mood disturbance class” (the *state* 1), with a score from 17 to 20, the “borderline clinical depression class” (the *state* 2), with a score from 21 to 30, the “moderate depression class” (the *state* 3), with a score from 31 to 40, the “severe depression class” (the *state* 4), and finally, with a score over 40, the “extreme depression class” (the *state* 5).

Now, it is well known that any psychological state produces (and is expressed by) particular behaviors, which can be taken as “symptoms” of that state. These behaviors can be in turn described by the values of physical parameters, for instance voice fundamental frequencies (pitches) in a sequence of sufficiently small time intervals, or the number and the length of empty (or silent) pauses, and the speed of speech and the choice of words when reading a text or simply talking about everyday life (see [36], also for more complete references). Thus, the problem of establishing in which cases the measurements of these and other parameters confirm the results of BDI arises spontaneously, and in which cases they modify the conclusions drawn from the questionnaire. The final diagnosis should result from a suitable “superposition” of all these results. More precisely, one takes a subject initially placed by his (or her) BDI score in a state  $k$ , then tries to determine the value of one of the physical parameters that could reveal depression or a normal state: for instance, he submits him or her to a check of the pitches of his or her voice while reading a text, or of the number and the length of his or her empty pauses when telling his or her experiences during the last week. As laid out above, the parameters defined by these tests, that are random variables, will be denoted by the symbols  $Y_1, Y_2, \dots, Y_m$ . If the  $i$ -th test gives the outcome  $Y_i = y_{ih}$  (i.e., the result of the *physical* interaction between the subject and the experimental device is  $y_{ih}$ ), provided one has defined for any  $j \in D \equiv \{0, 1, \dots, 5\}$  the probability

$\Phi_{k;ih}^j(1, 2) \equiv \Phi_{kh}^j$  that this result “moves” the subject from the state  $k$  (initial diagnosis) to the state  $j$ , system (5) allows us to find a *final* probability distribution on  $D$  (in this connection, notice that the initial probability distribution vector  $(f_{h0}^{(1)})_{1 \leq h \leq 5} \equiv (f_{h0})_{1 \leq h \leq 5}$  is assigned according to our confidence in the reliability of the BDI questionnaire: a good starting point could be to set  $f_{k0} = 1$  and  $f_{h0} = 0$  for any  $h \neq k$ ). Thus, the key step to make system (5) effective is to determine the transition probabilities  $\Phi_{kh}^j \equiv \Phi_{k;ih}^j(1, 2)$  ( $j, k \in D$ ). The way in which this can be performed will be discussed in detail in Section 4. At this moment, it will be appropriate to conclude the present Section with some important remarks.

First of all, it should be stressed that the first step to compute the transition probabilities  $\Phi_{kh}^j \equiv \Phi_{k;ih}^j(1, 2)$  ( $j, k \in D$ ) is to establish the range  $R(Y)$  of the values of  $Y$  that *confirm* the diagnosis delivered by BDI score, or, what is the same, to *synchronize* the responses of BDI and of the additional test. In other words, for any state  $k$ , one must know the class of values of the additional parameter  $Y$  that seems to “characterize”, in a suitable sense of the word, the state  $k$ . This will therefore be performed preliminarily in the next Section.

Thus, the procedure will enable the researchers to find the *final* probability distribution on the set of possible states of the subject, after *all* the tests devised to determine the level of his or her mental health.

**Remark 6.** *It should be carefully noted that the use of system (5) can be replaced by the repeated use of system*

$$\begin{aligned} f_h^{(1)}(t_k) - f_h^{(1)}(t_{k-1}) &= \sum_{i=1}^s \left[ \sum_{j=1}^{m_k} \eta_{i;kj}^{12} \Phi_{i;kj}^h(1, 2) f_i^{(1)}(t_{k-1}) f_{kj}^{(2)}(t_{k-1}) + \right. \\ &\quad \left. - \eta_{h;kj}^{12} \Phi_{h;kj}^i(1, 2) f_h^{(1)}(t_{k-1}) f_{kj}^{(2)}(t_{k-1}) \right], \quad (h = 1, \dots, s), \\ f_{kh}^{(2)}(t_k) - f_{kh}^{(2)}(t_{k-1}) &= 0, \quad (h = 0, 1, \dots, m_k), \quad (k = 1, \dots, m). \end{aligned} \quad (7)$$

More precisely, for any  $k \in \{1, \dots, m\}$ , the system for the parameter  $Y_k$  can be written, where the values  $f_h(t_{k-1})$  are delivered by the same system applied to the random variable  $Y_{k-1}$ . By the way, this will be the procedure which will be adopted in Section 6 in our application of the method to the detection of depression in human subjects.

**Remark 7.** *It should also be noted that the method outlined in Remark 6 and the use of system (5) are not equivalent. The former is applied and illustrated in Section 6, while the latter roughly produces a «superposition» of the results obtained in the first step of the procedure shown therein.*

**Remark 8.** *Analogously, consider a sequence  $\{Y_k\}_{1 \leq k \leq m}$  of parameters. One cannot in general expect that, if Equation (7) is applied according to the plot  $Y_1 \rightarrow Y_2 \rightarrow \dots \rightarrow Y_m$  or according to a plot  $Y_{i_1} \rightarrow Y_{i_2} \rightarrow \dots \rightarrow Y_{i_m}$  (where  $\{i_1, i_2, \dots, i_m\}$  is any permutation of the set  $\{1, 2, \dots, m\}$ ), the same result will be obtained, i.e., the probability distributions  $\{f_h^{(1)}(t_m)\}_{1 \leq h \leq s}$  and  $\{f_h^{(1)}(t_{i_m})\}_{1 \leq h \leq s}$  will be equal. As a matter of fact, as we shall see in Section 6, to obtain a single reliable distribution from the set of experiments, one needs to mediate between the different distributions obtained by following the different plots.*

#### 4. Detecting Symptoms of Depression: General Formulation

In this Section, the model described in purely theoretical terms in the previous one will be applied to a set of data, in order to exhibit a concrete example of the results of the outlined method. Three physical (behavioral) parameters will be introduced: (a) the *total variation*  $T$  of pitch over the speech while reading a prescribed text; (b) the *average variation*  $A$  of pitch over the same reading, and (c) the percentage  $O$  of *inversions*—also called *oscillations*—of the sign of two subsequent differences of pitches over the whole reading, and system (7) will be applied to them.

Accordingly, first of all, the problem of «synchronizing» the values of these parameters with BDI labels must be solved. From a theoretical viewpoint, this must be performed by the following procedure. First, a sufficiently large sample  $S$  of both healthy and depressed subjects is considered. Setting  $n = |S|$  and denoting by  $Y$  any one of the parameters  $T$ ,  $A$  and  $O$ , we also set

1.  $R_Y = \{Y(x) \mid x \in S\}$  and  $e'(Y) = \min R(Y)$ ,  $e''(Y) = \max R(Y)$ ;
2.  $C_h = \{x \in S \mid \text{the BDI state of } x \text{ is } h\}$ ;
3.  $|C_h| = n_h$  so that  $n = \sum_{h=0}^5 n_h$ ;
4.  $q_h \equiv n_h/n = \mathbf{P}(C_h)$  the fraction of the subjects in  $S$  who are placed by BDI in the class  $C_h$ ;
5.  $R_h(Y) = \{Y(x) \mid x \in C_h\}$  and  $e'_h(Y) = \min R_h(Y)$ ,  $e''_h(Y) = \max R_h(Y)$ ;
6.  $\mu_h(Y) = \frac{1}{n_h} \sum_{x \in C_h} Y(x)$ .
7.  $\sigma_h(Y) = \sqrt{\frac{1}{n_h} \sum_{x \in C_h} [Y(x) - \mu_h(Y)]^2}$ .

In general, one expects that  $h \neq k$  does not imply  $[e'_h(Y), e''_h(Y)] \cap [e'_k(Y), e''_k(Y)] = \emptyset$ . So, one has to determine, for any  $h \in \{0, 1, 2, 3, 4, 5\}$ , a set  $I_h(Y)$  in such a way that

1. for any  $h$ ,  $I_h(Y) \subseteq R_h(Y)$  is either a real interval or the join of a finite number of disjoint intervals;
2.  $\{I_h(Y)\}_{0 \leq h \leq 5}$  is a partition of  $R_Y$ ;
3. when a subject is assigned to class  $C_h$  by BDI, then the corresponding value of  $Y$  has a high probability to be in  $I_h(Y)$  (and a low probability to be in any interval  $I_k(Y)$  for  $k \neq h$ ).

In the present context, the three conditions above are assumed to be satisfied, and any description of the procedure one should follow to meet them is omitted. The reader eager for details can usefully consult any university textbook about statistics and data analysis. In the next Section, however, an example of such a procedure will be shown in a simplified case.

Accordingly, a family of 6 sets  $I_h \equiv I_h(Y)$  ( $h = 0, 1, \dots, 5$ ), some of which will be possibly split in two intervals, is assumed to be given. We now denote by  $n_{hk}$  the number of the subjects in  $C_h$  for which the value of the parameter  $Y$  belongs to  $I_k$ , so that

$$n_h = \sum_{k=0}^5 n_{hk},$$

and, in accordance with condition 3,

$$n_{hh} > \sum_{k \neq h} n_{hk}.$$

Obviously,

$$p_{hk} \equiv \mathbf{P}(I_k \mid C_h) = \frac{n_{hk}}{n_h}$$

and, setting  $p_h \equiv p_{hh}$ ,

$$\mathbf{P}(I_h^c \mid C_h) = \sum_{k \neq h} p_{hk}(Y) = 1 - p_h.$$

For the sake of completeness, it must be recalled that

$$\mathbf{P}(C_h \mid I_k) = \frac{p_{hk}q_h}{\sum_{i=0}^5 p_{ik}q_i} = \frac{n_{hk}}{\sum_{i=0}^5 n_{ik}}. \quad (8)$$

It must be stressed that (a) whenever an individual is taken randomly from the considered sample without checking its BDI class, given the value of  $Y$  associated to it, the above relation gives the probability that this individual belongs to any BDI class, and (b) the left-hand side of relation (8) only serves as a reminder of how one should compute

the required probability from the sample data: if one decides to export the values of probabilities  $q_h$  and  $p_{hk}$  to the whole, potentially infinite population of past, present and future human beings, then the middle side of relation (8) must be used to express the probability  $\mathbf{P}(C_h | I_k)$ .

Now, for any  $k \in \{0, \dots, 5\}$ , the law  $\mathbf{P}(\cdot | I_k) : h \in \{0, \dots, 5\} \rightarrow \mathbf{P}_k(h) = \mathbf{P}(C_h | I_k)$  is a probability distribution on the “state space”  $\{0, \dots, 5\}$  induced, for any examined subject, by the experimental result about the value of parameter  $Y$ , and it should be carefully noted that, when applied to the whole population, it *does not depend* on the BDI class to which the subject was assigned *a priori*, and this is quite natural since, as it has been just seen, each probability  $\mathbf{P}(C_h | I_k)$  is computed under the assumption to ignore that class. In order to determine the probabilities  $\mathbf{P}(C_h | C_r \cap I_k)$ , that take into account the information about the BDI class to which the subject is assumed to belong before the experiment, one has to solve the following system of equations:

$$\mathbf{P}(C_h | I_k) = \sum_{r=0}^5 \mathbf{P}(C_h | C_r \cap I_k) \mathbf{P}(C_r | I_k),$$

that is

$$n_{hk} = \sum_{r=0}^5 n_{rk} \mathbf{P}(C_h | C_r \cap I_k). \quad (9)$$

Notice that, referring to system (7), one has  $\mathbf{P}(C_h | C_r \cap I_j) = \Phi_{r,j}^h$  (where the index relative to the parameter has been dropped as useless for the sake of simplicity), so that the above system gives the transition probabilities needed to use system (7).

### 5. Detecting Symptoms of Depression: Computing Parameters on a Sample Study

Now, system (9) can be easily solved if one has enough time and patience, or simply a computer. Nevertheless, the general case will be left aside and the above procedure will be now applied to a reference sample already used to the same aim in [34]. It could be objected that the considered sample cannot be defined “sufficiently large”. As a matter of fact, the present application aims to be nothing more than an example of the way in which the method should work, and a suitable extension (and better validation) of the method outlined here is planned to be the object of a forthcoming paper. Furthermore, just in connection with the small size of our sample, the consideration of all the five possible states defined by BDI will be renounced, and only *two* states will be considered, 0 = healthy, and 1 = depressed. This simplifies the procedure and, at the same time, can provide the reader with a sufficiently clear plot of the method.

So, in the case under consideration, we have  $n \equiv |S| = 22$ ,  $n_0 \equiv |C_0| = n_1 \equiv |C_1| = 11$ . Consider first the parameter  $T$ , for which  $I_0(T) = [285.79, 422.88]$ ,  $I_1(T) = (115.88, 285.79] \cup [422.88, 1005.01)$ , and

$$\begin{aligned} n_{00}(T) &= 7 \\ n_{01}(T) &= 4 \\ n_{10}(T) &= 2 \\ n_{11}(T) &= 9. \end{aligned} \quad (10)$$

System (9) takes the form

$$\begin{aligned} n_{00} &= n_{00} \mathbf{P}(C_0 | C_0 \cap I_0) + n_{10} \mathbf{P}(C_0 | C_1 \cap I_0) \\ n_{01} &= n_{01} \mathbf{P}(C_0 | C_0 \cap I_1) + n_{11} \mathbf{P}(C_0 | C_1 \cap I_1) \\ n_{10} &= n_{00} \mathbf{P}(C_1 | C_0 \cap I_0) + n_{10} \mathbf{P}(C_1 | C_1 \cap I_0) \\ n_{11} &= n_{01} \mathbf{P}(C_1 | C_0 \cap I_1) + n_{11} \mathbf{P}(C_1 | C_1 \cap I_1). \end{aligned} \quad (11)$$

that is

$$\begin{aligned} 7 &= 7\mathbf{P}(C_0 | C_0 \cap I_0) + 2\mathbf{P}(C_0 | C_1 \cap I_0) \\ 4 &= 4\mathbf{P}(C_0 | C_0 \cap I_1) + 9\mathbf{P}(C_0 | C_1 \cap I_1) \end{aligned} \quad (12)$$

because a straightforward calculation, based on the relation  $\mathbf{P}(C_1 | C_0 \cap I_0) = 1 - \mathbf{P}(C_0 | C_0 \cap I_0)$ , shows at once that system (11) is redundant, and that the third and the fourth of Equation (11) are reproductions of the first and second ones respectively.

Now, as a system of two equations with four unknowns, system (12) does not allow us to determine all the required transition probabilities: two of them must be assigned as parameters. Therefore, setting  $\mathbf{P}(C_0 | C_0 \cap I_0) = \lambda$  and  $\mathbf{P}(C_1 | C_1 \cap I_1) = \mu$ , system (11) becomes

$$\begin{aligned} \mathbf{P}(C_0 | C_1 \cap I_0) &= \frac{7}{2}(1 - \lambda) \\ \mathbf{P}(C_0 | C_0 \cap I_1) &= \frac{9\mu - 5}{4} \end{aligned} \quad (13)$$

and the parameters must satisfy the conditions

$$\begin{aligned} \frac{5}{7} &\leq \lambda \leq 1, \\ \frac{5}{9} &\leq \mu \leq 1. \end{aligned} \quad (14)$$

The above procedure can be reproduced step by step for each of the two remaining parameters  $A$  and  $O$ . One only needs to insert in system (11) the values of  $n_{hk}$  resulting from Tables 1–6 in [34] for these parameters. These values are listed in the following Table 1 (while, for the sake of completeness, the intervals  $I_0$  and  $I_1$  are specified for all the three parameters  $T$ ,  $A$  and  $O$  in Tables 2–4).

So, with the same definition of parameters  $\lambda$  and  $\mu$  and the same calculations as before, we find for  $\lambda$  and  $\mu$  the conditions collected in Table 5.

**Remark 9.** As probabilities,  $\lambda$  and  $\mu$  belong to  $[0, 1]$ , and the above conditions, as derived by means of straightforward algebraic computations, simply give lower bounds to their values. As a consequence, the negative values in the third column simply mean that, for the parameter  $O$ ,  $\lambda$  and  $\mu$  are free to run over the whole interval  $[0, 1]$ .

Finally, it is possible to write the matrices of the transition probabilities for all three parameters in consideration. More precisely,

$$\Phi_0 \equiv (\Phi^r_{s0}) = \begin{pmatrix} \lambda & 1 - \lambda \\ \frac{7}{2}(1 - \lambda) & \frac{7}{2}\lambda - \frac{5}{2} \end{pmatrix}, \quad \Phi_1 \equiv (\Phi^r_{s1}) = \begin{pmatrix} \frac{9\mu - 5}{4} & \frac{9}{4}(1 - \mu) \\ 1 - \mu & \mu \end{pmatrix}, \quad (15)$$

for the parameter  $T$ ;

$$\Phi_0 \equiv (\Phi^r_{s0}) = \begin{pmatrix} \lambda & 1 - \lambda \\ 4(1 - \lambda) & 4\lambda - 3 \end{pmatrix}, \quad \Phi_1 \equiv (\Phi^r_{s1}) = \begin{pmatrix} 3\mu - 2 & 3(1 - \mu) \\ 1 - \mu & \mu \end{pmatrix}, \quad (16)$$

for the parameter  $A$ , and

$$\Phi_0 \equiv (\Phi^r_{s0}) = \begin{pmatrix} \lambda & 1 - \lambda \\ \frac{7}{8}(1 - \lambda) & \frac{1}{8}(1 + 7\lambda) \end{pmatrix}, \quad \Phi_1 \equiv (\Phi^r_{s1}) = \begin{pmatrix} \frac{3}{4}(1 - \mu) & \frac{1}{4}(1 + 3\mu) \\ 1 - \mu & \mu \end{pmatrix}, \quad (17)$$

for the parameter  $O$ .

**Remark 10.** Notice that, though the same symbols  $\lambda$  and  $\mu$  have been used for all three parameters  $T$ ,  $A$  and  $O$ , they need not to take the same values in all cases. On the contrary,  $\lambda$  and  $\mu$  will take in general different values from case to case. The reason for the differences will be explained in Remark 11.

**Table 1.** Values of  $n_{hk}$  for  $T$ ,  $A$ ,  $O$ .

$n$	$T$	$A$	$O$
$n_{00}$	7	8	7
$n_{01}$	4	3	4
$n_{10}$	2	2	8
$n_{11}$	9	9	3

**Table 2.** Confidence intervals for  $T$ .

$Y = T$	
$I_0(Y)$	[285.79, 422.88]
$I_1(Y)$	(115.88, 285.79] $\cup$ [422.88, 1200.01)

**Table 3.** Confidence intervals for  $A$ .

$Y = A$	
$I_0(Y)$	[14.62, 21.30]
$I_1(Y)$	[5.27, 14.62] $\cup$ [21.30, 35.89]

**Table 4.** Confidence intervals for  $O$ .

$Y = O$	
$I_0(Y)$	[51.56, 75.18]
$I_1(Y)$	[43.48, 51.56] $\cup$ [75.18, 84.21]

**Table 5.** Algebraic conditions on the least values of  $\lambda$  and  $\mu$  for  $T$ ,  $A$ ,  $O$ .

	$T$	$A$	$O$
$\lambda$	5/7	3/4	0
$\mu$	5/9	2/3	0

**Remark 11.** The meaning of the parametric nature of the solution as well as of these last conditions now requires careful discussion. As a matter of fact, beyond the mere algebraic condition, there is at least one deep conceptual reason why two of the above transition probabilities cannot be determined only on the base of the experienced relative frequencies of intervals  $I_k$ : a careful check of the above conditions on parameters, as well as of the consequences of assigning to them either the least or the greatest value allowed, shows that the choice of the values for  $\lambda$  and  $\mu$  is strictly related to our trust in the effectiveness of our experiments. More precisely, consider for simplicity only parameter  $T$ , and assume to have chosen  $\lambda = 5/7$ . It follows that  $\mathbf{P}(C_0 | C_0 \cap I_0) = 5/7$ ,  $\mathbf{P}(C_1 | C_0 \cap I_0) = 3/7$ ,  $\mathbf{P}(C_0 | C_1 \cap I_0) = 1$  and  $\mathbf{P}(C_1 | C_1 \cap I_0) = 0$ . In addition, set  $\mu = 1$ . It follows that  $\mathbf{P}(C_0 | C_0 \cap I_1) = 1$ ,  $\mathbf{P}(C_1 | C_0 \cap I_1) = 0$ ,  $\mathbf{P}(C_0 | C_1 \cap I_1) = 0$  and  $\mathbf{P}(C_1 | C_1 \cap I_1) = 1$ . So, this choice for parameters  $\lambda$  and  $\mu$  corresponds to ascribing to the experiment an heavy bias towards the result «healthy» (or «not depressed»). Indeed, a subject classified as «healthy» by BDI is still classified «healthy» after the experiment even if the value of  $T$  belongs to  $I_1$ , while a subject classified as «depressed» by BDI is classified «healthy» after the experiment when the value of  $T$  belongs to  $I_0$ . Conversely, choose  $\lambda = 1$  and  $\mu = 5/9$ . One finds  $\mathbf{P}(C_0 | C_0 \cap I_0) = 1$ ,  $\mathbf{P}(C_1 | C_0 \cap I_0) = 0$ ,  $\mathbf{P}(C_0 | C_1 \cap I_0) = 0$ ,  $\mathbf{P}(C_1 | C_1 \cap I_0) = 1$ ,  $\mathbf{P}(C_0 | C_0 \cap I_1) = 0$ ,  $\mathbf{P}(C_1 | C_0 \cap I_1) = 1$ ,  $\mathbf{P}(C_0 | C_1 \cap I_1) = 4/9$  and  $\mathbf{P}(C_1 | C_1 \cap I_1) = 5/9$ , which clearly correspond to a bias towards the state «depressed».

These considerations can be repeated word by word for parameters  $A$  and  $O$ , and lead us to the conclusion that in each case the values of parameters  $\lambda$  and  $\mu$  must be «tuned» according to our trust in the effectiveness of our experiment. At a first glance, the choice that seems not to introduce biases is

$$\lambda^* = \frac{\lambda_{\min} + \lambda_{\max}}{2}$$

$$\mu^* = \frac{\mu_{\min} + \mu_{\max}}{2}.$$

Of course, this choice requires to be revised and corrected after a sufficiently large number of additional experiments, by comparison with other evidence.

**Remark 12.** It is readily seen that with the above choice of the values of  $\lambda$  and  $\mu$  the matrices (15)–(17) assume the forms

$$\Phi_0 \equiv (\Phi^r_{s0}) = \begin{pmatrix} 6/7 & 1/7 \\ 1/2 & 1/2 \end{pmatrix}, \quad \Phi_1 \equiv (\Phi^r_{s1}) = \begin{pmatrix} 1/2 & 1/2 \\ 2/9 & 7/9 \end{pmatrix}, \quad (18)$$

for the parameter  $T$ ;

$$\Phi_0 \equiv (\Phi^r_{s0}) = \begin{pmatrix} 7/8 & 1/8 \\ 1/2 & 1/2 \end{pmatrix}, \quad \Phi_1 \equiv (\Phi^r_{s1}) = \begin{pmatrix} 1/2 & 1/2 \\ 1/6 & 5/6 \end{pmatrix}, \quad (19)$$

for the parameter  $A$ , and

$$\Phi_0 \equiv (\Phi^r_{s0}) = \begin{pmatrix} 1/2 & 1/2 \\ 7/16 & 9/16 \end{pmatrix}, \quad \Phi_1 \equiv (\Phi^r_{s1}) = \begin{pmatrix} 3/8 & 5/8 \\ 1/2 & 1/2 \end{pmatrix}, \quad (20)$$

for the parameter  $O$ , respectively.

Now, it is to be noted that while the transition matrices for parameters  $T$  and  $A$  have one row  $(1,0)$  or  $(0,1)$  corresponding to a certain transition of the subject from the state  $C_1$  to the state  $C_0$  or vice versa according to whether the value of the parameter belongs to  $I_0$  or to  $I_1$ , the corresponding rows for the parameter  $O$  are in both cases  $(1/2, 1/2)$ , expressing a complete uncertainty about the «final» state of the examined subject. This allows us to state that parameters  $T$  and  $A$  contribute information much richer than that transmitted by  $O$ , so that the computation of the value of parameter  $O$  could also be given up, or should be used to «put in doubt» the results of any previous sequence of experiments. This point will be illustrated more clearly through examples in the next Section.

## 6. Detecting Symptoms of Depression: How to Classify New Cases

The present Section is devoted to show the way in which the above model and the data obtained from the basic sample can be used to update the probability distribution on the states 0 (healthy) and 1 (depressed) for any additional case. In particular, two subjects not contained in the starting sample will be considered here, the former assigned by BDI in the class  $C_1$  (depressed), the latter assigned instead to the class  $C_0$  (healthy).

The first case, who from now on will be referred to as “S37”, a forty year-old female labeled without uncertainty as «depressed» by BDI, has obtained for the parameters  $T$ ,  $A$  and  $O$  the values shown in Table 6, which shows that S37 is depressed according to each experiment.

**Table 6.** Values of parameters  $T$ ,  $A$ ,  $O$  and confidence intervals for S37.

	Value	Interval
$T$	207.85	$I_1$
$A$	8.66	$I_1$
$O$	48%	$I_1$

Using these results, apply system (7) to each parameter for subject S37. As laid out in Section 4, we choose to repeatedly apply system (7), first to  $T$ , next to  $A$  and finally to  $O$ . Starting with parameter  $T$ , one has

$$\begin{aligned}
 f_0^{(1)}(t_1) &= \Phi_{1;1}^0(1,2)f_1^{(1)}(t_0)f_1^{(2)}(t_1) = \frac{2}{9} \approx 22\% \\
 f_1^{(1)}(t_1) &= 1 - \Phi_{1;1}^0(1,2)f_1^{(1)}(t_0)f_1^{(2)}(t_1) = \frac{7}{9} \approx 78\%.
 \end{aligned}
 \tag{21}$$

By applying the same system to  $A$ , starting from the updated probability distribution, we obtain

$$\begin{aligned}
 f_0^{(1)}(t_1) &= \frac{2}{9} + \Phi_{1;1}^0(1,2)f_1^{(1)}(t_0)f_1^{(2)}(t_1) - \Phi_{0;1}^1(1,2)f_0^{(1)}(t_0)f_1^{(2)}(t_1) = \\
 &= \frac{2}{9} + \frac{1}{6} \cdot \frac{7}{9} - \frac{1}{2} \cdot \frac{2}{9} = \frac{13}{54} \approx 24\% \\
 f_1^{(1)}(t_1) &= \frac{41}{54} \approx 76\%.
 \end{aligned}
 \tag{22}$$

Again, by repeating step-by-step the above procedure for the parameter  $O$ , we find

$$\begin{aligned}
 f_0^{(1)}(t_1) &= \frac{13}{54} + \Phi_{1;1}^0(1,2)f_1^{(1)}(t_0)f_1^{(2)}(t_1) - \Phi_{0;1}^1(1,2)f_0^{(1)}(t_0)f_1^{(2)}(t_1) = \\
 &= \frac{13}{54} + \frac{2}{3} \cdot \frac{41}{54} - \frac{1}{2} \cdot \frac{13}{54} = \frac{203}{324} \approx 63\% \\
 f_1^{(1)}(t_1) &= \frac{255}{432} \approx 37\%.
 \end{aligned}
 \tag{23}$$

In conclusion, the probability distribution on the «sample space»  $\{0, 1\}$  can be updated according the following plot, where in any pair of values, the first is the probability of the healthy state, while the second is that of the depressed state:

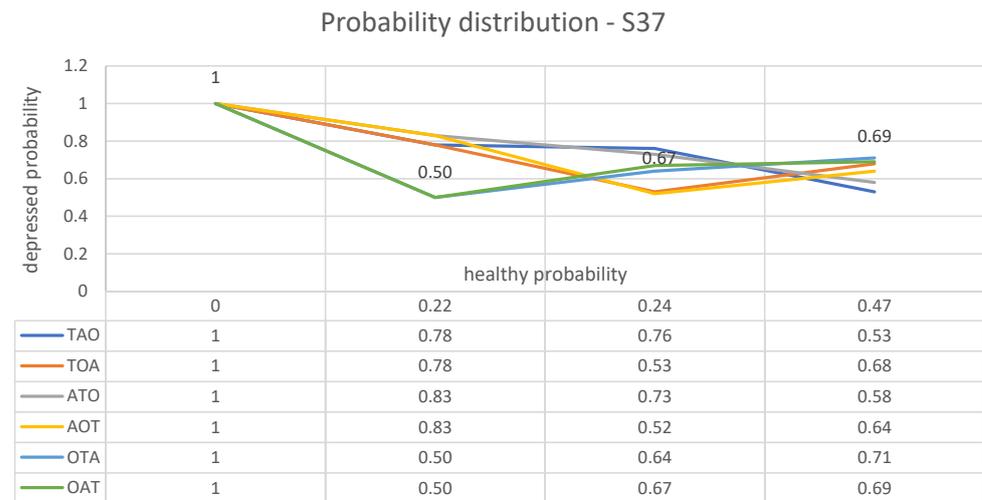
$$\begin{aligned}
 \xrightarrow{BDI} (0, 1) &\xrightarrow{T} (0.22, 0.78) \xrightarrow{A} (0.24, 0.76) \xrightarrow{O} (0.47, 0.53) \\
 \xrightarrow{BDI} (0, 1) &\xrightarrow{T} (0.22, 0.78) \xrightarrow{O} (0.47, 0.53) \xrightarrow{A} (0.32, 0.68) \\
 \xrightarrow{BDI} (0, 1) &\xrightarrow{A} (0.17, 0.83) \xrightarrow{T} (0.27, 0.73) \xrightarrow{O} (0.42, 0.58) \\
 \xrightarrow{BDI} (0, 1) &\xrightarrow{A} (0.17, 0.83) \xrightarrow{O} (0.48, 0.52) \xrightarrow{T} (0.36, 0.64) \\
 \xrightarrow{BDI} (0, 1) &\xrightarrow{O} (0.50, 0.50) \xrightarrow{T} (0.36, 0.64) \xrightarrow{A} (0.29, 0.71) \\
 \xrightarrow{BDI} (0, 1) &\xrightarrow{O} (0.50, 0.50) \xrightarrow{A} (0.33, 0.67) \xrightarrow{T} (0.31, 0.69).
 \end{aligned}$$

These paths from initial health and depression probabilities, as assigned to subject S37 by BDI, to final probabilities obtained according to different test arrangements are illustrated in Figure 2.

The second subject examined here, who from now on will be referred to as “S08”, is labelled (without uncertainty) as healthy by BDI, but turns out to be depressed by all the parameters  $T$ ,  $A$  and  $O$ , as results from Table 7.

**Table 7.** values of parameters  $T$ ,  $A$ ,  $O$  and confidence intervals for S08.

	Value	Interval
$T$	1119.51	$I_1$
$A$	33.92	$I_1$
$O$	47.06%	$I_1$

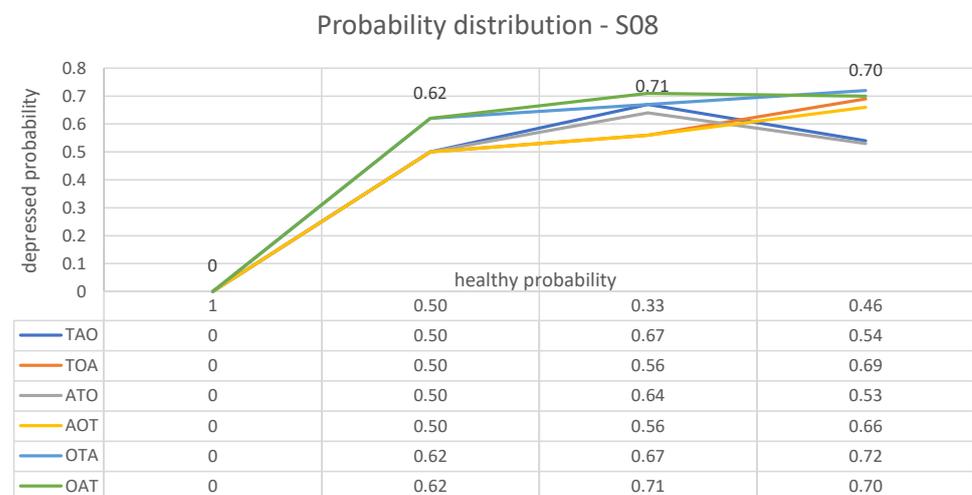


**Figure 2.** Graphic representation of the way in which the estimated probabilities of health and depression for subject S37 vary after each speech test, depending on the different relative importance assigned to each measured parameter. The coordinates of the edges of each plot are the health and depression probabilities of subject S37 after each test. More precisely, the *x*-coordinate is the health probability and the *y*-coordinate is the depression probability.

For subject S08, the following list of plots is obtained:

$$\begin{aligned}
 & \underline{BDI} \rightarrow (1, 0) \quad \underline{T} \rightarrow (0.50, 0.50) \quad \underline{A} \rightarrow (0.33, 0.67) \quad \underline{O} \rightarrow (0.46, 0.54) \\
 & \underline{BDI} \rightarrow (1, 0) \quad \underline{T} \rightarrow (0.50, 0.50) \quad \underline{O} \rightarrow (0.44, 0.56) \quad \underline{A} \rightarrow (0.31, 0.69) \\
 & \underline{BDI} \rightarrow (1, 0) \quad \underline{A} \rightarrow (0.50, 0.50) \quad \underline{T} \rightarrow (0.36, 0.64) \quad \underline{O} \rightarrow (0.47, 0.53) \\
 & \underline{BDI} \rightarrow (1, 0); \quad \underline{A} \rightarrow (0.50, 0.50) \quad \underline{O} \rightarrow (0.44, 0.56) \quad \underline{T} \rightarrow (0.34, 0.66) \\
 & \underline{BDI} \rightarrow (1, 0) \quad \underline{O} \rightarrow (0.38, 0.62) \quad \underline{T} \rightarrow (0.33, 0.67) \quad \underline{A} \rightarrow (0.28, 0.72) \\
 & \underline{BDI} \rightarrow (1, 0) \quad \underline{O} \rightarrow (0.38, 0.62) \quad \underline{A} \rightarrow (0.29, 0.71) \quad \underline{T} \rightarrow (0.30, 0.70).
 \end{aligned}$$

These paths from initial health and depression probabilities, as assigned to subject S08 by BDI, to final probabilities obtained according to different test arrangements are illustrated in Figure 3.



**Figure 3.** Graphic representation of the way in which the estimated probabilities of health and depression for subject S08 vary after each speech test, depending on the different relative importance assigned to each measured parameter.

Finally, consider again a subject—who will be identified as S15—labelled as depressed by BDI, but described as healthy by two indicators and as depressed by only one of them according to Table 8 below.

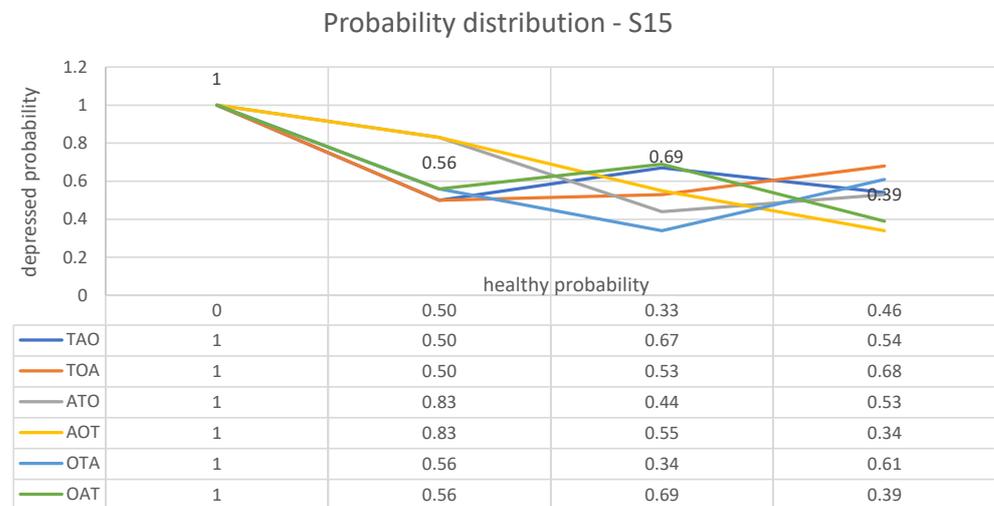
**Table 8.** Values of parameters *T*, *A*, *O* and confidence intervals for S15.

	Value	Interval
<i>T</i>	289.03	$I_0$
<i>A</i>	12.57	$I_1$
<i>O</i>	58.33%	$I_0$

For subject S15, the following list of plots follows:

$$\begin{aligned}
 & \underline{BDI} \rightarrow (0, 1) \quad \underline{T} \rightarrow (0.50, 0.50) \quad \underline{A} \rightarrow (0.33, 0.67) \quad \underline{O} \rightarrow (0.46, 0.54) \\
 & \underline{BDI} \rightarrow (0, 1) \quad \underline{T} \rightarrow (0.50, 0.50) \quad \underline{O} \rightarrow (0.47, 0.53) \quad \underline{A} \rightarrow (0.32, 0.68) \\
 & \underline{BDI} \rightarrow (0, 1) \quad \underline{A} \rightarrow (0.17, 0.83) \quad \underline{T} \rightarrow (0.56, 0.44) \quad \underline{O} \rightarrow (0.47, 0.53) \\
 & \underline{BDI} \rightarrow (0, 1); \underline{A} \rightarrow (0.17, 0.83) \quad \underline{O} \rightarrow (0.45, 0.55) \quad \underline{T} \rightarrow (0.66, 0.34) \\
 & \underline{BDI} \rightarrow (0, 1) \quad \underline{O} \rightarrow (0.44, 0.56) \quad \underline{T} \rightarrow (0.66, 0.34) \quad \underline{A} \rightarrow (0.39, 0.61) \\
 & \underline{BDI} \rightarrow (0, 1) \quad \underline{O} \rightarrow (0.44, 0.56) \quad \underline{A} \rightarrow (0.31, 0.69) \quad \underline{T} \rightarrow (0.61, 0.39).
 \end{aligned}$$

These paths from initial health and depression probabilities, as assigned to subject S15 by BDI, to final probabilities obtained according to different test arrangements are illustrated in Figure 4.



**Figure 4.** Graphic representation of the way in which the estimated probabilities of health and depression for subject S15 vary after each speech test, depending on the different relative importance assigned to each measured parameter.

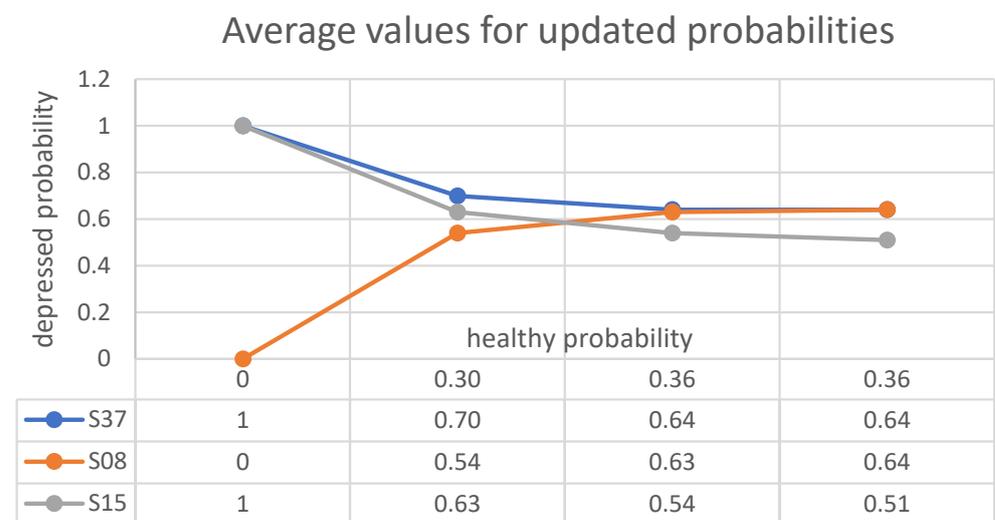
According to the above results, it seems that the probabilities of health and depression for each subject, when updated according to the outcomes of tests about indicators *T*, *A* and *O*, depend on the order in which the values of indicators are registered and compared. Nevertheless, at least when the indicators are all in accordance, the variations in the final probability distributions are very limited. This suggests referring to the average probabilities (on the set of all possible permutations of the triple (*T*, *A*, *O*)) rather than choosing any one of the values listed in the final column, and to compute explicitly the standard deviation in order to get a quantitative expression of their spreading. More precisely, let us denote by  $\mu_h$  and  $\mu_d$ , respectively, the average values of probabilities

for healthy and depressed state at each step (I, II or III), and by  $\sigma$  all the corresponding standard deviations. Thus, postponing to the next Section a discussion of the meaning of the results, we may collect them in Table 9.

**Table 9.** Average values and standard deviations for updated probabilities.

I			II			III			
$\mu_h$	$\mu_d$	$\sigma$	$\mu_h$	$\mu_d$	$\sigma$	$\mu_h$	$\mu_d$	$\sigma$	
0.30	0.70	0.16	0.36	0.64	0.10	0.36	0.64	0.07	<b>S37</b>
0.46	0.54	0.06	0.37	0.63	0.06	0.36	0.64	0.08	<b>S08</b>
0.37	0.63	0.16	0.46	0.54	0.13	0.49	0.51	0.13	<b>S15</b>

The results on  $\mu_h$  (and, consequently, on  $\mu_d$ ) listed in the above table are illustrated in Figure 5.



**Figure 5.** The plots above depict the evolution of the *average* depression probabilities for subjects S37, S08 and S15. The coordinates of the edges of the plot are the average health and depression probabilities of these subjects at each step, as updated by the *triples* of outcomes of experiments. A comparison of this picture with Figures 2–4 highlights that the final results of the different test arrangements are much more dispersed around the average results for subjects S08 and S15 than for subject S37. This seems to reflect the presence at each step of results in apparent mutual contrast.

## 7. Discussion and Research Perspectives

As it has been seen in Section 6, the order in which the values of parameters  $T$ ,  $A$  and  $O$  are considered influences the resulting probability distribution on the set {healthy, depressed}. Nevertheless, when the values of all these parameters are in agreement (i.e., they assign a subject to the same BDI class), the procedure turns out to be approximately «commutative», in the sense that the probability distributions obtained by following the different plots differ very little from each other. A perceivable difference can be observed when some parameters confirm BDI classification and others refute it. It is precisely this circumstance that suggests to use the averages of the outcomes of different plots, taking also into account the very small values of standard deviations at each step.

At a first glance, it can be considered as a loss of information to have transformed the BDI classification, which is assumed to be sharp, into a probabilistic statement. In this connection, however, it must be recalled that, strictly speaking, BDI classification is in turn probabilistic, and our choice to assign the values 0 and 1 to the initial probabilities simply expresses our absolute confidence in the response of the questionnaire. If one questions at least partially the veracity of this decision, and make different hypotheses about the reliability of the additional indicators (for instance, assuming them more accurate

in negative evaluations than in positive ones or vice versa), so that parameters  $\lambda$  and  $\mu$  are given different values, then the probabilistic character of all the results becomes evident.

Naturally, the assignment of the values of the initial probabilities and of the parameters  $\lambda$  and  $\mu$  should not be thought of as definitive, but in turn subjected to repeated updates by further experiments on more subjects. This immediately defines the natural development of the research and the problems to be tackled.

First, it is evident that the reference sample used to define the confidence intervals of the parameters introduced here should be greatly expanded; a sufficient expansion would certainly modify the ranges of possible values for  $\lambda$  and  $\mu$  and their choice. It can be expected that after a sufficiently large number of examined subjects, both ranges and choices will stabilize, especially if one establishes the «relative weight» that should be assigned to the BDI initial classification and to the values of parameters  $T$ ,  $A$  and  $O$ , respectively. So, the first step in the future development of this research must be to increase the size of the reference sample. In addition, the values of  $\lambda$  and  $\mu$  should be arbitrarily changed in order to determine through experimental controls the different degrees of belief that could be assigned to BDI and to parameters  $T$ ,  $A$  and  $O$ , respectively.

In connection with the expansion of the sample, of particular interest would be to conduct an analysis of the depressive effects of COVID on a large sample, both through the BDI and through speech analysis. Unfortunately, we do not yet have sufficient data on this subject, but we intend to develop research in this direction as soon as we have collected enough data.

A natural second step will then consist in automatizing the computation of the transition matrices and, as a consequence, of the solutions of Equation (7)<sub>1</sub>. In such a way, Equation (7)<sub>1</sub> will behave similarly to a tool for training (on the reference sample) and then to use on any subsequent subject in the form of software for diagnosing depression.

Finally, the question about the possibility of reproducing the method presented here for any kind of diagnostic problems arises spontaneously, once the experimental devices and the parameters of interest will be defined for them. The answer can be readily seen to be affirmative. A very good example of this kind of development of the studies about possible applications of the method presented here can be found in [37,38], where a study of mental health based on an electroencephalogram headset is proposed: the application of such a study should be of the greatest interest with regard to depression and the comparison of its results with those of the BDI using the method presented here. A more general exposition of the method, showing at once its independence of the context, and its effectiveness in solving any classification problem, is given in [14].

**Author Contributions:** Conceptualization, F.V., B.C. and A.E.; methodology, F.V., B.C. and A.E.; validation, F.V., B.C. and A.E.; formal analysis, B.C.; investigation, F.V.; resources, A.E.; data curation, A.E.; writing—original draft preparation, B.C.; writing—review and editing, B.C.; visualization, F.V.; supervision, A.E.; project administration, A.E.; funding acquisition, A.E. All authors have read and agreed to the published version of the manuscript.

**Funding:** The present paper has received funding from the project ANDROIDS, Università della Campania “Luigi Vanvitelli” V:ALERE 2019, D. R. 906/2019.

**Informed Consent Statement:** Informed consent was obtained from all subjects involved in the study.

**Data Availability Statement:** Data available on demand at the Department of Psychology of University of Campania “L. Vanvitelli”: the person responsible is dr Anna Esposito.

**Conflicts of Interest:** The authors declare no conflict of interest.

## References

1. Kanter, J.W.; Busch, A.M.; Weeks, C.E.; Labdes, S.J. The nature of clinical depression: Symptoms, syndromes, and behavior analysis. *Behav. Anal.* **2008**, *31*, 1–21. [\[CrossRef\]](#)
2. Remes, O.; Mendes, J.F.; Templeton, P. Biological, Psychological, and Social Determinants of Depression: A Review of Recent Literature. *Brain Sci.* **2021**, *11*, 1633. [\[CrossRef\]](#) [\[PubMed\]](#)
3. Esposito, A.; Raimo, G.; Maldonato, M.N.; Vogel, C.; Conson, M.; Cordasco, G. Behavioral Sentiment Analysis of Depressive States. In Proceedings of the 11th IEEE International Conference on Cognitive Infocommunications (CogInfoCom), Mariehamn, Finland, 23–25 September 2020; pp. 209–214.
4. Beck, A.T.; Alford, B.A. *Depression: Causes and Treatment*; University of Pennsylvania Press: Philadelphia, PA, USA, 2009.
5. Greco, C.; Matarazzo, O.; Cordasco, G.; Vinciarelli, A.; Callejas, Z.; Esposito, A. Discriminative Power of EEG-Based Biomarkers in Major Depressive Disorder: A Systematic Review. *IEEE Access* **2021**, *9*, 112850–112870. [\[CrossRef\]](#)
6. Lee, Y.-S.; Park, W.-H. Diagnosis of depressive disorder model on facial expression based on fast R-CNN. *Diagnostics* **2022**, *12*, 317. [\[CrossRef\]](#) [\[PubMed\]](#)
7. Likforman-Sulem, L.; Esposito, A.; Faundez-Zanuy, M.; Cléménçon, S.; Cordasco, G. EMOTHAW: A Novel Database for Emotional State Recognition from Handwriting and Drawing. *IEEE Trans. Hum.-Mach. Syst.* **2017**, *47*, 273–284. [\[CrossRef\]](#)
8. Liu, D.; Liu, B.; Lin, T.; Liu, G.; Yang, G.; Qi, D.; Qiu, Y.; Lu, Y.; Yuan, Q.; Shuai, S.C.; et al. Measuring depression severity based on facial expression and body movement using deep convolutional neural network. *Front. Psychiatry* **2022**, *13*, 1017064. [\[CrossRef\]](#)
9. Nolazco-Flores, J.A.; Faundez-Zanuy, M.; Velázquez-Flores, O.A.; Del-Valle-Soto, C.; Cordasco, G.; Esposito, A. Mood state detection in handwritten tasks using PCA—mFCBF and automated machine learning. *Sensors* **2022**, *22*, 1686. [\[CrossRef\]](#)
10. Tao, F.; Ge, X.; Ma, W.; Esposito, A.; Vinciarelli, A. Multi-Local Attention for Speech-Based Depression Detection. In Proceedings of the ICASSP 2023–2023 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP), Rhodes Island, Greece, 4–10 June 2023. [\[CrossRef\]](#)
11. Tao, F.; Esposito, A.; Vinciarelli, A. Spotting the Traces of Depression in Read Speech: An Approach Based on Computational Paralinguistics and Social Signal Processing. In Proceedings of the INTERSPEECH 2020, Shanghai, China, 25–29 October 2020; pp. 1828–1832. [\[CrossRef\]](#)
12. Xue, D.; Guo, X.; Li, Y.; Sheng, Z.; Wang, L.; Liu, L.; Cao, J.; Liu, Y.; Lou, J.; Li, H.; et al. Risk Factor Analysis and a Predictive Model of Postoperative Depressive Symptoms in Elderly Patients Undergoing Video-Assisted Thoracoscopic Surgery. *Brain Sci.* **2023**, *13*, 646. [\[CrossRef\]](#)
13. Alohoban, N.; Esposito, A.; Vinciarelli, A. What You Say or How You Say It? Depression Detection Through Joint Modelling of Linguistic and Acoustic Aspects of Speech. *Cogn. Comput.* **2021**, *14*, 1585–1598. [\[CrossRef\]](#)
14. Vitale, F.; Carbonaro, B.; Esposito, A. A methodological application of a time-discrete version of the kinetic-theoretic language to interactions between random variables. *J. Math. Phys.* **2023**, submitted.
15. Bellomo, N.; Degond, P.; Tadmor, E. (Eds.) *Active Particles, Volume 1: Advances in Theory, Models, and Applications*; Birkhäuser: Basel, Switzerland, 2017.
16. Boltzmann, L. *Lectures on Gas Theory*; Courier Corporation: Chelmsford, MA, USA, 2012.
17. Bellomo, N.; Carbonaro, B.; Gramani, L. On the kinetic and stochastic games theory for active particles: Some reasonings on open large living systems. *Math. Comput. Model.* **2008**, *48*, 1047–1054. [\[CrossRef\]](#)
18. Bianca, C. Thermostated kinetic equations as models for complex systems in physics and life sciences. *Phys. Life Rev.* **2012**, *9*, 359–399. [\[CrossRef\]](#) [\[PubMed\]](#)
19. Bianca, C.; Brézin, L. Modeling the antigen recognition by B-cell and T-cell receptors through thermostated kinetic theory methods. *Int. J. Biomath.* **2017**, *10*, 1750072. [\[CrossRef\]](#)
20. Bianca, C.; Riposo, J. Mimic therapeutic actions against keloid by thermostated kinetic theory methods. *Eur. Phys. J. Plus* **2015**, *130*, 159. [\[CrossRef\]](#)
21. Carbonaro, B. Modeling epidemics by means of the stochastic description of complex systems. *Comput. Math. Methods* **2021**, *3*, 1208–1220. [\[CrossRef\]](#)
22. Gramani, L. On the modeling of granular traffic flow by the kinetic theory for active particles trend to equilibrium and macroscopic behavior. *Int. J. Non-Linear Mech.* **2009**, *44*, 263–268. [\[CrossRef\]](#)
23. Iannini, M.L.L.; Dickman, R. Kinetic theory of vehicular traffic. *Am. J. Phys.* **2016**, *84*, 135–145. [\[CrossRef\]](#)
24. Bellomo, N.; Brezzi, F. Traffic, crowds and swarms. *Math. Model. Methods Appl. Sci.* **2008**, *18*, 1145–1148. [\[CrossRef\]](#)
25. Bertotti, M.L.; Modanese, G. Economic inequality and mobility in kinetic models for social sciences. *Eur. Phys. J. Spec. Top.* **2019**, *225*, 1945–1958. [\[CrossRef\]](#)
26. Bertotti, M.L.; Modanese, G. Microscopic models for welfare measures addressing a reduction of economic inequality. *Complexity* **2016**, *21*, 89–98. [\[CrossRef\]](#)
27. Maldarella, D.; Pareschi, L. Kinetic models for socio-economic dynamics of speculative markets. *Phys. A Stat. Mech. Its Appl.* **2012**, *391*, 715–730. [\[CrossRef\]](#)
28. Toscani, G.; Tosin, A.; Zanella, M. Multiple-interaction kinetic modeling of a virtual-item gambling economy. *Phys. Rev. E* **2019**, *100*, 012308. [\[CrossRef\]](#) [\[PubMed\]](#)
29. Bassetti, F.; Toscani, G. Explicit equilibria in bilinear kinetic models for socio-economic interactions. *ESAIM Proc. Surv.* **2014**, *47*, 1–16. [\[CrossRef\]](#)

30. Bianca, C.; Carbonaro, B.; Menale, M. On the Cauchy problem of vectorial thermostatted kinetic frameworks. *Symmetry* **2020**, *12*, 517–535. [[CrossRef](#)]
31. Ehmann, B.; Balázs, L.; Fülöp, ; Hargitai, R.; Kabai, P.; Péley, B.; Pólya, T.; Vargha, A.; László, J. Narrative psychological content analysis as a tool for psychological status monitoring of crews in isolated, confined and extreme settings. *Acta Astronaut.* **2011**, *68*, 1560–1566. [[CrossRef](#)]
32. Lopez-Otero, P.; Docio-Fernandez, L.; Garcia-Mateo, C. A study of acoustic features for depression detection. In Proceedings of the International Workshop on Biometrics and Forensics (IWBF), Valletta, Malta, 27–28 March 2014; pp. 1–6.
33. Beck, A.T. Cognitive models of depression. In *Clinical Advances in Cognitive Psychotherapy: Theory on Application*; Leahy, R.L., Dowd, E.T., Eds.; Springer Publishing Company: New York, NY, USA, 2002; pp. 29–61.
34. Vitale, F.; Carbonaro, B.; Cordasco, G.; Esposito, A.; Marrone, S.; Raimo, G.; Verde, L. A Privacy-Oriented Approach for Depression Signs Detection based on Speech Analysis. *Electronics* **2021**, *10*, 2986–3001. [[CrossRef](#)]
35. Cercignani, C. *The Boltzmann Equation and Its Applications*; Springer Science+Business Media: New York, NY, USA, 1988.
36. Esposito, A.; Esposito, A.M.; Likforman-Sulem, L.; Maldonato, N.M.; Vinciarelli, A. On the significance of speech pauses in depressive disorders: Results on read and spontaneous narratives. In *Recent Advances in Nonlinear Speech Processing, Smart Innovation, Systems and Technologies*; Esposito, A., Faundez-Zanuy, M., Esposito, A.M., Cordasco, G., Drugman, T., Solé-Casals, J., Morabito, F.C., Eds.; Springer: Cham, Switzerland, 2016; Volume 48, pp. 73–82.
37. Marin, I.; Dinescu, I.-A.; Deleanu, T.-C.; Sulaiman, L.A.; Al-Gayar, S.M.S.; Goga, N. Brain Performance Analysis based on an Electroencephalogram Headset. In Proceedings of the 12th International Conference on Electronics, Computers and Artificial Intelligence (ECAI), Bucharest, Romania, 25–27 June 2020.
38. Marin, I. Study of Mental Health and Learning Engagement during COVID-19 Pandemic based on an Electroencephalogram Headset. In Proceedings of the 13th annual International Conference of Education, Research and Innovation (ICERI2020), Online Conference, 9–10 November 2020.

**Disclaimer/Publisher’s Note:** The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.