Retro-information in Wheeler-Feynman Universe Model: Applications Over an Hypothetical Concept in Quantum Mechanics

Philippe Jacquet, Véronique Joly

INRIA

INSTITUT NATIONAL DE RECHERCHE EN INFORMATIQUE ET EN AUTOMATIQUE

Retro-information in Wheeler-Feynman universe model: Applications over an hypothetical concept in quantum mechanics

Philippe Jacquet et Véronique Joly

No 3530

THÈME 1

ISSN 0249-6399

INRIA
ROCQUENCOURT

Retro-information in Wheeler-Feynman universe model: Applications over an hypothetical concept in quantum mechanics

Philippe Jacquet et Véronique Joly

Thème 1 — Réseaux et systèmes
Projet Hipercom

Rapport de recherche n ° 3530 — — 26 pages

Abstract: Twisting the non-locality concept in quantum mechanics we* introduce the hypothetical concept of retro-information. We analyse the effect of paradoxal coupling on source of retro-information in order to quantify the new means of computing that could be derived from such an hypothetical concept.

Key-words: Information, communication, quantum mechanics, EPR paradox, Shannon law, integer factorization.

(Résumé : tsvp)

* The first author is with INRIA, HIPERCOM project, the second author is with ONERA Chatillon 92322 France

Unité de recherche INRIA Rocquencourt
Domaine de Voluceau, Rocquencourt, BP 105, 78153 LE CHESNAY Cedex (France)
Téléphone : 01 39 63 55 11 - International : +33 1 39 63 55 11
Télécopie : (33) 01 39 63 53 30 - International : +33 1 39 63 53 30

Rétro-information dans un modèle d'univers de Wheeler et Feynman: Applications sur un concept hypothétique en mécanique quantique

Résumé : En malmènant quelque peu les principes de non localité en mécanique quantique, nous introduisons le concept hypothétique de rétro-information. Nous analysons l'effet des couplages paradoxaux de façon à pouvoir quantifier la puissance des nouveaux moyens de calculs qui seraient basés sur ce concept hypothétique.

Mots-clé : Information, télécommunication, mécanique quantique, paradoxe EPR, loi de Shannon, factorisation d'entiers.

1 Introduction

Wheeler and Feynman developed a model of universe based on particle interaction [5]. This model has given rise to the standard model which is in use in most nuclear physics. complete acceptance of WF models needs the assumption that future events could influence past events. This is a mere consequence of non locality in Quantum Mechanics (QM). Non-locality means that under some conditions two particles with a common history continue to interact in the future, even when they are isolated in two distinct space locations. These concepts brought to evidence by Einstein, Podolsky and Rosen in a famous paper [4], are still nowadays hot issues. Non locality effect have been analyzed by Bell [3] and experimented by Aspect [1, 2].

Recently Shor [10], Bennett [8], Gershenfeld *et al.* [9] and Grover [7] have developed the concept of quantum computer which takes advantage from quantum mechanics methodology. Here we propose to get a little further in imagination and describe some potentials in Wheeler-Feynman model in order to allow *retro-information* on small time scale. Here we coin the term of retro-information to mean the ability to transfer informations on an arbitrary space time vector. Among such vectors, those which point toward the past are clearly the most challenging and will be mainly addressed in the present note.

In short retro-information would be possible to the cost of amending some of the product form states present in the standard model. We will not discuss how relevant these amendments could be in modern substitutes of standard model in Quantum Mechanics. We will also absolutely omit all the metaphysical problems possibly raised by retro-information. At least for the comfort of the reader we *a priori* assume that if nevertheless retro-information was eventually possible, then it would concern only very short range transfers (although cascading effects could be possible). For example we could imagine ranges of order some microseconds. Such ranges would certainly not suffice to make practical prediction in stock-exchange market

More interesting our point is to show that even restricted on short very range, retro-information would have the potential to bring interesting breakthroughs in information theory and computer science. Such benefit would be in the direct continuation of the already important benefits expected from quantum computers [10, 7].

It will be shown in the present paper that if retro-information were possible, then time paradox might occur between source and destination. The analysis of time paradox effect in an appropriate model shows that the use of paradoxal coupling would significantly improve the channel capacity from the case without coupling (Shannon case). For example we consider a binary channel with n inputs connected to n outputs. We assume a symmetric case where for each output probability vector is $(p, 1-p)$ when input setting is 0 and $(1-p, p)$ when input setting is 1, with $p > q$. The classical Shannon capacity of such channel is nC^S with

$$C^S = 1 + p\log_2 p + (1-p)\log_2(1-p) \ . \tag{1}$$

In presence of paradoxal coupling we will show that the capacity can be dopped up to nC^D with

$$C^D = \min\{1, \log_2 \frac{p}{1-p}\} \ . \tag{2}$$

Figure 1 displays C^S and C^D versus p.

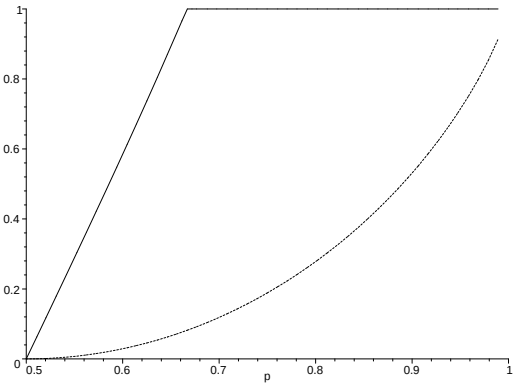

Figure 1: Retro-channel capacity, classical (dashed), dopped (plain)

Furthermore, retro-information could allow interesting improvement in computing, allowing, for example the factorization of any integer of n digit in one multiplication cost and $O(n)$ retro-memory volume. Below we show how such factorization could be programmed.

```
Procedure factorize(Z);
Int Z;
{
  Retro-loop((int X, Y), P)
  {
    P=false;
    If (X*Y=Z) then P=true;
  }
  if (X*Y=Z) then return(X,Y)
    else print ("Z assumed prime")
}
```

The first section introduces non-locality and Bell inequalities. The second section describes how an hypothetical asymmetric conditions could lead to retro-information creation, and retro-channel capacities are quantified. The third section is devoted to the analysis of paradoxal effect *via* micro-branching universe hypothesis. A fourth section discusses the perspective of retro-computing.

2 Non-locality and EPR systems

After a physical event S it is possible to create two twin particles, say particle 1 and particle 2, with exactly opposite spins. A well known experiment consists into measuring the two spins separately on two spin devices (or polarizers) 1 and 2, each of them oriented on respective axes $\mathcal{A}_{\theta_1}$ and $\mathcal{A}_{\theta_2}$ with respective angles θ_1 and θ_2. On each spin device the measure reads "$+1$" or "-1" (see figure 2). This measurement system is known to produce correlations which seem to create *a priori* faster than light distant effect. This paradox has been put in evidence by Einstein, Podolski and Rosen (EPR), and later analyzed by Bell [3]. We denote the pair of measures from the two spin devices by the tuple $(\sigma_1(\theta_1), \sigma_2(\theta_2))$. When the spin

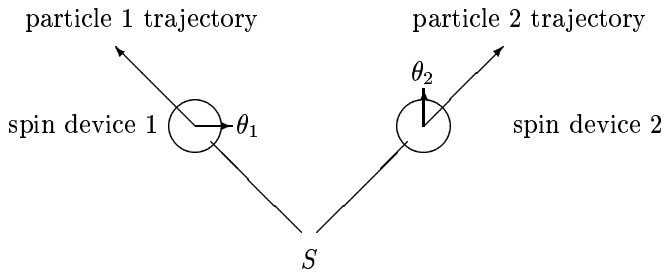

Figure 2: An EPR pair system

devices are aligned, *i.e.* the angles θ_1 and θ_2 are the same, we oberve that the only possible measure tuples are $(+1,-1)$ or $(-1,+1)$ distributed with probability $1/2$.

When the spin devices are misaligned: $\theta_1 - \theta_2 = \theta \neq 0$, measure tuples equal to $(+1,+1)$ and $(-1,-1)$ arise. Theory establishes that in case of particles of spin $1/2$:

$$\begin{cases} \Pr\{\sigma_1(\theta_1) = +1 \& \sigma_2(\theta_2) = -1\} &= \Pr\{\sigma_1(\theta_1) = -1 \& \sigma_2(\theta_2) = +1\} = \frac{1}{2}\cos^2(\frac{\theta}{2}) \\ \Pr\{\sigma_1(\theta_1) = +1 \& \sigma_2(\theta_2) = +1\} &= \Pr\{\sigma_1(\theta_1) = -1 \& \sigma_2(\theta_2) = -1\} = \frac{1}{2}\sin^2(\frac{\theta}{2}) \end{cases} \quad (3)$$

The above probability distribution is known to reveal distant interaction between device 1 and 2 during measurement, as depicted by Bell in 1964 [3].

Theorem 1 (Bell Theorem) *The EPR system is subject to faster than light interaction.*

Proof Assume on the contrary that the vector of all possible answers $\sigma_1(\theta)$ of particle 1 in device 1, for all possible θ, is not affected by measurement on device 2. We will prove that this hypothesis is wrong and thus a distant interaction should occur between device 1 and 2.

Since spins of particle 1 and 2 sums to 0, measuring particle 2 spin on one angle is equivalent to measuring particle 1 on the same angle but with opposite answer, *i.e* for all θ: $\Pr\{\sigma_1(\theta) + \sigma_2(\theta) = 0\} = 1$. Therefore we can determine the second order joint distribution of the answer vector of particle 1 just by refering to identity (3):

$$\Pr\{\sigma_1(\theta_1) = x \& \sigma_1(\theta_2) = y\} = \Pr\{\sigma_1(\theta_1) = x \& \sigma_2(\theta_2) = -y\}$$

The Bell inequalities applies on three events X, Y and Z. $\overline{Z}$ denotes the complementary event of Z:

$$\Pr\{X \& Y\} \leq \Pr\{X \& Z\} + \Pr\{Y \& \overline{Z}\}$$

Let us consider three angles θ_1, θ_2 and θ_3, we assume that event X is $\{\sigma_1(\theta_1) = +1\}$, the event Y: $\{\sigma_1(\theta_2) = +1\}$ and Z: $\{\sigma_1(\theta_3) = +1\}$. Applying Bell inequalities we should have:

$$\begin{aligned} \Pr\{\sigma_1(\theta_1) = +1 \& \sigma_2(\theta_2) = -1\} \quad \leq \quad & \Pr\{\Pr\{\sigma_1(\theta_1) = +1 \& \sigma_2(\theta_3) = -1\} + \\ & + \Pr\{\Pr\{\sigma_1(\theta_2) = +1 \& \sigma_2(\theta_3) = +1\} \end{aligned}$$

which translates into

$$\cos^2(\frac{\theta_1 - \theta_2}{2}) \leq \cos^2(\frac{\theta_1 - \theta_3}{2}) + \sin^2(\frac{\theta_2 - \theta_3}{2})$$

The above inequality does not hold for an infinity of angle values, for example $\theta_1 = 0$, $\theta_2 = \frac{\pi}{2}$ and $\theta_3 = \frac{2\pi}{3}$. Therefore a distant interaction occurs between device 1 and device 2 measurements. When both measurements simultaneously occur, there is therefore a distant effect which occurs faster than light. ■

Theorem 2 *The distance effect on EPR system cannot allow information transfers*

Proof since the probability of alternative answers on each device remains 1/2 whatever the angle setting on the other device. To make information transfers possible we must rely on an hypothetical twist.

3 Branching universes hypothesis and retro-information

3.1 The micro branching universe hypothesis

In 1945 Wheeler and Feynman introduced an original interpretation of QM based on transaction principle. The transaction principle purpose was to solve radiation problems. Wheeler

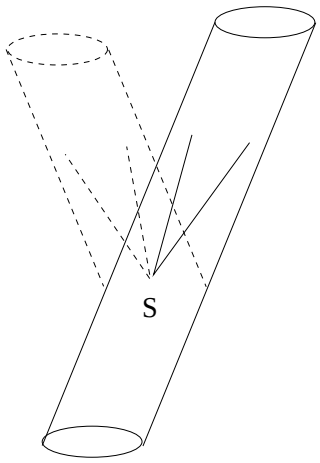

Figure 3: Micro-branching universes, the dead path time is dashed

and Feynman considered that in any QM experiment there is a source event S and a potential set T of target events. A target event is determined by its coordinates x in T. For example one can consider a source of light photons or other particles as event S, and as target set T, a simple screen put on particle trajectory. (with a two slit system in between there would be interference figures on the screen). A target event position x would be the 2D coordinates of particle impact on the screen.

In the transaction interpretation the probability of the target event x is given by the value of the forward wave function hitting position x, namely $\psi(x)$, multiplied by the value of the backward wave function from x hitting initial position S, namely $\psi^*(x)$: $|\psi(x)|^2$. It acts as if the backward wave function echoed on the future position s and comes back to past position S carrying therefore the convolution value $|\psi(x)|^2$.

The branching universe hypothesis is a very rough simplification of the forward-backward transaction hypothesis which assumes the same kind of "echo from the future". For each position x there is a probability $p(x)$ for the particle to hit it and a probability $r(x)$ to persist on it. With probability $1 - r(x)$, the particle rolls back to initial event S.

When rolling back to its initial position the system has no memory of the actual sequence of events between S and position x, which actually stand on the future of S. It is as if space-time universe branches on S on a dead time path and retracts on S *at the same time*. If the particle rollbacks several times, then there will be several dead time paths, and space-time universe will proceed on the "first" path time where the particle does not roll back. This live time path will be called the persistent time path. See figure 3 for an illustration

The persistence probability of a given set I of position has expression $\rho(I) = \int_I p(y)r(y)dy$.

The probability that the persistent time path contains the event "the particle hits the target at position x" has the quadratic expression:

$$\frac{p(x)r(x)}{\rho(T)} \tag{4}$$

According to basic quantum mechanic quantity $p(x)r(x)$ should be proportional to the square of wave function amplitude $|\psi(x)|^2$ or its equivalent. The global persistence probability is $\rho(T)$ and the global rollback probability $1-\rho(T)$. By virtue of the conservation of the flow of the squared wave function the unconditional rollback probability ρ should not depend on the target disposition in the future. We call this rule the *persistence conservation.*

Remark: There is a great temptation to identify in the quadratic expressions (4) $p(x)$ and $r(x)$ to be respectively proportional to $\psi(x)$ and its conjugate $\psi^*(x)$ or any equivalent, if one temporarily forgets the problems about complex numbers. Under this perspective the superposition of two rollback systems could be analoguous to the superposition of two wave functions in QM. These interesting considerations are not part of the present paper.

In the example of the source S of light the displacement of the screen would change the interference figure and therefore the function $r(x)$ but without changing the global persistence probability $\int_T p(y)r(y)dy$.

With EPR systems described in section 2 we should obtain

$$\rho(\{(\sigma_1(\theta_1),\sigma_2(\theta_2))=(+1,+1)\} = \rho(\{(\sigma_1(\theta_1),\sigma_2(\theta_2))=(-1,-1)\} = \frac{\rho}{2}\sin^2\frac{\theta_2-\theta_1}{2} \tag{5}$$

$$\rho(\{(\sigma_1(\theta_1),\sigma_2(\theta_2))=(+1,-1)\} = \rho(\{(\sigma_1(\theta_1),\sigma_2(\theta_2))=(-1,+1)\} = \frac{\rho}{2}\cos^2\frac{\theta_2-\theta_1}{2} \tag{6}$$

where ρ would be the unconditionnal persistence of the EPR system.

3.2 Retro-Information transfer with asymmetric persistence

An interesting case is when the persistence is asymmetric. For example we could imagine an hypothetical EPR system where the persistence distribution on state subset $\{\sigma_1(\theta_1)=+1\}$ would actually vary in function of settings on particle 2 device. For example one could imagine an EPR system created in anisotropic conditions such that $\mathrm{E}(\sigma_i(\theta)) \neq 0$ for some θ.

We assume that there exist at least two setting positions, setting 0 and setting 1, on device 2 such the following holds.

- The persistence of state subset $\{\sigma_1(\theta_1)=+1\}$ in setting 0, is $\rho_0(+1)$;
- the persistence of state subset $\{\sigma_1(\theta_1)=+1\}$ in setting 1, is $\rho_1(+1)$;
- $\rho_0(+1) \neq \rho_1(+1)$.

Let $\rho_0(-1)$ and $\rho_1(-1)$ respectively denotes the persistence of state subset $\{\sigma_1(\theta_1) = -1\}$ when setting are respectively 0 and 1.

Due to the persistence conservation rule we have $\rho_0(+1)+\rho_0(-1) = \rho_1(+1)+\rho_1(-1) = \rho$, where ρ is the global persistence of the EPR system. There exists (p_0, q_0) and (p_1, q_1) such that

$$\left\{ \begin{array}{rcl} \rho_0(0) & = & \rho p_0 \\ \rho_0(1) & = & \rho q_0 \ , \end{array} \right. \tag{7}$$

and

$$\left\{ \begin{array}{rcl} \rho_1(0) & = & \rho p_1 \\ \rho_1(1) & = & \rho q_1 \ . \end{array} \right. \tag{8}$$

and $p_0 + q_0 = p_1 + q_1 = 1$ and $p_0 \neq p_1$. It comes that

$$\Pr\{\sigma_1(\theta_1) = +1 \mid \text{setting } 0\} = p_0 \tag{9}$$

$$\Pr\{\sigma_1(\theta_1) = +1 \mid \text{setting } 1\} = p_1 \tag{10}$$

$$\tag{11}$$

Therefore information transfers are possible from device 2 to device 1.

3.3 Retro-information source, retro-channel capacity

In the sequel, we call active device, the device on particle 1, and passive device, the device on particle 2.

We can build the devices such that the difference of space-time location between active and passive devices can be any arbitrary vector. In particular the measurement and setting on active device could occurs in the future while measurement on passive device just occurs after source event S, see figure 4 for an illustration. Therefore an asymmetric EPR pair system can be a source of retro-information.

Since the angle setting on active device can alternate between two values, we have created a binary retro-information source. Using Shannon theory, the capacity of the retro-channel $C^S(x)$ per EPR pair when the proportion of angles at position 0 is x is given by

$$C^S(x) = h(xp_0 + (1-x)p_1) - xh(p_0) - (1-x)h(p_1)$$

with $h(y) = -y\log_2 y + (1-y)\log_2(1-y)$. The optimal value is given by $x = \frac{p_1 - \alpha q_1}{(q_0 - q_1)\alpha - (p_0 - p_1)}$ with $\alpha = (p_0^{p_0} q_0^{q_0} p_1^{-p_1} q_1^{-q_1})^{1/(p_0 - p_1)}$.

With "quasi-symmetric" EPR systems, when $p_0 = \frac{1}{2} - \delta p_0$ and $p_1 = \frac{1}{2} + \delta p_1$ with $|\delta p_0| + |\delta p_1| \ll 1$. In this case, the maximum capacity C^S per EPR pair will be

$$C^S \approx \frac{(\delta p_0 + \delta p_1)^2}{2\log 2} \ . \tag{12}$$

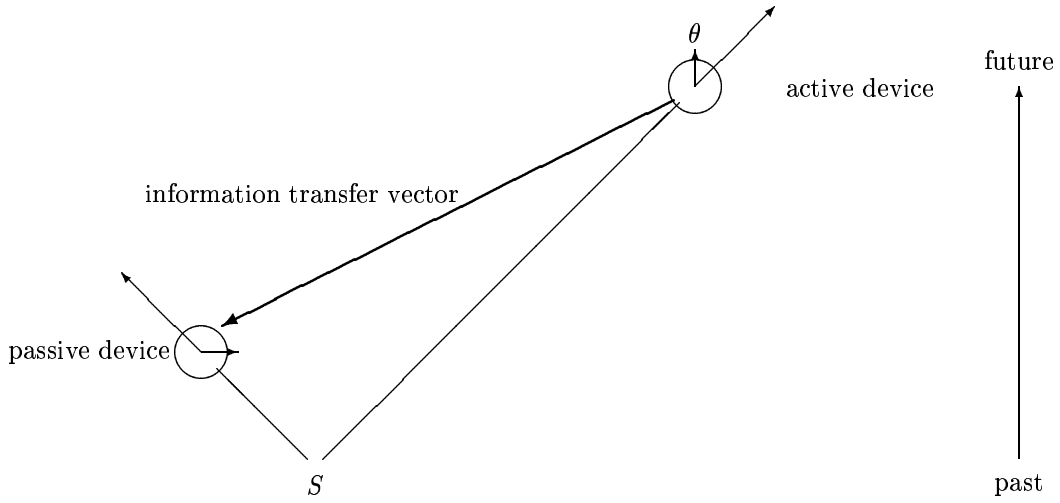

Figure 4: An hypothetical retro-information system build on EPR system

4 Information time paradoxes, paradoxal coupling

4.1 A Binary Oracle and the time paradox

Without loss of generality we assume $p_0 < p_1$. We set a real number μ strictly between p_0 and p_1.

One consider n EPR pairs of particles, respectively sourced on n "simultaneous" events: S_1, $S_2, \ldots, S_n$. We denote σ_i the spin measure on passive device i. We denote M_n the quantity $\sum_{i=1}^{i=n} \frac{\sigma_i+1}{2}$. When $M_n < \mu n$ we say that the oracle answer is "0", when otherwise $M_n \geq \mu n$, we say that the oracle answer is "1".

Theorem 3 *When settings on active devices are all in position x we have*

$$Pr\{M_n \geq \mu n\} = \sum_{k \geq \mu n}^{k=n} \binom{n}{k} p_x^k q_x^{n-k} \tag{13}$$

Corollary 1 *When $n \to \infty$, on setting 0 on all active devices $Pr\{M_n \geq \mu n\}$ tends to zero exponentially fast. On setting 1, $Pr\{M_n < \mu n\}$ tends to zero. In other words the oracle answer predicts the setting position with an exponentialy small error.*

Proof Using the asymptotic estimate over binomial factors:

$$\binom{n}{nx} p^\ell q^{n-\ell} \sim \frac{1}{\sqrt{2\pi n x(1-x)}} (f(x,p,q))^n \ , \tag{14}$$

with $f(x,p,q) = \frac{q^{1-x}p^x}{(1-x)^{1-x}x^x}$. We obtain:

- When all angle settings on the active devices are in position 0:

$$\Pr\{M_n \geq \mu n\} \sim \frac{1}{\sqrt{2\pi n\mu(1-\mu)}} \frac{1}{1-\frac{(1-\mu)p_0}{\mu q_0}} (f(\mu, p_0, q_0))^n$$

- When all angle settings are in position 1:

$$\Pr\{M_n < \mu n\} \sim \frac{1}{\sqrt{2\pi n\mu(1-\mu)}} \frac{1}{\frac{(1-\mu)p_1}{\mu q_1}-1} (f(\mu, p_1, q_1))^n \ ;$$

■

Of course the binary oracle gives rise to the following paradox. Nothing prevents a fast and unscrupulous operator who, after reading the oracle answer on passive devices will do all his best to contradict it on the active devices. To this end it would run fast enough to pass the flow of sister particles and change the settings on active devices before measurements occur. If the oracle said "0", then the operator sets all settings at position 1, if the answer was "1", then the operator sets settings at position 0, in both case contradicting the oracle prediction. We say that the oracle is in contrarious coupling.

This apparent unconsistency will be called an *information time paradox* since the operator strives contradicting a prediction. The consequence should be a kind of probability inversion and needs to be carefully analyzed. To this end we makes use of micro-branching universe hypothesis. To simplify the analysis we assume that when one EPR pair rollbacks, all the n EPR pairs rollback too.

Theorem 4 (Contrarious coupling) *The probabilities that the binary oracle with contrarious coupling answers 0 or 1 are respectively* $D_n(\mu)/(D_n(\mu)+U_n(\mu))$ *and* $U_n(\mu)/(D_n(\mu)+U_n(\mu))$, *with*

$$\left\{ \begin{array}{rcl} D_n(\mu) & = & \sum_{k\leq \mu n} \binom{n}{k} p_1^k q_1^{n-k} \\ U_n(\mu) & = & \sum_{k>\mu n} \binom{n}{k} p_0^k q_0^{n-k} \end{array} \right. \tag{15}$$

Proof: When $M_n = k$ with $k < \mu n$ all angle positions on active device are set in position 1 corresponding to persistence probability $\rho_1(1)$ for the EPR with answer 1 and $\rho_1(0)$ for the EPR pairs with answer 0. Therefore the persistence probability of the set of event $M_n = k$ will be $\binom{n}{k}(\rho_1(1))^k(\rho_1(0))^{n-k}$. When on the contrary $k \geq \mu n$ the angles are set in position 0, the event persistence probability becomes $\binom{n}{k}(\rho_0(1))^k(\rho_0(0))^{n-k}$.

Therefore the persistence of the set of events corresponding to the oracle answer equal to 0 is $\rho^n D_n(\mu)$ and the persistence of the complementary set of events corresponding to the answer 1 is $\rho^n U_n(\mu)$. The persistence of the whole oracle system is $(D_n(\mu)+U_n(\mu))\rho^n$ and the definitive answers probability are given by application of identity (4). ■

Corollary 2 *There is a value* μ_0 *such that when* $n \to \infty$, *if* $\mu < \mu_0$, *then the oracle in contrarious coupling answers 1 with high probability, otherwise when* $\mu > \mu_0$, *the oracle answers 0 with high probability.*

Proof Using the asymptotic estimate over the binomial expressions (14), it comes that the distribution of M_n/n is sharp around μ when n increases. Comparing the respective weight of $D_n(\mu)$ and $U_n(\mu)$ when n tends to infinity, we find a critical value μ_0 such that when $\mu > \mu_0$ the definitive probability of having oracle's answer equal to 0, *i.e* when $M_n < \mu n$, tends to be preponderant when n increases, and when $\mu < \mu_0$ the probability of having oracle's answer equal to 1 tends to be preponderant. We have $(\frac{p_0}{p_1})^{1-\mu_0}(\frac{q_0}{q_1})^{\mu_0} = 1$.

Remark: If instead of an contrarious operator we have a zealous operator which strives to have prediction realized whatever the prediction is: *i.e.* always switches the angle to predicted position. In this case the pair of values of (p_0, q_0) and (p_1, q_1) are switched in the expressions of $D_n(\mu)$ and $U_n(\mu)$. When n increases both expressions tend to 1. In other words, it turns out that the probabilities of oracle answer equal to 0 or 1 both tend to 1/2. We display the distribution of M_n when $n = 100$, $p_0 = 1/3$, $p_1 = 2/3$, $\mu = 1/2$: zealous coupling on figure 5 and contrarious coupling on figure 6.

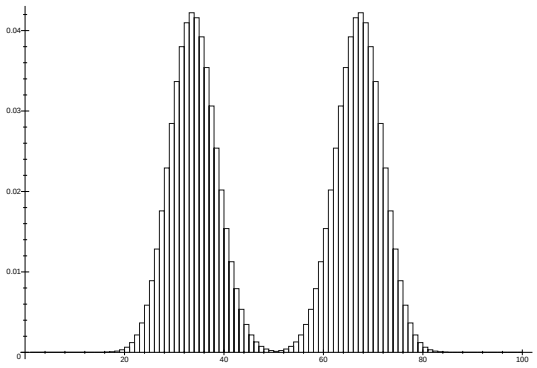

Figure 5: Oracle answer with zealous coupling

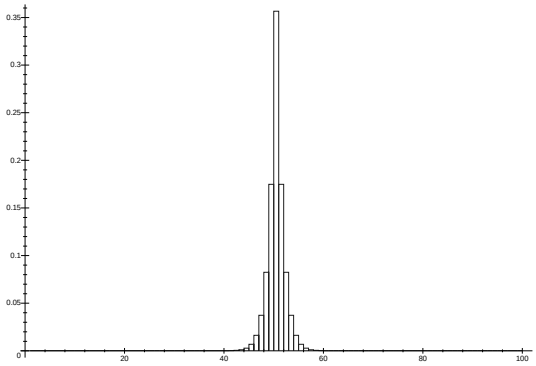

Figure 6: Oracle answer with contrarious coupling

4.2 Retro channel capacity with general paradoxal coupling

One unexpected by-product of the previous analysis is that paradoxal coupling can very significantly improve the capacity bound of retro-channel. Dopping channel capacity consists into introducing a paradoxal coupling between passive and active devices. To be precise we assume that the result of measurement passive devices are provided to active device *via* a reliable channel.

As before we consider n pairs of passive and active devices. We denote Y_n the code-word made of the spin measurements on passive devices. We call R_n the code-word made of the setting on active devices. We assume that both R_n and Y_n are binary words over alphabet $\{0,1\}$ (it means that the ith digit of Y_n is $\frac{\sigma_i+1}{2}$, if we refer to EPR spin terminology).

Information theory and Shannon theorem apply to retro-channel and in absence of paradoxal coupling the channel capacity C_n^S is

$$\frac{C_n^S}{n} = \max_{x\in[0,1]}\{h(xp_0+(1-x)p_1) - xh(p_0) - (1-x)h(p_1)\} \tag{16}$$

with $h(y) = -y \log y - (1-y)\log(1-y)$.

When we are in infinitesimal conditions: $p_1 = p + \delta p_1$ and $p_0 = p + \delta p_0$ with $|\delta p_0| + |\delta p_1| \ll 1$, we get

$$\frac{C_n^S}{n} \sim \frac{(\delta p_1 - \delta p_0)^2}{4pq} \tag{17}$$

We will show the following theorem.

Theorem 5 *In presence of paradoxal coupling the channel capacity C_n^D satisfies:*

$$\frac{C_n^D}{n} \geq \min\{\log 2, \log \frac{\max_i\{p_i\} + \max_i\{q_i\}}{\min_i\{p_i\} + \min_i\{q_i\}}\} \tag{18}$$

Remark: The above theorem proves that $C_n^D \ll C_n^S$. Indeed in infinitesimal conditions we get

$$C_n^D > 2|\delta p_1 - \delta p_0| + O((\delta p_1 - \delta p_0)^2)$$

which is significantly larger than the estimate in (17).

Moreover, in the symmetric case where $p_0 > q_0$ and $q_1 = p_0 = p$ and $p_1 = q_0 = q$,we get

$$\begin{array}{rcl} C_n^S & = & n(\log 2 + p\log p + q \log q) \\ C_n^D & = & n \min\{\log 2, \log \frac{p}{q}\} \end{array}$$

Proof: If we don't have pradoxal coupling, then the maximum channel capacity is obtained by the classical Shannon formula

$$C_n = \max_{R_n}(H(Y_n) - H(Y_n|R_n)) \tag{19}$$

where $H(.)$ denotes the entropy of a random variable. This formula translates into (16) since it can be proved that the maximum is attained for a Bernoulli source model of R_n where the 0 are chosen with probability x, and the 1 with probability $1-x$.

In case of paradoxal coupling we introduce a coupling where R_n is function of both the codeword transmitted X_n and the codeword received Y_n: $R_n = R_n(X_n, Y_n)$. In this case

$$C_n^D = \max_{R_n, X_n}(H(Y_n) - H(Y_n|X_n)) \ . \tag{20}$$

The case without paradoxal coupling is equivalent to have setting codeword R_n as function only of X_n, since maximum is attained when $R_n = X_n$ we fall back on (19).

In the sequel we will show that there exists a specific coupling function $R_n(X_n, Y_n)$ which yields formula (18). We assume that $p_0 > p_1$; therefore $q_0 < q_1$. Quantity $\overline{Y_n}$ denotes the opposite codeword of Y_n, *i.e.* the codeword made by exchanging 0 and 1 in Y_n. When $R_n = Y_n$ we say that the setting are zealous and when $R_n = \overline{Y_n}$ we say that the setting are contrarious.

We assume a certain integer $k \leq n$ which we will precise later. The paradoxal coupling is the following:

1. when X_n and Y_n agree on the k first characters then $R_n = Y_n$, *i.e.* the setting are zealous;
2. otherwise $R_n = \overline{Y_n}$, *i.e.* the setting are contrarious.

The goal is to find an estimate of $H(Y_n|X_n)$. We call $D_k(X_n)$ the set of codewords which agree with X_n on the k first characters, and $\overline{D_k}(X_n)$ is the complementary set. The persistence of set $D_k(X_n)$, $\rho(D_k(X_n)$ is

$$\frac{\rho(D_k(X_n))}{\rho^n} = p_k(X_n)(p_0 + q_1)^{n-k} \tag{21}$$

where $p_k(X_n)$ is the factor obtained by replacing each 0 by p_0 and each 1 by q_1 in the first k characters of X_n.

Conversely

$$\frac{\rho(\overline{D_k}(X_n))}{\rho^n} = (p_1 + q_0)^n - \overline{p_k}(X_n)(p_1 + q_0)^{n-k} \tag{22}$$

where $\overline{p_k}(X_n)$ is the factor obtained by replacing each 0 by p_1 and each 1 by q_0 in the first k characters of X_n. If we assume a Bernoulli model for X_n with probability x for the 0's and probability $1 - x$ for the 1's, almost surely when $k \to \infty$:

$$\log p_k(X_n) \sim (x \log p_1 + (1 - x) \log q_0)k \tag{23}$$

The game consists into selecting k such that

$$\lim_{n\to\infty} \frac{\rho(\overline{D_k}(X_n))}{\rho(D_k(X_n))} = 0 \tag{24}$$

In this case the random variable Y_n conditionned by a fixed X_n will tends to have its support in $D_k(X_n)$ following on the last $n-k$ digit a Bernoulli distribution with probability $\frac{p_1}{p_1+q_0}$ for the 0's, and probability $\frac{q_0}{p_1+q_0}$ for the 1's. Consequently we will get

$$\frac{H(Y_n|X_n)}{n} \sim \frac{n-k}{n} h(\frac{p_1}{p_1 + q_0}) \tag{25}$$

Similarly, condition (24) would imply:

$$H(Y_n) = h(x)k + H(Y_n|X_n) \tag{26}$$

which yields a channel capacity $C_n^D(x) = h(x)k$.

In order to have condition (24) satisfied we need

$$(x \log p_1 + (1 - x) \log q_0)k + (n - k) \log(p_1 + q_0) > n \log(p_0 + q_1)$$

therefore condition (24) holds for all k such that

$$\frac{k}{n} < \frac{-1}{x\log p_1 + (1-x)\log q_0 - \log(p_1+q_0)} \log \frac{p_1+q_0}{p_0+q_1} \tag{27}$$

In the case the right hand-side of above (27) is not greater than 1, since nevertheless we must have $k \leq n$ we can plug the upper bound in the expression of $C_n^D(x)$. It comes that

$$\begin{aligned} C_n^D &\geq \max_{x\in[0,1]} C_n^D(x) && (28)\\ &= n \max_{x\in[0,1]} \left\{ \frac{-h(x)}{x\log p_1 + (1-x)\log q_0 - \log(p_1+q_0)} \log \frac{p_1+q_0}{p_0+q_1} \right\} && (29)\\ &= n \log \frac{p_1+q_0}{p_0+q_1} && (30) \end{aligned}$$

noticing that the maximum is attained for $x = \frac{p_1}{p_1+q_0}$. ■

One immediately notices that we would obtain capacity greater than $n\log 2$ if $\frac{p_1+q_0}{p_0+q_1} > 2$ which should not be possible since the codeword Y_n has only n digits. In fact this apparent unconsistency is only the consequence of side effect when the strict application of (27) would imply $k > n$. The strict obedience to the condition $k \leq n$ will restore $C_n^D \leq n\log 2$. Therefore we have

$$C_n^D \geq n \min\{\log 2, \log \frac{p_1+q_0}{p_0+q_1}\}$$

The case $\frac{p_1+q_0}{p_0+q_1} > 2$ is nevertheless interesting because it allows to limit paradoxal coupling on a fraction of active devices. For example Let ℓ be the integer part of $\frac{n}{\log 2}\log\frac{p_1+q_0}{p_0+q_1}$, the coupling $R_n(Y_n, X_n)$ can be the following:

1. The ℓ first characters of R_n are the ℓ first characters of Y_n when $Y_n = X_n$, otherwise the ℓ first characters of R_n equal the ℓ first characters of $\overline{Y_n}$.
2. the $n-\ell$ last characters of R_n do not depend on Y_n (*e.g* equal to the $n-\ell$ last characters of X_n or are constant).

This example is interesting because it illustrates the fact that the coupling can affect the information received from active devices with indifferent setting. Of course this property and the previous theorem is greatly dependent on the assumption that all devices rolbacks as soon one device rollbacks.

When the state measurement on each passive device is V-ary instead of binary, *i.e.* can have V values instead of 2. If the set of setting $\mathcal{S}$ on each active device is of arbitrary cardinality, instead of having only two positions we would have

$$C_n^D \geq \log \frac{\sum_{s\in\mathcal{S}} \max_i p_s(i)}{\sum_{s\in\mathcal{S}} \min_i p_s(i)}$$

where I varies between 1 and V and $p_s(i)$ denotes the probability of state i under setting s. Of course the C_n^D is limited by $\log V$.

5 Retro-computing

5.1 A Trace theorem

In this section we try to generalize the idea of paradoxal coupling with arbitrary operators, indifferently contrarious or zealous. Let a fixed integer k. We consider kn EPR pairs divided into k identical classes of size n, corresponding to k binary oracles indexed from 1 to k.

We consider a memoryless Markov process with 2^k states defined by a $2^k \times 2^k$ transition matrix $\mathbf{P} = [p_{ij}]$. Each state corresponds to any possible vector of answers made by the k simultaneous binary oracles. We suppose that the transition matrix has a non identically null diagonal, *i.e.* there exists at least one coefficient p_{ii} which is non zero.

We assume that when the operator reads the vector of answers from the k binary oracle, identifying the state i of the system, it has a probability p_{ij} to switch the angle position on active devices to those corresponding to the vector of answers of state j. The transition can be external to oracle device, *i.e.* determined by a random trial done by the operator on his path from passive device to active device. The transition can be internal to oracle, *i.e.* commanded by the very values of the kn device answers, for example the last device of the ith oracle gives the ith digits of the switched state.

We denote $\mathrm{Tr}(\mathbf{P}) = \sum_{j=1}^{j=2^k} p_{jj}$, the trace of matrix $\mathbf{P}$.

Theorem 6 (Trace theorem) *When the Markov matrix p_{ij} has a non identically null diagonal, then, when n increases, the definitive state probability distribution of state i tends to be $p_{ii}(\mathrm{Tr}(\mathbf{P}))^{-1}$.*

Proof Let $g(\mu) < \rho$, ρ being the persistence of the EPR pair $\rho = \rho_0(0) + \rho_0(1) = \rho_1(0) + \rho_1(1)$ such that $g(\mu)^n$ is greater than both $\rho^n D_n(\mu)$ and $\rho^n U_n(\mu)$ as described in theorem 4. For example $g(\mu) > \max\{f(\mu, \rho_1(0), \rho_1(1)), f(\mu, \rho_0(0), \rho_0(1))$.

Let us select a state i with binary expresion $i_1 \ldots i_k$. The persistence probability of state i has the approximate expression $p_{ii}\rho^{kn} + \sum_{j \neq i} p_{ij}\rho_{ij}$ where each of the terms ρ_{ij} are smaller than $\rho^{(k-\ell)n} g(\mu)^{\ell n}$ where ℓ is the number of different digits between i and j (Hamming distance). It is clear that the probability persistence tends to be equivalent to $p_{ii}\rho^{kn}$. Using formula (4) the definitive state distribution comes. ■

Remark: We notice that with the contrarious operator paradox analyzed in section 4.1, we have $k = 1$ but the Markov matrix has an identically null diagonal, thus the theorem conditions do not hold. But with the zealous operator paradox the Markov matrix is just identity and we can use the trace theorem to obtain that the definitive state distribution is uniform.

5.2 Retro-computing, or EPR versus RSA

An interesting objective is to use retro-information in order to improve computation. It could be affordable even if retro-information were possible only on short space-time distance. In

particular retro-computing could be used in order to make possible some of the computations which are not tractable with classical means. For example the problem of factorizing integer is one of such computations. One way of using retro-information could be to let a classical computer to factorize one given long integer Z over one billion of years needed by the present technology and then to send back the results obtained one billion year before in order to make the computation looking instantaneous. But this should be a kind of cheating since the computation would have needed a total cost of one billion years of classical computation plus one billion year of retro-information: therefore no complexity improvement. Of course our aim here is to really improve cost complexity in order to really obtain computations with much less effort.

Let us take as an example the difficult problem of integer factorization whose difficulty is the basis of RSA crypto-algorithms [6]. Let X be a number of k binary digits. To simplify we assume that X is the product of two prime numbers Y and Z.

We assume that nk EPR pairs are available and we realize a paradoxal coupling as described in the section devoted to Trace Theorem. To simplify we assume that each state is a binary sequence of length k which represents a pair of numbers (y, z) with product less than 2^k. In fact we should need $n \log_2 k$ more EPR pairs, to indicate the length of the binary representation of the first number in the couple, but for simplifying the analysis we omit this part.

The coupling will be the following $p_{(y,z)(y,z)} = 1$ when $yz = X$ and $p_{(y,z)(\overline{y},\overline{z})} = 1$ when otherwise $yz \neq X$ ($\overline{x}$ indicates the number which has opposit digits of x, or $\overline{x} = 2^k - 1 - x$). In other words the operator will take the answers of the oracle, make the product and be zealous when the product equal X, or be contrarious otherwise.

The markov matrix has only two non zero coefficient on its diagonal, corresponding to (Y, Z) and (Z, Y), and the trace is equal to 2. According to Trace theorem, when $n \to \infty$ the distribution probability of oracle answer exponentially converges to be $\frac{1}{2}$ on (Y, Z) and $\frac{1}{2}$ on (Z, X). In other words the retro-computing gives the right factorization of a number X, *i.e.* brakes RSA crypto-encoding, in $O(1)$ time and $O(n \log X)$ devices with probability $1 - \varepsilon_n$, with ε_n exponentially small. This is an example of retro-computing on functions which are hardly reversible with "usual" computing. In passing one could conjecture that most of functions used in cryptography would be easy to break in presence of retro-computing facilities.

6 Retro-programming

6.1 Retro-loop and retro-memory

The art of programming retro-computer is not a difficult one. We only introduce a new primitive called *retro-loop* attached to the concept of retro-memories. Retro-memories could be made of EPR pairs in cascade. The retro-loop uses retro-memories with paradoxal coupling.

A retro-loop procedure can be initiated as follows:

```
Retro-loop(passive_variables,coupling_boolean)
{
  loop_instructions
}
```

where

- `passive_variables` is a list of variables read on the passive devices of the retro-memories which will be affected by a paradoxal coupling on their active devices. In passing this declaration fixes the volume of retro-memories needed by the retro-loop.
- `coupling_boolean` is the name of the boolean variable whose value at the end of the retro-loop determines the setting of the active devices to be either zealous (when `coupling-boolean=true`) or contrarious (when `coupling_boolean=false`).
- the `loop_instructions` contain the instructions needed to compute the value of the `coupling_boolean` which will be used for the active device setting. Of course the `loop_instructions` must be performed between the read of the passive devices and the setting of the active devices of the retro- memories involved by the loop.

As an example we gain consider the program for integer factorization and we write it with a retro-loop.

```
Procedure factorize(Z);
Int Z;
{
  Retro-loop((int X, Y), P)
  {
    P=false;
    If (X*Y=Z) then P=true;
  }
  if (X*Y=Z) then return(X,Y)
    else print ("Z assumed prime")
}
```

Remark: as we know retro-information is subject to random alteration. Therefore a retro-program provides a correct answer with probability $1 - \varepsilon$, for ε arbitrarily small, provided that enough redundancy in retro-memories are available.

6.2 Playing with retro-programming: the multiple door problem

Assume a room with n doors, all closed. Behind each door there is a certain amount of money, not necessarily the same. With help of retro-information how many doors need to be opened in order to know a door which hides:

1. an amount of money equal to 10 FF;
2. the largest amount of money;
3. the kth largest amount of money.

Surprisingly the answers of those questions are the following:

1. one door plus one memory unit and $O(\log n)$ retro-memory units;
2. two doors plus two memory units and $O(\log n)$ retro-memory units;
3. $k+1$ doors plus $k+1$ memory units and $O((k+1)\log n)$ retro-memory units.

Therefore retro-information is in direct continuation of the improvement brought my quantum computers which need to open $O(\sqrt{n})$ doors with binary *qubits* to answer the first question, n being the number of doors in the room [7] or much less with general quantum computers [10]. To answer these questions with classic information theory and algorithmic would need:

1. $n/2$ doors in average plus one memory unit;
2. n doors plus two memory units;
3. n doors plus $k+1$ memory units.

Below we provide the retro-programs to answer these questions. If `X` is an integer we call `G(X)` the amount of money hidden behind door number `X`.

6.2.1 Ten franc finding

Below is the retro-program to answer question 1:

```
Procedure find_ten_FF;
{
  int K;
    retro-loop(int X, I)
    {
      I=false;
      K=G(X);
      If (K=10FF) then I=true;
     }
     if (K=10FF) then return(X)
       else print("no door hides amount 10FF");
}
```

Notice that in the living time path only one value of function `G(.)` is computed: only one door is opened. Notice that $O(\log n)$ retro-memories are needed in order to define `X` which needs $\log_2 n$ digits.

6.2.2 Maximum finding

Below is the retro-program to answer question 2:

```
Procedure find_maximum;
{
  int K1, K2;
  retro-loop(int X1,I1)
  {
    I1=false;
    K1=G(X1);
    Retro-loop(int X2,I2);
    {
      I2=false;
      K2=G(X2);
      If (K2 > K1) then I2=true;
    }
    if (K2<=K1) then I1=true;
  }
  return(X1,K1);
}
```

Notice that after inner retro-loop, there is an inversion of the "exit" condition attached to the external loop. This means that

1. inner retro-loop strives to find an amount `K2` greater than `K1`;
2. external retro-loop strives to find `K1` such that such `K2` does not exist: `K1` is the maximum of function `G(.)`.

One interesting application is the solution of the traveler problem: finding on a graph of n vertices the path which visits all vertices with the minimum cost. This problem is known to be NP hard. With retro-computing it is just polynomial in n.

```
Procedure min_path
{
  int K1, K2;
  retro-loop(path1,I1)
  {
    I1=false;
    K1=cost(path1);
    Retro-loop(path2,I2)
    {
      I2=false;
      K2=cost(path2);
```

```
      If (K2<K1) then I2=true;
    }
    if (K2>=K1) then I1=true;
  }
  return(path1);
}
```

Since a path is a specific permutation over the n vertices, and there at most $n!$ of such permutations, the total number of retro-memories needed is $O(n \log n)$ for a $O(n)$ computation time (computing the cost of the two paths).

6.2.3 Order finding

Below is the program which finds the m-th amount in decreasing order.

```
Procedure find_order(m);
int m;
{
  int K0, K1, .., Km;
  retro-loop(int X1, .., Xm, I1)
  {
    I1=false;
    K1=G(X1); .. Km=G(Xm);
    Retro-loop(int X0,I2)
    {
      I2=false;
      K0=G(X0);
      If (K0>K1) or .. or (K0>Km) then I2=true;
    }
    if (K0<=K1) and .. and (K0<=Km) then I1=true;
  }
  Ki=min(K1,..,Km);
  Return(Ki,Xi);
}
```

Remark: the procedure "min" can be performed with retro-programming too, to the additional cost of $O(\log m)$ retro-memories and 2 comparisons. Therefore the total cost of the program is $O((m+1)\log n)$.

Conjecture 1 *We need to open at least* $\min\{m, n-m\}+1$ *doors to find the* m*th order amount with retro-programming.*

This is an example where retro-computing cannot make polynomial all NP hard problems. Suppose that we have 2^n doors all indexed by a binary index of length n. The problem which consists into finding the 2^{n-1}th order amount would need 2^{n-1} to be opened.

7 Conclusion

The main question which naturally remains after this note is: does our universe be a Wheeler Feynman universe which allows retro-information? Of course we don't have even the beginning of the answer. It is clear that the standard model would need to be significantly amended, in particular about the product form state expressions (see appendix). Our only aim with respect to pure physics is to say that unitarity and causality would probably survive in an universe with reto-information. Moreover retro-information might offer very interesting perspectives in information theory and computer technology. Therefore it might be worthy to keep this in mind for future investigations, just in case...

References

[1] A. Aspect, "Trois tests expérimentaux des inégalités de Bell par mesure de corrélation de polarisation de photons," Thèse, Université Paris-Sud, 1983.

[2] A. Aspect, J. Dalibard, G. Roger, "Experimental Test of Bell's Inequalities Using Time-varying Analyzers," in *Physical Review Letters*, 49, 1804-1807, 1982.

[3] J.S. Bell, "On the Einstein-Podolsky-Rosen Paradox," in *Physics*, 1, 195-200, 1964.

[4] A. Einstein, B. Podolsky, N. Rosen, "Can Quantum Mechanical Description of Physical Reality be considered complete?" in *Physical Review*, 47, 777-780, 1935.

[5] J.A. Wheeler, R.P. Feynman, "Interaction with the Absorber as the Mechanism of Radiation," in *Reviews of Modern Physics*, 17, 157-181, 1945.

[6] R.L. Rivest, A. Shamir, L. Adleman, "A method for obtaining digital signatures and public-key cryptosystems," *Commun. ACM*, vol 21, pp. 120-126, 1978.

[7] L. Grover, "Quantum mechanics helps in searching for a needle in a haystack," in *Physical Review Letters*, vol. 79, no 2, pp. 325-328, 1997.

[8] C. Bennett, "Quantum information and computation," in *Physics Today*, vol 48, no 10, pp. 24-30, 1995.

[9] N. Gershenfeld, I. Chuang, "Bulk spin-resonance quantum computations," in *Science*, vol 275, pp. 350-356, 1997.

[10] P. Shor, "Polynomial-time algorithms for prime factorization and discrete logarithm on quantum computer," in proc. *35th Annual Symposium on Foundation of Computer Science*, Santa Fe, IEEE Computer society Press, pp. 124-139, 1994.

Appendix

7.1 Spin wave functions

The wave function of the spin of kind $\frac{1}{2}$ measured on axe $\mathcal{A}_\theta$ is distributed over the two state functions $\mathbf{u}(\theta)$ and $\mathbf{d}(\theta)$ corresponding respectively to spin value $+1$ and -1. The correspondance between state functions relative to two axes $\mathcal{A}_{\theta_1}$ and $\mathcal{A}_{\theta_2}$ satisfies the classical rotation identity:

$$\left\{ \begin{array}{rcl} \mathbf{u}(\theta_2) & = & \cos\frac{\theta_1-\theta_2}{2}\mathbf{u}(\theta_1) + \sin\frac{\theta_1-\theta_2}{2}\mathbf{d}(\theta_1) \\ \mathbf{d}(\theta_2) & = & -\sin\frac{\theta_1-\theta_2}{2}\mathbf{u}(\theta_1) + \cos\frac{\theta_1-\theta_2}{2}\mathbf{d}(\theta_1) \end{array} \right. \tag{31}$$

7.2 Symmetric EPR

The wave function ψ of the symmetrical EPR spin system $(\sigma_1(\theta), \sigma_2($
$th))$ *modulo* a proportional factor, satisfies the identity:

$$\psi = \mathbf{u}_1(\theta)\mathbf{d}_2(\theta) - \mathbf{d}_1(\theta)\mathbf{u}_2(\theta) \tag{32}$$

where $\mathbf{u}_x(\theta)$ and $\mathbf{d}_x(\theta)$, for $x \in \{1,2\}$ denote the state function applied to spin $\sigma_x(\theta)$.

Identity (32) is invariant when angle θ varies. When the angle of measure on spin 1 and 2 are different, namely θ_1 and θ_2, the state distribution classically becomes:

$$\begin{array}{rcl} \psi & = & \mathbf{u}_1(\theta_1)\mathbf{d}_2(\theta_1) - \mathbf{d}_1(\theta_1)\mathbf{u}_2(\theta_1) \\ & = & \mathbf{u}_1(\theta_1)(\sin\frac{\theta_1-\theta_2}{2}\mathbf{u}_2(\theta_2) + \cos\frac{\theta_1-\theta_2}{2}\mathbf{d}_2(\theta_2)) - \\ & & -\mathbf{d}_1(\theta_1)(\cos\frac{\theta_1-\theta_2}{2}\mathbf{u}_2(\theta_2) - \sin\frac{\theta_1-\theta_2}{2}\mathbf{d}_2(\theta_2)) \\ & = & \sin\frac{\theta_1-\theta_2}{2}\mathbf{u}_1(\theta_1)\mathbf{u}_2(\theta_2) + \sin\frac{\theta_1-\theta_2}{2}\mathbf{d}_1(\theta_1)\mathbf{u}_2(\theta_2) + \\ & & + \cos\frac{\theta_1-\theta_2}{2}\mathbf{u}_1(\theta_1)\mathbf{d}_2(\theta_2) - \cos\frac{\theta_1-\theta_2}{2}\mathbf{d}_1(\theta_1)\mathbf{u}_2(\theta_2) \end{array} \tag{33}$$

The squares of the coefficient before each state function correspond to the probability presented in identity (3).

7.3 Asymmetric EPR spin systems

In this section we introduce a tentative physical model about asymmetric EPR spin system. We consider state functions of the form $(f,g)_x$ where f and g are either $\mathbf{u}(\theta)$ or $\mathbf{d}(\theta)$ and $x \in \{1,2\}$ is called "precedence". The state function f is related to spin 1 and g to pin 2.

We asume the following transition rules:

For precedence 1 and for any state function f:

$$\left\{ \begin{array}{rcl} (f,\mathbf{u}(\theta_2))_1 & = & \cos\frac{\theta_1-\theta_2}{2}(f,\mathbf{u}(\theta_1))_1 + \sin\frac{\theta_1-\theta_2}{2}(f,\mathbf{d}(\theta_1))_1 \\ (f,\mathbf{d}(\theta_2))_1 & = & -\sin\frac{\theta_1-\theta_2}{2}(f,\mathbf{u}(\theta_1))_1 + \cos\frac{\theta_1-\theta_2}{2}(f,\mathbf{d}(\theta_1))_1 \end{array} \right. \tag{34}$$

For precedence 1 and opposite spin functions:

$$\left\{ \begin{array}{rcl} (\mathbf{u}(\theta_2),\mathbf{d}(\theta_2))_1 & = & \cos\frac{\theta_1-\theta_2}{2}(\mathbf{u}(\theta_1),\mathbf{d}(\theta_1))_1 + \sin\frac{\theta_1-\theta_2}{2}(\mathbf{d}(\theta_1),\mathbf{u}(\theta_1))_1 \\ (\mathbf{d}(\theta_2),\mathbf{u}(\theta_2))_1 & = & -\sin\frac{\theta_1-\theta_2}{2}(\mathbf{u}(\theta_1),\mathbf{d}(\theta_1))_1 + \cos\frac{\theta_1-\theta_2}{2}(\mathbf{d}(\theta_1),\mathbf{u}(\theta_1))_1 \end{array} \right. \tag{35}$$

For precedence 2 and for any state function g

$$\left\{\begin{array}{lcl}(\mathbf{u}(\theta_2),g)_2 & = & \cos\frac{\theta_1-\theta_2}{2}(\mathbf{u}(\theta_1),g)_2+\sin\frac{\theta_1-\theta_2}{2}(\mathbf{d}(\theta_1,g)_2\\(\mathbf{d}(\theta_2),g)_2 & = & -\sin\frac{\theta_1-\theta_2}{2}(\mathbf{u}(\theta_1),g)_2+\cos\frac{\theta_1-\theta_2}{2}(\mathbf{d}(\theta_1,g)_2\end{array}\right. \tag{36}$$

For precedence 2 and for opposite spin state functions:

$$\left\{\begin{array}{lcl}(\mathbf{d}(\theta_2),\mathbf{u}(\theta_2))_2 & = & \cos\frac{\theta_1-\theta_2}{2}(\mathbf{d}(\theta_1),\mathbf{u}(\theta_1))_2+\sin\frac{\theta_1-\theta_2}{2}(\mathbf{u}(\theta_1),\mathbf{d}(\theta_1))_2\\(\mathbf{u}(\theta_2),\mathbf{d}(\theta_2))_2 & = & -\sin\frac{\theta_1-\theta_2}{2}(\mathbf{d}(\theta_1),\mathbf{u}(\theta_1))_2+\cos\frac{\theta_1-\theta_2}{2}(\mathbf{u}(\theta_1),\mathbf{d}(\theta_1))_2\end{array}\right. \tag{37}$$

For any precedence x and for identical spin state functions:

$$\left\{\begin{array}{lcl}(\mathbf{u}(\theta_2),\mathbf{u}(\theta_2))_x & = & \cos\frac{\theta_1-\theta_2}{2}(\mathbf{u}(\theta_1),\mathbf{u}(\theta_1))_x+\sin\frac{\theta_1-\theta_2}{2}(\mathbf{d}(\theta_1),\mathbf{d}(\theta_1))_x\\(\mathbf{d}(\theta_2),\mathbf{d}(\theta_2))_x & = & -\sin\frac{\theta_1-\theta_2}{2}(\mathbf{u}(\theta_1),\mathbf{u}(\theta_1))_x+\cos\frac{\theta_1-\theta_2}{2}(\mathbf{d}(\theta_1),\mathbf{d}(\theta_1))_x\end{array}\right. \tag{38}$$

An arbitrary asymmetric EPR system based on opposite spins has wave function of the form

$$\psi = x_1(\mathbf{u}(\theta),\mathbf{d}(\theta))_1+y_1(\mathbf{d}(\theta),\mathbf{u}(\theta))_1+x_2(\mathbf{u}(\theta),\mathbf{d}(\theta))_2+y_2(\mathbf{d}(\theta),\mathbf{u}(\theta))_2 \tag{39}$$

with $|x_1|^2+|y_1|^2+|x_2|^2+|y_2|^2=1$. On two wave functions ψ and $\psi' = x_1'(\mathbf{u}(\theta),\mathbf{d}(\theta))_1 + y_1'(\mathbf{d}(\theta),\mathbf{u}(\theta))_1+x_2'(\mathbf{u}(\theta),\mathbf{d}(\theta))_2+y_2'(\mathbf{d}(\theta),\mathbf{u}(\theta))_2$, we denote the hilbertian product $\langle\psi',\psi\rangle$:

$$\langle\psi',\psi\rangle = x_1'x_1^*+y_1'y_1^*+x_2'x_2^*+y_2'y_2^*\ , \tag{40}$$

where z^* denotes the conjugate number of a complex number z.

In particular we introduce the symmetrical wave functions:

$$s_1(\theta) = \frac{1}{\sqrt{2}}((\mathbf{u}(\theta),\mathbf{d}(\theta))_1+i(\mathbf{d}(\theta),\mathbf{u}(\theta))_1) \tag{41}$$

$$s_2(\theta) = \frac{1}{\sqrt{2}}((\mathbf{u}(\theta),\mathbf{d}(\theta))_2-i(\mathbf{d}(\theta),\mathbf{u}(\theta))_2) \tag{42}$$

For any θ_1, θ_2 and $x\in\{1,2\}$ we have $s_x(\theta_2)=e^{-i(\theta_2-\theta_1)/2}s_x(\theta_1)$.

It is clear that the $s_x(\theta)$ corresponds to the symmetrical EPR systems. When ψ corresponds to a symmetrical EPR system we have both $\langle\psi,s_1(\theta_1)\rangle\langle\psi^*,s_1(\theta_1)\rangle=0$ and $\langle\psi,s_2(\theta_1)\rangle\langle\psi^*,s_2(\theta_1)\rangle=0$.

Taking an arbitrary asymmetric EPR wave function ψ as defined in (39), the aim is to compute the probability distribution of $(\sigma_1(\theta_1),\sigma(\theta_2))$. We have the identities:

$$\left\{\begin{array}{lcl}\Pr\{(\sigma_1(\theta_1),\sigma(\theta_2))=(+1,+1)\} & = & |\langle\psi,(\mathbf{u}(\theta_1),\mathbf{u}(\theta_2))_1|^2+|\langle\psi,(\mathbf{u}(\theta_1),\mathbf{u}(\theta_2))_2|^2\\\Pr\{(\sigma_1(\theta_1),\sigma(\theta_2))=(+1,-1)\} & = & |\langle\psi,(\mathbf{u}(\theta_1),\mathbf{d}(\theta_2))_1|^2+|\langle\psi,(\mathbf{u}(\theta_1),\mathbf{d}(\theta_2))_2|^2\\\Pr\{(\sigma_1(\theta_1),\sigma(\theta_2))=(-1,+1)\} & = & |\langle\psi,(\mathbf{d}(\theta_1),\mathbf{u}(\theta_2))_1|^2+|\langle\psi,(\mathbf{d}(\theta_1),\mathbf{u}(\theta_2))_2|^2\\\Pr\{(\sigma_1(\theta_1),\sigma(\theta_2))=(-1,-1)\} & = & |\langle\psi,(\mathbf{d}(\theta_1),\mathbf{d}(\theta_2))_1|^2+|\langle\psi,(\mathbf{d}(\theta_1),\mathbf{d}(\theta_2))_2|^2\end{array}\right. \tag{43}$$

The game consists into expanding ψ on the state functions $(f,g)_x$, where $f \in \{\mathbf{u}(\theta_1), \mathbf{d}(\theta_1)\}$, $g \in \{\mathbf{u}(\theta_2), \mathbf{d}(\theta_2)\}$ and $x \in \{1,2\}$. This is trivially done via the use of the above correspondence tables. In particular it comes out that

$$\Pr\{\sigma_1(\theta_1) = +1\} = \frac{1}{2} + \frac{1}{2}\Re(\langle\psi, s_1(\theta_1)\rangle\langle\psi^*, s_1(\theta_1)\rangle) + \quad (44)$$

$$+\frac{1}{4}\Re((1 + e^{-i(\theta_2-\theta_1)/2})\langle\psi, s_2(\theta_1)\rangle\langle\psi^*, s_2(\theta_1)\rangle) \quad (45)$$

Quantity $\Re(z)$ denotes the real part of a complex number z. It is clear that when $\langle\psi, s_2(\theta_1)\rangle\langle\psi^*, s_2(\theta_1)\rangle \neq 0$, the above right-hand side varies in function of θ_2 and therefore retro-information can occur.

In *quasi*-symmetrical EPR system, *i.e.* when quantities $\langle\psi, s_x(\theta_1)\rangle\langle\psi^*, s_1(\theta_1)\rangle$ for $x \in \{1,2\}$ are very small compared to unity, we can use formula (12) to compute the maximum retro-channel binary capacity:

$$C^S \approx \frac{|\langle\psi, s_1(\theta_1)\rangle\langle\psi^*, s_1(\theta_1)\rangle|^2}{8\log 2} \ . \quad (46)$$

and, in case of paradoxal coupling

$$C^D \approx \frac{2|\langle\psi, s_1(\theta_1)\rangle\langle\psi^*, s_1(\theta_1)\rangle|}{\log 2} \ . \quad (47)$$

Unit´e de recherche INRIA Lorraine, Technopôle de Nancy-Brabois, Campus scientifique,
615 rue du Jardin Botanique, BP 101, 54600 VILLERS LÈS NANCY
Unit´e de recherche INRIA Rennes, Irisa, Campus universitaire de Beaulieu, 35042 RENNES Cedex
Unit´e de recherche INRIA Rhône-Alpes, 655, avenue de l'Europe, 38330 MONTBONNOT ST MARTIN
Unit´e de recherche INRIA Rocquencourt, Domaine de Voluceau, Rocquencourt, BP 105, 78153 LE CHESNAY Cedex
Unit´e de recherche INRIA Sophia-Antipolis, 2004 route des Lucioles, BP 93, 06902 SOPHIA-ANTIPOLIS Cedex

Éditeur
INRIA, Domaine de Voluceau, Rocquencourt, BP 105, 78153 LE CHESNAY Cedex (France)
http://www.inria.fr
ISSN 0249-6399