



# Capacity of Retro-Information Channels

Philippe Jacquet, Véronique Joly

► **To cite this version:**

Philippe Jacquet, Véronique Joly. Capacity of Retro-Information Channels. [Research Report] RR-3836, INRIA. 1999. <inria-00072821>

**HAL Id: inria-00072821**

**<https://hal.inria.fr/inria-00072821>**

Submitted on 24 May 2006

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

*Capacity of retro-information channels*

Philippe Jacquet, Véronique Joly

**N° 3836**

1999

THÈME 1

  
*Rapport  
de recherche*



## Capacity of retro-information channels

Philippe Jacquet\*, Véronique Joly†

Thème 1 — Réseaux et systèmes  
Projet HIPERCOM

Rapport de recherche n° 3836 — 1999 — 22 pages

**Abstract:** In this paper we investigate how retro-information is possible in a non unitary universe. In particular we give estimate of asymptotic capacity of retro-information channels in parallel or in series. These results are significantly divergent from classical information theory results and open interesting perspectives.

**Key-words:** Quantum physics, unitarity, retro-information, EPR spin systems, forward coupling functions, channel capacity.

*(Résumé : *tsvp*)*

\* Email: [philippe.jacquet@inria.fr](mailto:philippe.jacquet@inria.fr)

† Email: [joly@onera.fr](mailto:joly@onera.fr)

# Capacité des canaux de rétro-information

**Résumé :** Dans ce papier, nous examinons comment la rétro-information est rendue possible dans des univers non unitaire. En particulier nous donnons des estimations des capacités asymptotiques des canaux de rétro-information en parallèle ou en série. Les résultats présentent des divergences significatives par rapport aux résultats classiques de théorie de l'information et offrent des perspectives intéressantes.

**Mots-clé :** Physique quantique, unitarité, rétro-information, systèmes de spin EPR, fonctions de couplage avancé, capacité des canaux

# 1 Introduction

For more than fifty years the most prominent researchers in physics have looked in vain for a consistent theory that would unify the forces of the universe (gravitation, electromagnetism, *etc*). *Superstring* theory [4] is the last attempt to date, this (very) complicated theory however needs to be embedded in an universe with at least 26 dimensions to make unification. Numerical estimate based on superstring theory leads to gravitational constants  $10^{120}$  greater of what is observed. The quest for unification will surely continue again for several years, but many physicists support the assertion that such unification may simply not exist. Under this perspective the non unified model of universe would systematically present symmetry violations, gauge exception and in particular unitarity violations probably at (very) short scale.

We present the properties of information transfers based on the unitarity violations in the non-unifiable universe model. We will show that such transfers may occur backward in time and therefore may give rise to inedit combinations in information theory. We coin here the term *retro-information* to denote information transfer backward in time (at short scale). Conversely we can prove that retro-information implies unitarity violation. Since retro-information is based on quantum theory rigorous axiomatic, all paradoxical effects can be precisely analysed. In particular an interesting combination occurs where the retro-information channel is submitted to a forward coupling function: *i.e.* when the result of the transmission is made available to the transmitter *before* the transmission occurs. In this very case the capacity of the channel can be greatly improved from its estimate in classical Shannon theory [2].

The paper is organized as follow. Section 2 is devoted to a short survey of the physical background which allows to make unitarity violation in correspondence with retro-information transfers. In particular we prove the fundamental result that retro-information is equivalent to unitarity violation. Section 3 describes an hypothetical non-unitary EPR spin system which could act like a retro-information binary channel on arbitrary space-time vector. Section 4 addresses in details the results about retro-information channels in two cases: when the channels are in parallel and when they are in series.

## 2 Unitarity violations and retro-information

### 2.1 Overview of probability axiomatic in quantum theory

A classical physical experiment consists to measure the coordinates  $(x, y, z)$  of particle impact on a screen located at distance  $z$  from a source of particles. More generally, to characterize a measurement of a physical effect, one needs

1. the measurement setting, e.g. the distance  $z$  of the screen to the source  $O$ ;
2. the measurement value obtained under this setting, e.g. the coordinate  $(x, y, z)$  of the particle impact on the screen.

Let  $S$  be a set of measurement values of a physical effect under a given setting (for example a set of tuple  $(x, y, z)$  under the same  $z$  value). Let  $\psi$  be the wave function of the physical effect defined on the measurement set (for example the wave function of the particle defined on  $(x, y, z)$  coordinates). Let  $A$  be a subset of measurement values of the said physical effect (for example the sets of all coordinates  $(x, y, z)$  such that  $x > 0$ , in other words measurement on right half screens), for example see figure 1. It is classical in quantum theory that the conditional probability  $P(A|S)$  that the measured value belongs to  $A$ , given that it belongs to  $S$  is given by:

$$P(A|S) = \frac{\int_{A \cap S} \psi^* \times \psi}{\int_S \psi^* \times \psi}. \quad (1)$$

Quantity  $\psi^*$  denotes the conjugate value of  $\psi$ . In equation 1, quantity  $\psi \times \psi^*$  plays the role of a probability density.

### 2.2 Unitarity principle

We say that the measurement set  $S$  is a universe cross section for a physical effect when it contains all possible measurement values for this physical effect. For example in the screen impact measurement,  $S$  can be the set all coordinate  $(x, y, z)$  which share same value for  $z$  coordinate. There is unitarity when the

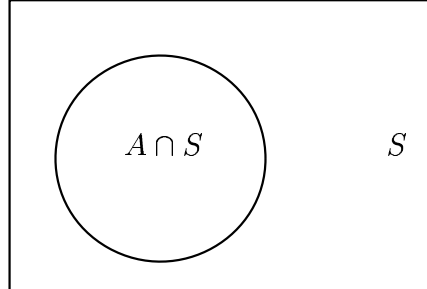


Figure 1: Conditional probabilities in quantum physics

integral  $\int \psi^* \times \psi$  is invariant when the set  $S$  varies (for example when the screen position varies). More precisely, if  $S_1$  and  $S_2$  are universe cross sections (see figure 2), then the flow of the squared wave function  $\psi^* \times \psi$  through  $S_1$  is the same as through  $S_2$ <sup>1</sup>.

$$\int_{S_1} \psi^* \times \psi = \int_{S_2} \psi^* \times \psi . \quad (2)$$

In equation 2 quantity  $\psi^* \times \psi$  plays the role of a probability current through hypersurfaces  $S_1$  and  $S_2$ . To simplify the presentation, we call *persistence* [1] of a set  $S$ , and we denote by  $\rho(S)$  the quantity  $\int_S \psi \times \psi^*$ . The unitarity principle is therefore equivalent to the persistence conservation over universe cross sections. The consequence of unitarity principle is that the probability that the particle hit  $S$  at a certain given subset  $A$  in  $S$  does not depend on the overall shape of  $S$  outside the spot. Unitarity is generally assumed to

---

<sup>1</sup>in relativist quantum mechanics, and for a particle with spin 0 the squared wave function current should be a 4-vector,  $i\psi^*\nabla\psi - i\psi\nabla\psi^*$ , particles with spin of higher need more dimensions and tensor product space-times



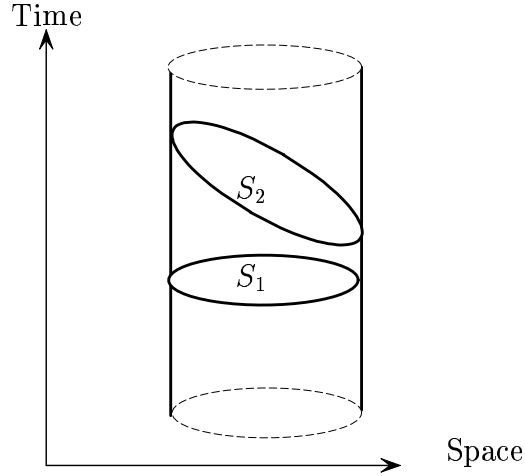


Figure 2: Probability conservation in quantum physics

hold in most physical measurement. The unitarity principle holds as soon the divergence of the probability current, symbolized by  $\psi \times \psi^*$ , is zero every where.

If the unitarity principle were to be violated, then the probability  $P(A|S)$ , which is equal to  $\rho(A)/\rho(S)$ , would vary with the shape of  $S$ . In other words an information transfer would occur between  $S$  and  $A$ :

$$P(A|S_1) \neq P(A|S_2) . \quad (3)$$

### 2.3 Retro-information on unitarity violation

We will prove in this section that unitarity violation implies retro-information and *vice versa*. We focus on the physical effect which is the measurement of a particle presence in coordinate  $(x, y, z, t)$  in space time. The hyper-surface made of points  $(x, y, z, t)$  which have same value  $t$  is a universe cross section. But a universe cross section can have more general shape. In particular it can

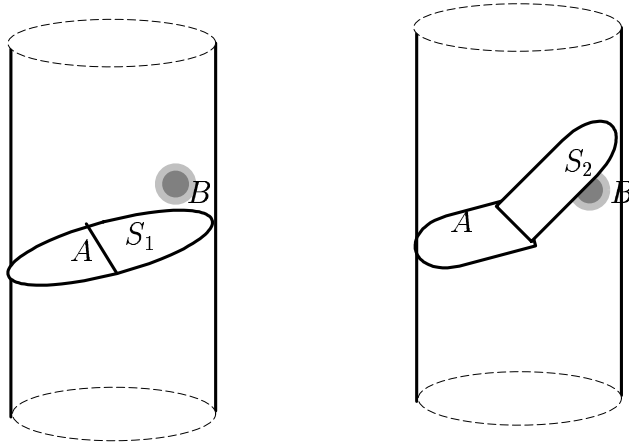


Figure 3: Information transfer over an unitarity singularity

contain a set  $A$  and other points which are in the future of  $A$  (i.e. such that the distance vector to  $A$  are of kind time backward). Let  $S_1$  and  $S_2$  be two universe cross hyper-surfaces which contains a same subset  $A$  and differs only on points which stands on the future of  $A$ , say around a location  $B$  (see figure 3). If we assume that a unitarity violation occurs between  $S_1$  and  $S_2$ , for example the probability current has non zero divergence in  $B$ , then  $P(A|S_1) \neq P(A|S_2)$ : there is an information transfer from  $B$  to  $A$ . Since this transfer is backward in time, therefore unitarity violation implies retro-information. In appendix we describe a hypothetical non-unitary system based on paired spin system which gives rise to retro-information.

We are going to prove the very simple following theorem which has the advantage to exercise the axiomatic of quantum axiomatic on retro-information:

**Theorem 1** *Retro-information is equivalent to unitarity violation.*

**Proof:** We already know that unitarity violation implies retro-information. We now prove the opposite. Let the retro-information channel be a physical

experiment made at a location  $A$  with two possible settings  $S_1$  and  $S_2$  controlled from a location  $B$  in the future of  $A$ . Without loss of generality we consider that the output channel is  $Y$  and is binary:  $Y \in \{1, 2\}$ . We define  $A_1$  as the set of measurement values which gives 1 as channel output and  $A_2$  as the set measurement which gives value 2 (thus  $A_2$  is the complementary set of  $A_1$ ).

By definition for all  $i \in \{1, 2\}$  and  $j \in \{1, 2\}$ :

$$P(Y = i|S_j) = \frac{\int_{A_i \cap S_j} \psi \times \psi^*}{\int_{S_j} \psi \times \psi^*} = \frac{\rho(A_i \cap S_j)}{\rho(S_j)}. \quad (4)$$

Since there is information transfer  $P(Y = 1|S_1) \neq P(Y = 1|S_2)$ . The identity  $S_j = A_1 \cap S_j + A_2 \cap S_j$  valid for all  $j \in \{1, 2\}$  implies:

$$\rho(S_i) = \rho(A_1 \cap S_i) + \rho(A_2 \cap S_i). \quad (5)$$

Assume *a contrario* that unitarity holds:  $\rho(S_1) = \rho(S_2)$ . Therefore for all  $i \in \{1, 2\}$

$$\rho(A_i \cap S_1) \neq \rho(A_i \cap S_2). \quad (6)$$

We now consider a minor extension of the system where a reliable channel brings the result of measurement from  $A$  to  $B$ . In the extended system we force the setting to be at  $S_1$  if the measurement in  $A$  is 1, otherwise the setting are at  $S_2$  (*i.e.* when  $Y = 2$ ). In this case the set  $A_1 \cap S_1 + A_2 \cap S_2$  is a universe cross section since output channel  $Y = i$  and setting  $S_j$  never occur together when  $i \neq j$ . We have

$$\rho(A_1 \cap S_1 + A_2 \cap S_2) = \rho(A_1 \cap S_1) + \rho(A_2 \cap S_2) \quad (7)$$

which is different of both  $\rho(S_1)$  and  $\rho(S_2)$  by virtue of equation (5) and (6). Therefore a unitarity violation occurs. ■

### 3 Retro-information over non-unitary EPR spin system.

After a collision on an arbitrary origin  $O$ , two twin particles, say particle 1 and particle 2 (for example two protons), carry exactly opposite spins. This relation

between the spins is non local since the particles continue to be correlated by their spins even if they are far apart. 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  $\theta_1$  and  $\theta_2$ . On each spin device the measure reads “+1” or “-1” (see figure 4). 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 [5]), and later analyzed by Bell [6]. We denote the pair of measures from the two spin devices by the tuple  $(\sigma_1(\theta_1), \sigma_2(\theta_2))$ .

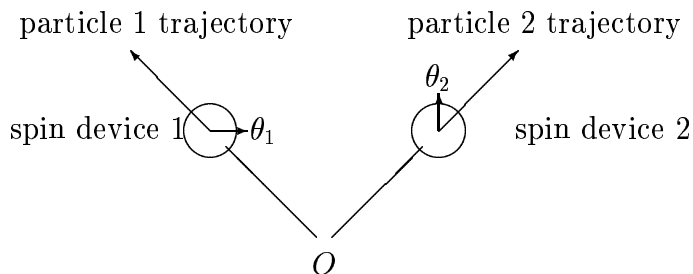


Figure 4: An EPR pair system

The wave function  $\mathbf{u}(\theta)$  (respectively  $\mathbf{d}(\theta)$ ) for up position (resp. for down position) when a spin detector is oriented to angle  $\theta$ . In unitarity scheme when a polarizer rotates the spin function also rotates with the relation:

$$\begin{aligned} \mathbf{u}(\theta) &= R_{\theta/2}(\mathbf{u}(0)) \\ \mathbf{d}(\theta) &= R_{\theta/2}(\mathbf{d}(0)) \end{aligned} \quad (8)$$

Where  $R_{\theta/2}$  is the rotation of angle  $\theta/2$ :

$$\begin{cases} R_{\theta/2}(\mathbf{u}(0)) &= \cos \frac{\theta}{2} \mathbf{u}(0) + \sin \frac{\theta}{2} \mathbf{d}(0) \\ R_{\theta/2}(\mathbf{d}(0)) &= -\sin \frac{\theta}{2} \mathbf{u}(0) + \cos \frac{\theta}{2} \mathbf{d}(0) . \end{cases} \quad (9)$$

Two particle with paired spins have stable expression wave function  $\psi = \mathbf{u}_1(\theta)\mathbf{d}_2(\theta) - \mathbf{d}_1(\theta)\mathbf{u}_2(\theta)$ , with  $\mathbf{u}_i(\theta)$ , resp.  $\mathbf{d}_i(\theta)$ , the wave function for up position (resp. down position) for particle  $i$ . Therefore  $\psi$  the wave function of the EPR spin system is

$$\psi = \mathbf{u}_1(0)\mathbf{d}_2(0) - \mathbf{d}_1(0)\mathbf{u}_2(0) . \quad (10)$$

In order to predict the effect of the second polarizer rotated of  $\theta_2 = \theta$ , the wave function expression is also

$$\begin{aligned} \psi &= \cos \frac{\theta}{2} \mathbf{u}_1(0)\mathbf{u}_2(\theta) - \sin \frac{\theta}{2} \mathbf{u}_1(0)\mathbf{d}_2(\theta) \\ &\quad - \sin \frac{\theta}{2} \mathbf{d}_1(0)\mathbf{u}_2(\theta) - \cos \frac{\theta}{2} \mathbf{d}_1(0)\mathbf{d}_2(\theta) . \end{aligned} \quad (11)$$

If the two particles collide over a unitarity singularity located on  $O$  (for example if they have enough energy to reach the non-unitary range, see figure 5) then the EPR system will catch and keep the non-unitarity correlation over arbitrary space-time distance. An hypothetical (simple) non-unitary correlation can be imagined over a slight modification of the rotating identity:

$$\begin{aligned} \mathbf{u}(\theta) &= (a(\theta))^{-1} R_{\theta/2}(\mathbf{u}(0)) \\ \mathbf{d}(\theta) &= (a(\theta))^{-1} R_{\theta/2}(\mathbf{d}(0)) \end{aligned} \quad (12)$$

A candidate expression for  $a(\theta)$  is  $\exp(\nu e^{2i\theta} - \nu)$ , we call  $\nu$  the unitarity coefficient supposed to be very small. The stable expression of the coupled system is still  $\mathbf{u}_1(\theta)\mathbf{d}_2(\theta) - \mathbf{d}_1(\theta)\mathbf{u}_2(\theta)$ .

Therefore the wave function  $\psi$  of the non-unitary system satisfies the identity:

$$\begin{aligned} \psi &= \mathbf{u}_1(0)\mathbf{d}_2(0) - \mathbf{d}_1(0)\mathbf{u}_2(0) \\ &= a(\theta) \cos \frac{\theta}{2} \mathbf{u}_1(0)\mathbf{u}_2(\theta) - a(\theta) \sin \frac{\theta}{2} \mathbf{u}_1(0)\mathbf{d}_2(\theta) \\ &\quad - a(\theta) \sin \frac{\theta}{2} \mathbf{d}_1(0)\mathbf{u}_2(\theta) - a(\theta) \cos \frac{\theta}{2} \mathbf{d}_1(0)\mathbf{d}_2(\theta) . \end{aligned} \quad (13)$$

Let  $A_i(+1)$  is the measurement subset corresponding to spin +1 for particle  $i$ , and let  $A_i(-1)$  corresponding to spin -1. Let  $S(\theta_1, \theta_2)$  be the setting set corresponding to polarizer 1 oriented with angle  $\theta_1$  and polarizer 2 oriented with angle  $\theta_2$ . The above identities leads to the persistence expression

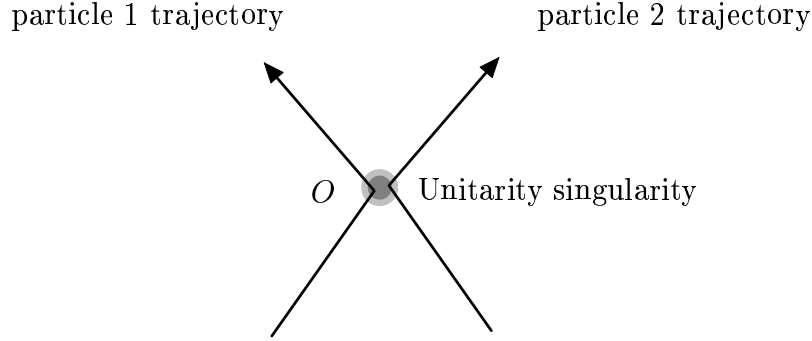


Figure 5: An EPR system over an unitarity singularity

$$\begin{cases} \rho(A_1(+1) \cap A_2(-1) \cap S(0, 0)) = \rho(A_1(-1) \cap A_2(+1) \cap S(0, 0)) = 1 \\ \rho(A_1(+1) \cap A_2(+1) \cap S(0, 0)) = \rho(A_1(-1) \cap A_2(-1) \cap S(0, 0)) = 0 \end{cases} \quad (14)$$

and

$$\begin{cases} \rho(A_1(+1) \cap A_2(-1) \cap S(0, \theta)) = \rho(A_1(-1) \cap A_2(+1) \cap S(0, \theta)) \\ \quad = |a(\theta)|^2 (\cos \frac{\theta}{2})^2 \\ \rho(A_1(+1) \cap A_2(+1) \cap S(0, \theta)) = \rho(A_1(-1) \cap A_2(-1) \cap S(0, \theta)) \\ \quad = |a(\theta)|^2 (\sin \frac{\theta}{2})^2 \end{cases} \quad (15)$$

Thus

$$\begin{cases} \rho(A_1(+1) \cap S(0, 0)) = \rho(A_1(-1) \cap S(0, 0)) = 1 \\ \rho(A_1(+1) \cap S(0, \theta)) = \rho(A_1(-1) \cap S(0, \theta)) = |a(\theta)|^2 \end{cases} \quad (16)$$

This identity implies that an information transfer can occur between the two polarizers as illustrated by figure 6.

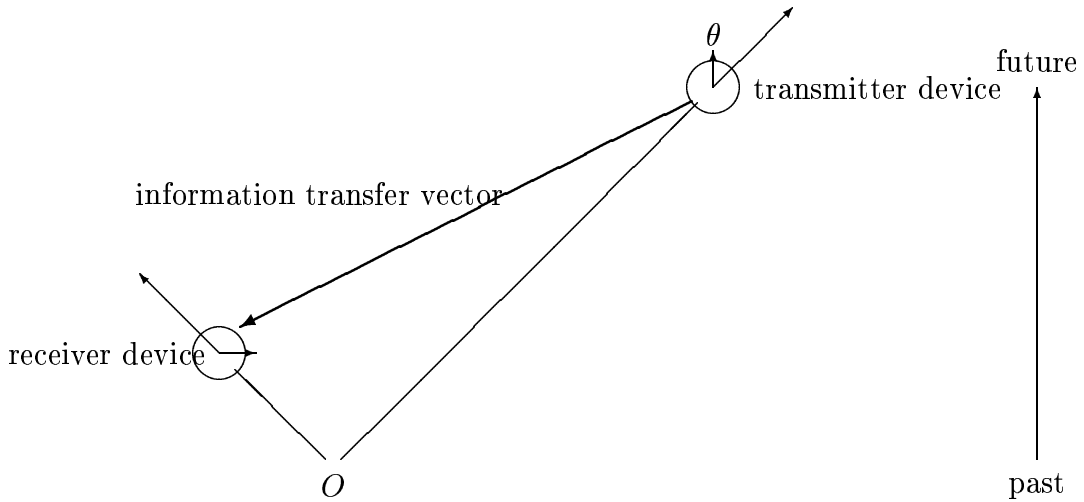


Figure 6: An hypothetical retro-information system build on EPR system

## 4 Capacity of retro-information channels

Like in classical information theory a  $V$ -ary channel with  $M$  setting is modeled via its transfer operator  $\mathbf{T}$ , which is a  $V \times M$  where the  $(i, j)$  coefficient is  $p_j(i) = \rho(S_i \cap A_j)$ : in other words  $p_j(i)/(\sum_{i'} p_j(i'))$  is equal to the In other words, for  $1 \leq i \leq V$  and  $1 \leq j \leq M$ ,  $p_j(i)/(\sum_{i'} p_j(i'))$  is equal to the probability that the channel output equal symbol number  $i$ , for  $1 \leq i \leq V$ , when the channel setting is  $S_j$ .

In the non-unitary EPR system described in the previous section,  $V = M = 2$  and the operator  $\mathbf{T}$  has expression:

$$\mathbf{T} = \begin{bmatrix} 1 & |a(\theta)|^2 \\ 1 & |a(\theta)|^2 \end{bmatrix} \quad (17)$$

The setting depends on a point situated in the future of the channel output. We say that the transfer operator is unitarity when all rows sum to the same constant. As we have discussed in the previous sections, the operator is not necessarily unitary.

In this section we consider  $n$  i.i.d. retro-information channels, each of them modelled by the transfer operator  $\mathbf{T}$ . In a first subsection we consider that

the  $n$  channels are in parallel, *i.e.* the  $n$  settings depend on a point which is situated in the future of the  $n$  output of the channels. In the second sub-section the channel are in *cascade*, *i.e.* the output of channel  $k$  is in the future of the output of channel  $k - 1$ , but the setting of channel  $k$  depend on a point in the future of the output of channel  $k - 1$ . Under this configuration, information can be transferred via multi-hops without range limitation.

#### 4.1 Capacity of retro-information channels in parallel

Quantity  $X_n = x_1 x_2 \dots x_n$  denotes the  $V$ -ary codeword to be sent and  $Y_n = y_1 y_2 \dots y_n$  is the  $V$ -ary codeword received, see figure 7. We denote by  $Z_n = z_1 z_2 \dots z_n$  the  $M$ -ary word made by the vector of settings applied to the set of  $n$  retro-channels.

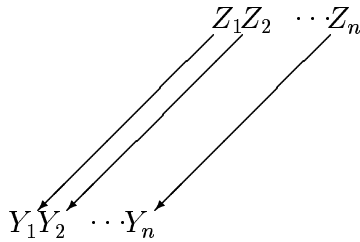


Figure 7: Channels in parallel

**Lemma 1** *The transfer operator of the  $n$  i.i.d retro-information channels in parallel is  $\mathbf{T}_n = \mathbf{T}^{\otimes n}$  when  $\mathbf{T}$  is the transfer operator of one channel.*

**Proof:** This result is obvious in classical information theory [3] but must be carefully revisited for in retro-information, since the operator might be non-unitary. Since the channels are independent, the wave functions of the system is equal to the product of the wave functions on each channel. If we call  $S(Z_n)$



the collective setting set of the channel corresponding to word  $Z_n$ , then we have

$$S(Z_n) = S_{z_1} \times S_{z_2} \times \cdots \times S_{z_n} . \quad (18)$$

Similarly, we denote  $A(Y_n)$  the subset of measurement which corresponds to the output codeword  $Y_n$ :

$$A(Y_n) = A_{y_1} \times A_{y_2} \times \cdots \times A_{y_n} . \quad (19)$$

Therefore the persistence of  $A(Y_n) \cap S(Z_n)$  satisfies the product form expression:

$$\rho(A(Y_n) \cap S(Z_n)) = \rho(A_{y_1} \cap S_{z_1}) \times \cdots \times \rho(A_{y_n} \cap S_{z_n}) . \quad (20)$$

■

The probability  $P(Y_n|Z_n)$  that output  $Y_n$  comes when setting codeword is  $Z_n$  has the expression:

$$P(Y_n|Z_n) = \frac{\rho(A(Y_n) \cap S(Z_n))}{\sum_{Y'_n} \rho(A(Y'_n) \cap S(Z_n))} . \quad (21)$$

In classical information theory, the setting  $Z_n$  depends on the codeword  $X_n$ :  $Z_n = Z_n(X_n)$  and the information transferred  $I(X_n, Y_n)$  is given by the formula

$$I(X_n, Y_n) = H(Y_n) - H(Y_n|X_n) \quad (22)$$

where  $H(Y_n)$  is the entropy of the variable  $Y_n$  and  $H(Y_n|X_n)$  the entropy of the variable  $Y_n$  conditionned by  $X_n$ . The above expression is computable as long we know the probabilities  $P(Y_n|Z_n)$  and the function  $Z_n(X_n)$ .

The classical capacity  $C_n$  of the channels in parallel is given by

$$C_n = \max_{X_n, Z_n(X_n)} I(X_n, Y_n) . \quad (23)$$

In retro-information transfer we assume that a reliable (classical) channel collects the output  $Y_n$  of the retro-channels and brings the result just before the setting of the channel is actually made. Under this configuration, a Forward

Coupling Function (FCF) can be defined to make  $Z_n$  dependent on both  $X_n$  and  $Y_n$ :  $Z_n = Z_n(X_n, Y_n)$ .

In this case identity (21) becomes:

$$P(Y_n|X_n) = \frac{\rho(A(Y_n) \cap S(Z_n(X_n, Y_n)))}{\sum_{Y'_n} \rho(A(Y'_n) \cap S_n(Z_n(X_n, Y'_n)))}. \quad (24)$$

In presence of Forward Coupling Function the channel capacity  $C_n^F$  is given by

$$C_n^F = \max_{X_n, Z_n(X_n, Y_n)} I(X_n, Y_n). \quad (25)$$

An interesting objective is to find a forward coupling function (FCF) which optimizes the information transfer capacity. In fact we will exhibit a set of FCF's which greatly improves the channel capacity with regard to classic channel coding.

**Theorem 2** *The capacity  $C_n^F$  of  $n$  retro-information channels in parallel satisfies the following inequality*

$$C_n^F \geq n \min\{\log V, \log \frac{\sum_i \max_j p_j(i)}{\sum_i \min_j p_j(i)}\}. \quad (26)$$

**Remark:** This new capacity is much greater than the classical capacity  $C_n$  without FCF. In particular, assuming  $\mathbf{T}$  is unitary for convenience, and  $p_1(1) - p_2(1) = \delta \ll 1$ , we have  $C_n \approx \frac{1}{4} n \frac{\delta^2}{(1-p_1)p_1}$ , and  $C_n^F \approx 2n\delta$ . If we assume a symmetric channel with

$$\mathbf{T} = \begin{bmatrix} p & 1-p \\ 1-p & p \end{bmatrix}$$

with  $p \in [0.5, 1]$ , then in absence of FCF the capacity is  $nC$  with  $C = 1 + p \log p + (1-p) \log(1-p)$ , while  $C^F = \min\{\log 2, \log \frac{p}{1-p}\}$ . Figure 8 displays  $C$  and  $C^F$  versus  $p$  expressed in *binary* (bit).

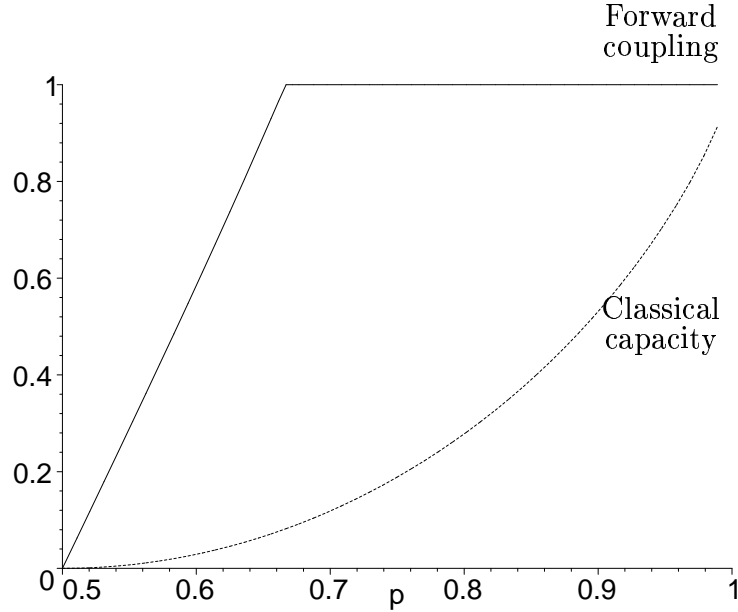


Figure 8: Parallel retro-information channel capacity, with as a function of  $p$ , classical ( $C/\log 2$ ), with forward coupling function ( $C^F/\log 2$ )

**Remark:** The estimate for capacity is self consistent: if we define  $C^F = C^F(\mathbf{T})$  as a function of transfer operator  $\mathbf{T}$ , then  $C^F(\mathbf{T}_1 \otimes \mathbf{T}_2) = C^F(\mathbf{T}_1) + C^F(\mathbf{T}_2)$ .

**Proof of theorem 2:** Without loss of generality we assume that the channels are binary: *i.e.*  $V = M = 2$ .

For each  $i$  we denote  $\bar{p}(i)$  the quantity  $\max_j \{p_j(i)\}$  and  $\bar{i}$  a value of  $j$  which attains this maximum. We denote also  $\underline{p}(i)$  the quantity  $\min_j \{p_j(i)\}$  and  $\underline{i}$  a value of  $j$  which attains this minimum. For a  $V$ -ary codeword  $Y_n = y_1 \dots y_n$  we denote  $\overline{Y}_n$  the  $M$ -ary codeword  $\overline{y}_1 \dots \overline{y}_n$ . Similarly we denote  $\underline{Y}_n$  the  $M$ -ary codeword  $\underline{y}_1 \dots \underline{y}_n$ .

We describe an FCF such that channel capacity tends to be equivalent to  $nC^F$  when  $n \rightarrow \infty$  with  $C^F = \min\{\log 2, \log \frac{\bar{p}(1)+\bar{p}(2)}{\underline{p}(1)+\underline{p}(2)}\}$ . To this end for each

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

1. when  $X_n$  and  $Y_n$  agree on the  $k$  first bits then  $Z_n = \overline{Y_n}$ , *i.e.* the setting is "parallel";
2. otherwise  $S_n = \underline{Y_n}$ , *i.e.* the setting is "in opposition".

The goal is to find an estimate of the entropy  $H(Y_n|X_n)$ . 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 the sum of the persistence measure on the set  $D_k(X_n)$  assuming parallel settings

$$\rho(D_k(X_n)) = \overline{p}_k(X_n)(\overline{p}(1) + \overline{p}(2))^{n-k} \quad (27)$$

where  $\overline{p}_k(X_n)$  is the factor obtained by replacing each 1 by  $\overline{p}(1)$  and each 2 by  $\overline{p}(2)$  in the first  $k$  characters of  $X_n$ .

Conversely the persistence of  $\overline{D_k}(X_n)$  is the sum of persistence measure on the set  $\overline{D_k}(X_n)$  assuming settings in opposition:

$$\rho(\overline{D_k}(X_n)) = (\underline{p}(1) + \underline{p}(2))^n - \underline{p}_k(X_n)(\underline{p}(1) + \underline{p}(2))^{n-k} \quad (28)$$

where  $\underline{p}_k(X_n)$  is the factor obtained by replacing each 1 by  $\underline{p}(1)$  and each 2 by  $\underline{p}(2)$  in the first  $k$  characters of  $X_n$ .

The probability that  $Y_n$  belongs to  $D_k(X_n)$  is

$$\frac{\rho(D_k(X_n))}{\rho(D_k(X_n)) + \rho(\overline{D_k}(X_n))}$$

We assume that  $X_n$  is generated via a Bernoulli model for  $X_n$  with probability  $x$  for symbol 1 and probability  $1 - x$  for symbol 2. For all  $X_n$  we have almost surely when  $k \rightarrow \infty$ :

$$\log \overline{p}_k(X_n) \sim (x \log \overline{p}(1) + (1 - x) \log \overline{p}(2))k \quad (29)$$

The game consists into selecting  $k$  such that

$$\lim_{n \rightarrow \infty} \frac{\rho(\overline{D}_k(X_n))}{\rho(D_k(X_n))} = 0 \quad (30)$$

In other words  $Y_n$  will agree with  $X_n$  on the first  $k$  digits with probability 1 when  $n \rightarrow \infty$ . Furthermore the random variable  $Y_n$  conditioned by a fixed  $X_n$  will tend to have its last  $n - k$  digit a Bernoulli distribution with probability  $\frac{\overline{p}(1)}{\overline{p}(1) + \overline{p}(2)}$  for the 1, and probability  $\frac{\overline{p}(2)}{\overline{p}(1) + \overline{p}(2)}$  for the one. Consequently we will get

$$\frac{H(Y_n|X_n)}{n} \sim (n - k)h\left(\frac{\overline{p}(1)}{\overline{p}(1) + \overline{p}(2)}\right) \quad (31)$$

Similarly, condition (30) would imply:

$$H(Y_n) = h(x)k + H(Y_n|X_n) \quad (32)$$

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

In order to have condition (30) satisfied we need

$$(x \log \overline{p}(1) + (1 - x) \log \overline{p}(2))k + (n - k) \log(\overline{p}(1) + \overline{p}(2)) > n \log(\underline{p}(1) + \underline{p}(2))$$

therefore condition (30) holds for all  $k$  such that

$$\frac{k}{n} < \frac{-1}{x \log \overline{p}(1) + (1 - x) \log \overline{p}(2) - \log(\overline{p}(1) + \overline{p}(2))} \log \frac{\overline{p}(1) + \overline{p}(2)}{\underline{p}(1) + \underline{p}(2)} \quad (33)$$

In the case the right hand-side of above (33) is not greater than 1, since nevertheless we must have  $k \leq n$  we can plug the upper bound in the expression of  $C_n^F(x)$ . It comes that

$$\begin{aligned}
C_n^F &\geq \max_x C_n^F(x) \\
&= n \max_x \left\{ \frac{-h(x)}{x \log \bar{p}(1) + (1-x) \log \bar{p}(2) - \log(\bar{p}(1) + \bar{p}(2))} \log \frac{\bar{p}(1) + \bar{p}(2)}{\underline{p}(1) + \underline{p}(2)} \right\} \\
&= n \log \frac{\bar{p}(1) + \bar{p}(2)}{\underline{p}(1) + \underline{p}(2)}
\end{aligned}$$

■

**Remark:** the maximum is attained for  $x = \frac{\bar{p}(1)}{\bar{p}(1) + \bar{p}(2)}$ .

One immediately notices that we would obtain capacity greater than  $n \log 2$  when  $\frac{\bar{p}(1) + \bar{p}(2)}{\underline{p}(1) + \underline{p}(2)} > 2$  which should not be possible since the codeword  $Y_n$  has only  $n$  digits. In fact this apparent inconsistency is only the consequence of side effect when the strict application of (33) 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

$$C_n^F \geq n \min\{\log 2, \log \frac{\bar{p}(1) + \bar{p}(2)}{\underline{p}(1) + \underline{p}(2)}\}$$

The case  $\frac{\bar{p}(1) + \bar{p}(2)}{\underline{p}(1) + \underline{p}(2)} > 2$  is nevertheless interesting because it allows to limit FCF on a fraction of retro-channels. For example Let  $\ell$  be the integer part of  $\frac{n}{\log 2} \log \frac{\bar{p}(1) + \bar{p}(2)}{\underline{p}(1) + \underline{p}(2)}$ , the coupling  $Z_n(Y_n, X_n)$  can be the following:

1. The  $\ell$  first characters of  $Z_n$  are the  $\ell$  first characters of  $\bar{Y}_n$  when  $Y_n = X_n$ , otherwise the  $\ell$  first characters of  $Z_n$  equal the  $\ell$  first characters of  $\underline{Y}_n$ .
2. the  $n - \ell$  last characters of  $S_n$  do not depend on  $Y_n$  (e.g equal to the  $n - \ell$  last characters of  $X_n$  or are constant).

## 4.2 Capacity of retro-information channels in cascade

In this section we consider  $n$  i.i.d. retro-information channels with a  $V \times V$  transfer operator. The  $n$  channels are set in cascade, see figure 9. We take the trivial case in *series* where the setting of the  $k$ -th channel is equal to the output of the  $k - 1$ -th channel. To this end the transfer operator  $\mathbf{T}$  is a  $V \times V$  matrix.

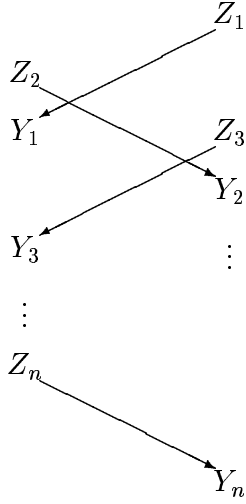


Figure 9: Channels in cascade

**Lemma 2** *The transfer operator of  $n$  i.i.d. channels in series is  $\mathbf{T}^n$ .*

**Proof:** A valid measurement-setting tuple  $(X_n, Z_n)$  is such that  $y_i = z_{i+1}$  for all  $1 \leq i < n$ . The persistence of the set  $A(X_n) \cap S(Z_n)$  given that  $z_1 = j$

$$\begin{aligned} \rho(A(X_n) \cap S(Z_n)) &= \rho(A_{y_n} \cap S_{y_{n-1}}) \times \cdots \times \rho(A_{y_1} \cap S_j) \\ &= p_{y_{n-1}}(y_n) \times \cdots \times p_{y_1}(y_2) p_j(y_1). \end{aligned} \quad (34)$$

Therefore the persistence of the set of valid measurement tuples  $(X_n, Z_n)$  such that  $y_n = i$  and  $z_1 = j$  is obtained by summing the persistence of  $\rho(A(Y_n) \cap S(Z_n))$  over all possible  $(Y_n, Z_n)$  tuples, which is equivalent to the extraction of the  $(i, j)$  coefficient of  $\mathbf{T}^n$ . ■

In classical information theory the capacity of  $\mathbf{T}^n$  tends to zero when  $n$  tends to infinity. We will show that when  $\mathbf{T}$  is non-unitary the capacity of the  $n$  channels in series remains a certain threshold when  $n$  tends to infinity.

We denote  $\Pi$  and  $\eta$  respectively the right and left eigenvector of the main eigenvalue of operator  $\mathbf{T}$ . We know that  $\eta \neq (1, \dots, 1)$ . We denote  $\eta = (\eta_1, \dots, \eta_V)$ . We have the following result.

**Theorem 3** *The capacity of the  $n$  retro-information channels in series when  $n$  tends to infinity tends to the capacity of the single retro-channel with transfer operator  $\Pi \otimes \eta$ .*

**Proof:** Denoting by  $\lambda$  the main eigen-value of  $\mathbf{T}$  and  $\lambda'$  the second eigenvalue, one has  $\mathbf{T}^n = \lambda^n \Pi \otimes \eta + O((\lambda')^n)$  when  $n$  tends to infinity. After renormalisation by  $\lambda^n$ , the lemma holds. ■

**Corollary 1** *The capacity of  $m$  series of  $n$  channels in series, in presence of forward coupling, tends to be equal to  $m \min_i \{\log V, \log \max_i \{\eta_i\} - \log \min_i \{\eta_i\}\}$ .*

## 5 Conclusion

We have presented a new information transfer architecture in information theory based on hypothetical physical assumption. Retro-information *a priori* seems to be a logical challenge (even if restricted on short space-time distance), however it can be framed in a consistent axiomatic which can take into account paradoxical effects due to forward coupling. The unexpected results about channel capacity in presence of forward coupling contributes to make retro-information a very promising area of investigation.

## References

- [1] P. JACQUET AND V. JOLY, “Retro-information in Wheeler-Feynman universe model: application over an hypothetical concept in quantum mechanics,” INRIA RR-3530, 1998.
- [2] B. SCHUMACHER, M.D. WESTMORELAND, “Sending classical information via noisy quantum channel,” *Phys. Rev. A*, vol. 56, pp. 131-138, 1997.
- [3] T. OGAWA, H. NAGAOKA, “Strong converse to the quantum channel coding theorem,” Tech. Rep. UEC-IS-1998-6, University of Electro-Communicatyon, 1998
- [4] J.H. SCHWARZ, E. WITTEN, *Superstring Theory*, Cambridge University Press, 1987.



- [5] A. EINSTEIN, B. PODOLSKY, N. ROSEN, “Can Quantum Mechanical Description of Physical Reality be considered complete?” in *Physical Review*, 47, 777-780, 1935.
- [6] J.S. BELL, “On the Einstein-Podolsky-Rosen Paradox,” in *Physics*, 1, 195-200, 1964.



---

Unit e de recherche INRIA Lorraine, Technop le de Nancy-Brabois, Campus scientifique,  
615 rue du Jardin Botanique, BP 101, 54600 VILLERS L ES NANCY  
Unit e de recherche INRIA Rennes, Irista, Campus universitaire de Beaulieu, 35042 RENNES Cedex  
Unit e de recherche INRIA Rh ne-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

---

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