An Algorithm to Calculate a Quantum Probability Space

Massimo Melucci
University of Padua, Italy
massimo.melucci@unipd.it

1 Introduction

Many scientific disciplines need probability spaces to fit observed data, predict future data or explain relationships. To this end, a *probability space* should be utilized to represent events by sets and leverage the closure under conjunction, disjunction and negation [6] to examine more complex events according to the tie with the Boolean logic [3].

Context¹ is viewed as the complex of experimental conditions in which uncertain and related events are observed in terms of data. Kolmogorov implicitly assumes a context when he refers to "a complex of conditions which allows of any number of repetitions" [6, page 3]. In this paper, we explicitly state that a probability space is the mathematical representation of a context because the space provides a representation of contextual events, their relationships and measures of the uncertainty affecting the observed and the predicted data. The uniqueness of the context from which data are observed is crucial when an event is predicted conditionally to the observation of other events. As a context coincides with one probability space, the prediction of the event conditionally to the other events would be possible if the observed data could define a single probability space and then only one context.

¹Context origins from Latin *contextus*, from *con-*'together' + *texere* 'to weave'.

In this paper we address the problem of using one probability space for estimating parameters and predicting future data when the observed data come from multiple contexts and thus from distinct spaces. We explain that a set-based probabilistic space might be suboptimal in the case of multiple contexts. To overcome suboptimality and reconcile multiple contexts in one space, the paper introduces the Quantum Probability Space (QPS) whose foundations were described in [4]. We also present an algorithm to calculate the QPS for data observed from multiple contexts and provide a web application that implements the algorithm².

2 Probability Spaces

In probability, a random experiment is an experimental context where the data are observed in conditions of uncertainty. As the events need a numerical representation, a random variable is a function mapping an event observed in a random experiment to a point of an interval of the n-dimensional data space; for example, a binary variable maps coin toss to $\{0,1\}$. Random variables provide a succinct and sufficient description of events because the events mapped to a certain data are in the same subspace; for example, the movies a person does (or does not) like are in the same subspace labeled by A (or respectively \bar{A}) including all the movies that the user likes (or does not like); a random variable maps A (or \bar{A}) to 1 (or 0).

A probability space is the mathematical representation of a random experiment. Formed by a set of events and by a probability function, or density operator, a probability space assigns a probability to each subspace with three fundamental properties. The empty subspace is mapped to 0, the whole space is mapped to 1, and for any pairwise disjoint subspaces, the probability of the disjunction is the sum of the probabilities assigned to each subspace. (Two subspaces are disjoint when one is included in the orthogonal subspace of the other.)

²http://isotta.dei.unipd.it/cgi-bin/qps/qps-w-form.py

Suppose A_1, A_2, A_3 be three subsets corresponding to n=3 binary variables that assign an event to a subset if and only if the value observed for the corresponding variable is one. After measuring the three variables for a certain sample, the marginal probabilities like $P(A_2A_3)$ are available to calculate $P(A_1A_2A_3) = P(A_1|A_2A_3) P(A_2A_3)$; however, the number of subsets is 2^n , thus making the calculation of probability for any number of variables infeasible for non-small n. Conditional independence can be a helpful workaround for overcoming the problem of the exponential number of subsets. Two variables A_2 and A_3 are conditionally independent on A_1 when

$$P(A_2A_3|A_1) = P(A_2|A_1)P(A_3|A_1)$$
(1)

Conditional independence requires marginal probability values. Howver, a set-based probability space implies some statistical inequalities between the marginal probability values. Therefore, the violation of an inequality implies the inexistence of a single set-based probability space. For example, suppose that the following marginal probabilities are provided by some distinct contexts:

$$P(A_1 A_2) = \frac{9}{20} \qquad P(A_1 A_3) = \frac{9}{20} \qquad P(A_2 A_3) = \frac{1}{10}$$

$$P(\bar{A}_1) = \frac{1}{2} \qquad P(\bar{A}_2) = \frac{1}{2} \qquad P(\bar{A}_3) = \frac{1}{2}$$
(2)

A probability space cannot exist for A_1 , A_2 , A_3 because no $P(A_1A_2A_3)$ can be calculated; otherwise, we would have to accept sets of negative volume. The situation is similar to Euclid's theorem according to which the inner angles of any triangle shall sum to π only if placed on a plane. If placed on a type of surface other than a plane, the angles of the triangles of stars might not sum to π . If the observed angles violates Euclid's theorem, planarity does not hold [1].

A test of existence of a single probability space for any n is provided in [8]. Let A_1, A_2, A_3 be three binary variables. Suppose we are provided with $P(A_1), P(A_2), P(A_3), P(A_1A_2), P(A_1A_3), P(A_2A_3)$ from distinct experi-

mental contexts. Let

$$\ell = \max\{0, P(A_1A_2) + P(A_1A_3) - P(A_1), P(A_1A_2) + P(A_2A_3) - P(A_2), P(A_1A_3) + P(A_2A_3) - P(A_3)\}$$
 and

$$v = \min\{P(A_1A_2), P(A_1A_3), P(A_2A_3), 1 - (P(A_1) + P(A_2) + P(A_3) - P(A_1A_2) - P(A_1A_3) - P(A_2A_3))\}$$

The following inequality holds if A_1, A_2, A_3 refer to the same space:

$$\ell \le P(A_1 A_2 A_3) \le v \tag{3}$$

In other words, one context would be possible only if (3) held. Although there are other inequalities to consider [2], there is a general result that holds for every n [8].

In other words, if (3) does not hold, then $A_1A_2A_3$ cannot be defined, and distributivity, that is

$$A_1(A_2 \vee A_3) = (A_1 A_2) \vee (A_1 A_3) = (A_1 A_2 \bar{A}_3) \vee (A_1 A_2 A_3) \vee (A_1 \bar{A}_2 A_3)$$

cannot hold. Since distributivity is a feature of sets, eventually one set-based probability space does not exist for all the variables A_1, A_2, A_3 when (3) does not hold.

One approach to calculating $P(A_1A_2A_3)$ when $A_1A_2A_3$ cannot exist is to utilize conditional independence (1). In this way, $P(A_1A_2A_3)$ can be approximated even though $A_1A_2A_3$ cannot exist. However, some efficiency is lost when using conditional independence as explained in Section 3.

3 Probabilistic Ranking

Let A_1, A_2, A_3 be three binary variables corresponding to three subspaces. A decider has to split the set of observed values for A_1, A_2, A_3 in an acceptance region (\mathcal{A}) and its complement; the acceptance region is the subset of triples of binary values such that the decider takes one out of two options; for

example, a classifier decides for each class whether a triple of values observed for an event is in the acceptance region and therefore if the event should be put in a class. Let

$$P_i(\mathcal{A}) = \sum_{A_1, A_2, A_3 \in \mathcal{A}} P_i(A_1 A_2 A_3)$$

be the likelihood of class i given the observed data. The system's decision is optimal when the set of values is split in such a way as to maximize $P_1(A)$ while keeping the likelihood of the complement class small.

The Neyman-Pearson's Lemma (NPL) states that optimal decision can be obtained when $P_1(\mathcal{A}) - cP_0(\mathcal{A})$ is maximum provided that $P_0(\mathcal{A}) \leq \alpha$ [7]. While varying c, the decider produces a ranking; at the top of the ranking, the decider puts the first items, while the least preferred items are ranked at the bottom of the ranking.

However, optimal decision require the existence of $A_1A_2A_3$ which cannot be taken for granted if there are distinct experimental contexts, although it might not be calculated nor approximated by conditional independence.

Suppose A_1, A_2, A_3 are measured in three distinct contexts corresponding to the pairs (A_1, A_2) , (A_1, A_3) and (A_2, A_3) , of which marginal probabilities may violate (3). Suppose the system ranks items by assuming conditional independence (1). If following marginal probabilities are estimated

$$P(A_1 A_2) = \frac{1}{4}$$
 $P(A_1 A_3) = \frac{1}{4}$ $P(A_2 A_3) = \frac{1}{4}$

$$P(\bar{A}_1) = \frac{1}{2}$$
 $P(\bar{A}_2) = \frac{1}{2}$ $P(\bar{A}_3) = \frac{1}{2}$

the following triples, $\bar{A}_1A_2A_3$, $A_2\bar{A}_1\bar{A}_3$, $\bar{A}_1\bar{A}_2A_3$, $\bar{A}_1\bar{A}_2\bar{A}_3$, $A_1A_2\bar{A}_3$, $A_1\bar{A}_2A_3$, $A_1\bar{A}_2\bar{A}_3$ will be equally ranked. According to (3), $P(A_1A_2A_3)$ ranges between $\ell=0$ and $v=\frac{1}{4}$; therefore, $A_1A_2A_3$ may be ranked before, after or coincidentally with any other triple depending on the probability space and provided the same marginal probabilities. As there may be an infinity of $P(A_1A_2A_3)$

satisfying (3) provided the aforementioned marginal probabilities, the approximation of $P(A_1A_2A_3)$ by $P(A_2|A_1)P(A_3|A_1)$ is only one out of the infinity of admittable values whereas the true and unknown value of $P(A_1A_2A_3)$ might be in any of the points of the interval $[\ell, v]$. As $P(A_2|A_1)P(A_3|A_1)$ is only one out of the infinity of admittable values, the ranking resulting from the assumption of conditional independence may place $A_1A_2A_3$ on the top or on the bottom, thus causing suboptimal ranking. To overcome suboptimality, $P(A_1A_2A_3)$ should be calculated and not only approximated through conditional independence. However, when the number of variables is not small, the number of probability calculation becomes exponentially large.

In the next two sections we show that the QPS calculates $P(A_1A_2A_3)$, avoids suboptimality and optimally ranks events even though (3) is violated.

4 The Quantum Probability Space

Sets are not the only framework of a probabilistic space. Instead of sets, the QPS utilizes vectors and the operators thereof. In this section, we first define the QPS in terms of subspaces and probability function; then, we introduce an algorithm to compute the subspaces and the probability function.

4.1 Definition of the QPS

Let \mathcal{H} be a vector space and $x \in \mathcal{H}$ be a vector, i.e. the simplest subspace. Let 0 be the null vector. The join $x \oplus y$ is the smallest subspace including both x and y. Let Q be a probability function; the QPS is given by \mathcal{H} and Q where $Q(x), x \in \mathcal{H}$ is the probability of x and

- 1. Q(0) = 0
- $2. \ Q(\mathcal{H}) = 1$
- 3. $Q(x \oplus y) = Q(x) + Q(y)$ for any pair of disjoint subspaces x and y.

We label the output probability space as "quantum" because it is based on the mathematical formalism of Quantum Mechanics. For this reason, we use Q instead of P, which is left for set-based probability spaces.

Essential to the QPS is Gleason's theorem [5]. If an event is represented by a vector $x \in \mathcal{H}$, the probability function Q(x) is necessarily a density matrix ρ such that

$$Q(x) = x'\rho x \tag{4}$$

is a bilinear quadratic form. The hard part of Gleason's theorem is calculating ρ . In this paper, we explain an algorithm to the aim of calculating ρ that can reproduce the marginal probabilities and calculate the probabilities of the events when they cannot be expressed by the set-based space.

The QPS can measure variables outside the framework based on sets. In the framework based on vector spaces, Q may still be calculated even though the marginal probabilities violate (3). As a consequence, a single QPS can be defined although the marginal probabilities come from distinct contexts. Of course, the probabilities provided by the QPS might differ from those provided by a set-based space even though both spaces provide the same marginal probabilities. Therefore, the QPS may give another ranking of items. Whether this alternative ranking is better than the ranking provided by a set-based space is a matter of experimentation.

4.2 An Algorithm to Calculate the QPS

In this section, we explore the search for a unique probability space to move to a theoretical framework other than Kolmogorov's. The basic idea is that optimality of ranking may be recovered if (3) can be violated. To this end, the algorithm introduced in this paper takes marginal probabilities as input and gives one QPS as output, although the variables are measured from distinct experimental contexts and may violate (3).

In general, the problem is as follows. Suppose that there are n binary variables A_1, \ldots, A_n and m = n(n+1)/2 univariate or bivariate marginal

probabilities

$$P(\bar{A}_i)$$
 $P(A_iA_j)$ $i=1,\ldots,n-1$ $j=i+1,\ldots,n$

Let $b_n = i_1 i_2 \cdots i_n \in \{0,1\}^n$ be a *n*-digit binary string³. We have that the event

$$A_1 = i_1, \dots, A_n = i_n$$

can be shortly written in terms of canonical vectors of $\{0,1\}^n$ as follows:

$$A_{b_n} = i_1 \cdots i_n$$

The algorithm calculates ρ such that

$$P(\bar{A}_i) = \left(\bigoplus_{b \in b_{i-1}0b_{n-i}} x_b'\right) \rho \left(\bigoplus_{b \in b_{i-1}0b_{n-i}} x_b\right)$$

$$i = 1, \dots, n$$

$$P(A_i A_j) = \left(\bigoplus_{b \in b_{i-1}1b_{j-i}1b_{n-j-1}} x_b'\right) \rho \left(\bigoplus_{b \in b_{i-1}1b_{j-i}1b_{n-j-1}} x_b\right)$$

$$i = 1, \dots, n-1$$

$$j = i+1, \dots, n$$

and

$$0 \le x_b' \rho x_b \le 1$$
 $\sum_{b \in \{0,1\}^n} x_b' \rho x_b = 1$

particular, a generic element of the Λ 's diagonal is defined as

$$\lambda_i = \begin{cases} P(\bar{A}_i) & i = 1, \dots, n \\ P(A_k A_j) & i = (k-1)n + j \quad k = 1, \dots, n - 1 \quad j = k+1, \dots, n \end{cases}$$

³When n = 0, b_0 is an empty string.

```
Require: n binary variables A_1, \ldots, A_n such that either \bar{A}_i or A_i.
Require: P(\bar{A}_i) for i = 1, ..., n.
Require: P(A_i A_j) for i = 1, ..., n - 1 and j = i + 1, ..., n.
 1: m \leftarrow \frac{n(n+1)}{2}
2: N \leftarrow 2^n
 3: {Build the m \times N matrix K as follows:}
 4: for all i = 1, ..., n do
       {Fill row i by alternating 2^{n-i} ones and 2^{n-i} zeros.}
 6: end for
 7: {Fill the remaining \frac{n(n-1)}{2} rows as follows:}
 8: for all \ell = 1, ..., N do
 9:
       k \leftarrow n
10:
       for all i = 0, ..., n - 1 do
          for all j = i + 1, \ldots, n do
11:
             if K[i, \ell] = K[j, \ell] = 0 then
12:
                K[k,\ell] \leftarrow 1
13:
             else
14:
                K[k,\ell] \leftarrow 0
15:
             end if
16:
             k \leftarrow k + 1
17:
18:
          end for
       end for
19:
20: end for
21: for all i = 1, ..., n do
22:
       \lambda_i \leftarrow P(\bar{A}_i)
23: end for
24: for all i = 1, ..., n - 1 do
       for all j = i + 1, ..., n do
25:
          k \leftarrow \frac{1}{2} (n (n-1) - (n-i) (n-i-1)) + j
26:
          \lambda_k \leftarrow P(A_i A_i)
27:
       end for
28:
29: end for
30: \{Compute (8)\}
31: {Compute (9)}
32: {Compute (10)}
33: {Compute (11)}
```

Figure 1: An algorithm for computing the density matrix given the marginal probabilities of n binary variables.

The general algorithm can be described in Figure 1. First, a $n \times N$ binary matrix K is generated (lines 1–20). The algorithm proceeds with lines 21–29 where the input marginal probabilities are arranged in a diagonal matrix Λ . Finally, the matrices introduced in the rest of this section are computed (lines 30–33). In particular, as the algorithm has to reproduce all the marginal probabilities, we must calculate a matrix R such that

$$KRK' = \Lambda$$

where K is $m \times N$, K' is the transpose conjugate, K^+ is called pseudo-inverse, R is $N \times N$, and Λ is $m \times m$. For any complex $m \times N$ matrix K, the pseudo-inverse of K is any $N \times m$ matrix K^+ such that

$$KK^{+}K = K$$
 $KK^{+} = (KK^{+})'$ (5)

$$K^{+}KK^{+} = K^{+}$$
 $K^{+}K = (K^{+}K)'$ (6)

where K' is the conjugate transpose of K. One can prove that K^+ is unique and certainly exists. Moreover, any complex $m \times N$ matrix of rank k is pseudo-diagonal when it has only zeros except for k diagonal elements.

For any complex $m \times N$ matrix K of rank k, the Singular Value Decomposition (SVD) of K is the product

$$K = V S U' \tag{7}$$

where V is a $m \times m$ unitary matrix, S is a $m \times N$ pseudo-diagonal matrix, and U is a $N \times N$ unitary matrix. Any complex $m \times N$ matrix K of rank k admits a SVD, although not unique. When a matrix is Hermitean, the SVD is an eigen-decomposition where V = U and Σ is a real matrix.

After multiplying both sides of (7) by K' on the left and by K on the right, we obtain

$$K'KRK'K = K'\Lambda K$$

Let

$$J = K'K \tag{8}$$

As J is a Hermitean matrix on a finitely dimensional vector space,

$$J = U \Sigma U' \tag{9}$$

where $U = (u_1, \ldots, u_N)$ and Σ is a diagonal matrix such that $\operatorname{diag}(\Sigma) = (\sigma_1, \ldots, \sigma_N)$. If $\sigma_i \neq \sigma_j$ unless i = j then (9) is unique. If (9) is applied we have that:

$$JRJ = K' \Lambda K$$

$$U \Sigma U'RU \Sigma U' = K' \Lambda K$$

$$\Sigma U'RU \Sigma = U' K' \Lambda K U$$

$$U'RU = \Sigma^{+} U'K' \Lambda K U \Sigma^{+}$$

and finally

$$R = U \Sigma^{+} U' K' \Lambda K U \Sigma^{+} U'$$

$$\tag{10}$$

The marginal probabilities can be restored by computing the following expression:

Finally, we have that

$$\rho = \operatorname{tr}(R)^{-1} R \tag{11}$$

The probability function of the QPS can also be expressed as a system of quadratic equations as follows. Let

$$W = U \Sigma^+ U' K'$$

and $w_{b,i}$ be an element of W. The b-th diagonal element of (11) is the probability provided by the QPS of the event represented by x_b and can be

written as

$$Q(x_b) = \sum_{i=1}^{N} w_{b,i}^2 \lambda_i \tag{12}$$

Note that the number of columns of K grows to an exponential order of n, since there are $N=2^n$ combinations of binary values. Therefore, the number of variables should then be kept as small as possible. In some applications such as Information Retrieval (IR) the number of variables equals the number of query terms, which is usually small. When n is not small, some heuristics can ameliorate the computational cost. Consider the n(n-1)/2 marginal bivariate probabilities and in particular those of events such as A_1A_2 and $A_1\bar{A}_2$. Each event is represented by a row of K with two 1's and N-2 zeros. As such a row is very sparse, it can be efficiently stored in compressed format, thus only requiring memory space for storing the column index of the 1's. Moreover, the first n rows of K might not be stored since such rows may result from the join of the corresponding univariate events; for example, the row of A_1 results from the join of the rows of A_1A_2 and $A_1\bar{A}_2$ and can be calculated only if needed.

4.3 Two Examples of QPS

For starters, consider three binary variables A_1, A_2, A_3 . Consider also their conjunctions (A_1A_2) , (A_1A_3) , (A_2A_3) in the sense that each conjunction is measured in one experimental context. Then, we correspond each triple, such as $A_1A_2A_3$, to a canonical vector; for example, $\bar{A}_1\bar{A}_2\bar{A}_3$ corresponds to the vector $x_0' = (1, 0, 0, 0, 0, 0, 0, 0)$ and $\bar{A}_1\bar{A}_2A_3$ corresponds to the vector $x_1' = (0, 1, 0, 0, 0, 0, 0, 0)$; the join of the two vectors results in the plane corresponding to $\bar{A}_1\bar{A}_2$, that is,

$$x_0' \oplus x_1' = (1, 1, 0, 0, 0, 0, 0, 0)$$

and $x_0' \oplus x_1' \oplus x_2' \oplus x_3'$ will be vector

$$(1 \ 1 \ 1 \ 1 \ 0 \ 0 \ 0 \ 0)$$

If the join is repeated for all the events corresponding to the six measured variables, the following matrix is obtained

$$K = \begin{pmatrix} 1 & 1 & 1 & 1 & 0 & 0 & 0 & 0 \\ 1 & 1 & 0 & 0 & 1 & 1 & 0 & 0 \\ 1 & 0 & 1 & 0 & 1 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 \end{pmatrix}$$

where m is the number of marginal probabilities and N is the number of basis vectors. The first row corresponds to $P(\bar{A}_1)$ because it is obtained by the join of the subspaces representing $\bar{A}_1\bar{A}_2\bar{A}_3$, $\bar{A}_1\bar{A}_2A_3$, $\bar{A}_1A_2\bar{A}_3$ and $\bar{A}_1A_2A_3$. It can be easily checked that, if K[i] is the i-th row of K we have that

$$K[1] \rho K[1]' = P(\bar{A}_1)$$
 $K[4] \rho K[4]' = P(A_1 A_2)$
 $K[2] \rho K[2]' = P(\bar{A}_2)$ $K[5] \rho K[5]' = P(A_1 A_3)$
 $K[3] \rho K[3]' = P(\bar{A}_3)$ $K[6] \rho K[6]' = P(A_2 A_3)$

Therefore, the problem of calculating one probability space can be stated as the problem of calculating a density matrix ρ such that

$$P(\bar{A}_{1}) = (x'_{0} \oplus x'_{1} \oplus x'_{2} \oplus x'_{3}) \ \rho \ (x_{0} \oplus x_{1} \oplus x_{2} \oplus x_{3})$$

$$P(\bar{A}_{2}) = (x'_{0} \oplus x'_{1} \oplus x'_{4} \oplus x'_{5}) \ \rho \ (x_{0} \oplus x_{1} \oplus x_{4} \oplus x_{5})$$

$$P(\bar{A}_{3}) = (x'_{0} \oplus x'_{2} \oplus x'_{4} \oplus x'_{6}) \ \rho \ (x_{0} \oplus x_{2} \oplus x_{4} \oplus x_{6})$$

$$P(A_{1}A_{2}) = (x'_{6} \oplus x'_{7}) \ \rho \ (x_{6} \oplus x_{7})$$

$$P(A_{1}A_{3}) = (x'_{5} \oplus x'_{7}) \ \rho \ (x_{5} \oplus x_{7})$$

$$P(A_{2}A_{3}) = (x'_{3} \oplus x'_{7}) \ \rho \ (x_{3} \oplus x_{7})$$

and

$$0 \le x_b' \, \rho \, x_b \le 1$$
 for all $b \in \{0, 1\}^3$ such that $\sum_{b \in \{0, 1\}^3} x_b' \, \rho \, x_b = 1$

where $\rho = R/\text{tr}(R)$. A numerical example is provided in the following. Suppose

$$\Lambda = \operatorname{diag}\left(\frac{1}{2}, \frac{1}{2}, \frac{1}{2}, \frac{9}{20}, \frac{9}{20}, \frac{1}{10}\right)$$

which violates (3) and then does not admit a set-based probability space. We have that the probability distributed along the diagonal of ρ is

$$\{0.0105, 0.242, 0.242, 0.0469, 0.201, 0.115, 0.115, 0.0274\}$$

The sum of the first four values is $P(\bar{A}_1)$. Moreover, suppose

$$\Lambda = \operatorname{diag}\left(\frac{1}{2}, \frac{1}{2}, \frac{1}{2}, \frac{1}{20}, \frac{9}{20}, \frac{1}{10}\right)$$

which does not violate (3). Accordingly, $P(x_1x_2x_3) = \frac{1}{20}$ in a set-based probability space. We have that the probabilities distributed along the diagonal

of ρ are

$$\{0.0134, 0.227, 0.287, 0.0469, 0.234, 0.135, 0.0343, 0.0217\}$$

Therefore, the set-based distribution is not necessarily reproduced although (3) holds. However, the marginal probabilities have again been restored even in the new space:

$$K[1] \rho K[1]' = 0.50$$
 $K[4] \rho K[4]' = 0.05$
 $K[2] \rho K[2]' = 0.50$ $K[5] \rho K[5]' = 0.45$
 $K[3] \rho K[3]' = 0.50$ $K[6] \rho K[6]' = 0.10$

Using (12), an alternative expression of the probability space can be provided in terms of equations as follows:

$$P(\bar{A}_1\bar{A}_2\bar{A}_3) = \frac{0.02\lambda_1 + 0.02\lambda_2 + 0.02\lambda_3 + 0.001\lambda_4 + 0.001\lambda_5 + 0.001\lambda_6}{0.7\lambda_1 + 0.7\lambda_2 + 0.7\lambda_3 + \lambda_4 + \lambda_5 + \lambda_6}$$

$$P(\bar{A}_1\bar{A}_2A_3) = \frac{0.2\lambda_1 + 0.2\lambda_2 + 0.3\lambda_3 + 0.4\lambda_4 + 0.1\lambda_5 + 0.1\lambda_6}{0.7\lambda_1 + 0.7\lambda_2 + 0.7\lambda_3 + \lambda_4 + \lambda_5 + \lambda_6}$$

$$P(\bar{A}_1A_2\bar{A}_3) = \frac{0.2\lambda_1 + 0.3\lambda_2 + 0.2\lambda_3 + 0.1\lambda_4 + 0.4\lambda_5 + 0.1\lambda_6}{0.7\lambda_1 + 0.7\lambda_2 + 0.7\lambda_3 + \lambda_4 + \lambda_5 + \lambda_6}$$

$$P(\bar{A}_1A_2A_3) = \frac{0.001\lambda_1 + 0.001\lambda_2 + 0.001\lambda_3 + 0.07\lambda_4 + 0.07\lambda_5 + 0.6\lambda_6}{0.7\lambda_1 + 0.7\lambda_2 + 0.7\lambda_3 + \lambda_4 + \lambda_5 + \lambda_6}$$

$$P(A_1\bar{A}_2\bar{A}_3) = \frac{0.3\lambda_1 + 0.2\lambda_2 + 0.2\lambda_3 + 0.1\lambda_4 + 0.1\lambda_5 + 0.4\lambda_6}{0.7\lambda_1 + 0.7\lambda_2 + 0.7\lambda_3 + \lambda_4 + \lambda_5 + \lambda_6}$$

$$P(A_1\bar{A}_2A_3) = \frac{0.001\lambda_1 + 0.001\lambda_2 + 0.001\lambda_3 + 0.07\lambda_4 + 0.6\lambda_5 + 0.07\lambda_6}{0.7\lambda_1 + 0.7\lambda_2 + 0.7\lambda_3 + \lambda_4 + \lambda_5 + \lambda_6}$$

$$P(A_1A_2\bar{A}_3) = \frac{0.001\lambda_1 + 0.001\lambda_2 + 0.001\lambda_3 + 0.6\lambda_4 + 0.07\lambda_5 + 0.07\lambda_6}{0.7\lambda_1 + 0.7\lambda_2 + 0.7\lambda_3 + \lambda_4 + \lambda_5 + \lambda_6}$$

$$P(A_1A_2A_3) = \frac{0.001\lambda_1 + 0.001\lambda_2 + 0.001\lambda_3 + 0.6\lambda_4 + 0.07\lambda_5 + 0.07\lambda_6}{0.7\lambda_1 + 0.7\lambda_2 + 0.7\lambda_3 + \lambda_4 + \lambda_5 + \lambda_6}$$

In the event of 4 binary observables, we have the following:

and

$$\operatorname{diag}(\Lambda) = (P(\bar{A}_1), P(\bar{A}_2), P(\bar{A}_3), P(\bar{A}_4), P(A_1A_2), P(A_1A_3), P(A_1A_4), P(A_2A_3), P(A_2A_4), P(A_3A_4))$$

The diagonal of R is the distribution of probability of A_1, A_2, A_3, A_4 . The

linear equation system can be written as follows:

$$P(\bar{A}_1\bar{A}_2\bar{A}_3\bar{A}_4) = \frac{0.003(\lambda_1 + \lambda_2 + \lambda_3 + \lambda_4 + \lambda_5 + \lambda_6 + \lambda_7 + \lambda_8 + \lambda_9 + \lambda_{10})}{0.6(\lambda_1 + \lambda_2 + \lambda_3 + \lambda_4) + 0.8(\lambda_5 + \lambda_6 + \lambda_7 + \lambda_8 + \lambda_9 + \lambda_{10})}$$

$$P(\bar{A}_1\bar{A}_2\bar{A}_3\bar{A}_4) = \frac{0.05(\lambda_1 + \lambda_2 + \lambda_3 + \lambda_4) + 0.8(\lambda_5 + \lambda_6 + \lambda_7 + \lambda_8 + \lambda_9 + \lambda_{10})}{0.6(\lambda_1 + \lambda_2 + \lambda_3 + \lambda_4) + 0.8(\lambda_5 + \lambda_6 + \lambda_7 + \lambda_8 + \lambda_9 + \lambda_{10})}$$

$$P(\bar{A}_1\bar{A}_2\bar{A}_3\bar{A}_4) = \frac{0.05(\lambda_1 + \lambda_2 + \lambda_3 + \lambda_4) + 0.8(\lambda_5 + \lambda_6 + \lambda_7 + \lambda_8 + \lambda_9 + \lambda_{10})}{0.6(\lambda_1 + \lambda_2 + \lambda_3 + \lambda_4) + 0.8(\lambda_5 + \lambda_6 + \lambda_7 + \lambda_8 + \lambda_9 + \lambda_{10})}$$

$$P(\bar{A}_1\bar{A}_2\bar{A}_3\bar{A}_4) = \frac{0.02(\lambda_1 + \lambda_2) + 0.01(\lambda_3 + \lambda_4) + 0.1\lambda_5 + 0.03(\lambda_6 + \lambda_7 + \lambda_8 + \lambda_9 + \lambda_{10})}{0.6(\lambda_1 + \lambda_2 + \lambda_3 + \lambda_4) + 0.8(\lambda_5 + \lambda_6 + \lambda_7 + \lambda_8 + \lambda_9 + \lambda_{10})}$$

$$P(\bar{A}_1\bar{A}_2\bar{A}_3\bar{A}_4) = \frac{0.05\lambda_1 + 0.03\lambda_2 + 0.05(\lambda_3 + \lambda_4) + 0.07\lambda_5 + 0.06(\lambda_6 + \lambda_7 + \lambda_8 + \lambda_9 + \lambda_{10})}{0.6(\lambda_1 + \lambda_2 + \lambda_3 + \lambda_4) + 0.8(\lambda_5 + \lambda_6 + \lambda_7 + \lambda_8 + \lambda_9 + \lambda_{10})}$$

$$P(\bar{A}_1\bar{A}_2\bar{A}_3\bar{A}_4) = \frac{0.02\lambda_1 + 0.01\lambda_2 + 0.02\lambda_3 + 0.01\lambda_4 + 0.03\lambda_5 + 0.06(\lambda_6 + \lambda_7) + 0.07(\lambda_8 + \lambda_9) + 0.05\lambda_{10}}{0.6(\lambda_1 + \lambda_2 + \lambda_3 + \lambda_4) + 0.8(\lambda_5 + \lambda_6 + \lambda_7 + \lambda_8 + \lambda_9 + \lambda_{10})}$$

$$P(\bar{A}_1\bar{A}_2\bar{A}_3\bar{A}_4) = \frac{0.02\lambda_1 + 0.01\lambda_2 + 0.02\lambda_3 + 0.01\lambda_4 + 0.03\lambda_5 + 0.06(\lambda_6 + \lambda_7) + 0.07(\lambda_8 + \lambda_9 + \lambda_{10})}{0.6(\lambda_1 + \lambda_2 + \lambda_3 + \lambda_4) + 0.8(\lambda_5 + \lambda_6 + \lambda_7 + \lambda_8 + \lambda_9 + \lambda_{10})}$$

$$P(\bar{A}_1\bar{A}_2\bar{A}_3\bar{A}_4) = \frac{0.02\lambda_1 + 0.01\lambda_2 + 0.02\lambda_3 + 0.01\lambda_4 + 0.03\lambda_5 + \lambda_6 + \lambda_7 + \lambda_8 + \lambda_9 + \lambda_{10})}{0.6(\lambda_1 + \lambda_2 + \lambda_3 + \lambda_4) + 0.8(\lambda_5 + \lambda_6 + \lambda_7 + \lambda_8 + \lambda_9 + \lambda_{10})}$$

$$P(\bar{A}_1\bar{A}_2\bar{A}_3\bar{A}_4) = \frac{0.02\lambda_1 + 0.01\lambda_2 + \lambda_3 + \lambda_4 + 0.04\lambda_5 + \lambda_6 + \lambda_7 + \lambda_8 + \lambda_9 + \lambda_{10})}{0.6(\lambda_1 + \lambda_2 + \lambda_3 + \lambda_4) + 0.8(\lambda_5 + \lambda_6 + \lambda_7 + \lambda_8 + \lambda_9 + \lambda_{10})}$$

$$P(\bar{A}_1\bar{A}_2\bar{A}_3\bar{A}_4) = \frac{0.02\lambda_1 + 0.01\lambda_2 + \lambda_3 + \lambda_4 + 0.04\lambda_5 + \lambda_6 + \lambda_7 + \lambda_8 + \lambda_9 + \lambda_{10})}{0.6(\lambda_1 + \lambda_2 + \lambda_3 + \lambda_4) + 0.8(\lambda_5 + \lambda_6 + \lambda_7 + \lambda_8 + \lambda_9 + \lambda_{10})}$$

$$P(\bar{A}_1\bar{A}_2\bar{A}_3\bar{A}_4) = \frac{0.03\lambda_1 + 0.004(\lambda_2 + \lambda_3 + \lambda_4) + 0.8(\lambda_5 + \lambda_6 + \lambda_7 + \lambda_8 + \lambda_9 + \lambda_{10})}{0.6(\lambda_1 + \lambda_2 + \lambda_3 + \lambda_4) + 0.8(\lambda_5 + \lambda_6 + \lambda_7 + \lambda_8 + \lambda_9 + \lambda_$$

An implementation of the QPS's algorithm can be utilized at

http://isotta.dei.unipd.it/cgi-bin/qps/qps-w-form.py

by using any $3 \le n \le 14$.

References

- [1] L. Accardi. The probabilistic roots of the quantum mechanical paradoxes. In S. Diner, D. Fargue, G. Lochak, and F. Selleri, editors, *The Wave-Particle Dualism: A Tribute to Louis de Broglie on his 90th Birthday*, pages 297–330, Dordrecht, 1984. Springer Netherlands.
- [2] L. Accardi and A. Fedullo. On the statistical meaning of complex numbers in quantum mechanics. *Lettere al nuovo cimento*, 34(7):161–172, June 1982.
- [3] G. Boole. An Investigation of the laws of Thought. Walton and Maberly, 1854.
- [4] R. Feynman, R. Leighton, and M. Sands. *The Feynman Lectures On Physics*. Addison-Wesley, 1965.
- [5] A. M. Gleason. Measures on the closed subspaces of a Hilbert space. Journal of Mathematics and Mechanics, 6:885–893, 1957.
- [6] A. Kolmogorov. Foundations of the Theory of Probability. Chelsea Publishing Company, New York, II edition, 1956.
- [7] J. Neyman and E. Pearson. On the problem of the most efficient tests of statistical hypotheses. *Philosophical Transactions of the Royal Society, Series A*, 231:289–337, 1933.
- [8] I. Pitowsky. Correlation polytopes: Their geometry and complexity. *Mathematical Programming*, 50:395–414, 1991.