Quantum observer networks and quantum Shannon theory

A. Feller[1], B. Roussel[1], I. Frérot[4], O.Fawzi[3], and P. Degiovanni[2]
(1) European Space Agency - Advanced Concepts Team, ESTEC, Keplerlaan 1, 2201 AZ Noordwijk, The Netherlands.
(2) Univ Lyon, Ens de Lyon, Université Claude Bernard Lyon 1, CNRS, Laboratoire de Physique, F-69342 Lyon, France
(3) Univ Lyon, Ens de Lyon, Université Claude Bernard Lyon 1, CNRS, INRIA Laboratoire d'Informatique du Parallélisme, F-69342 Lyon, France and
(4) ICFO-Institut de Ciencies Fotoniques, The Barcelona Institute of Science and Technology, Av. Carl Friedrich Gauss 3, 08860, Castelldefels, (Barcelona), Spain
(Dated: March 24, 2023)

Quantum theory is the fundamental framework which describes all of modern physics except gravity. Despite its tremendous successes, fundamental questions regarding its foundations and interpretation are still actively investigated. A key question is to understand the quantum-to-classical transition without relying on ad-hoc postulates. The decoherence program, a significant step towards answering this question, has recently been renewed by considering the information flow from a system into its environment. This led to a paradigm shift called "quantum Darwinism" in which the environment is now seen as a quantum communication channels between the system and observers. In this study, we propose to push this idea forward by considering the full (quantum) network of observers in the quantum Shannon theory framework.

Our main result is the clarification of the information theoretical criteria for independent and shared objectivity as well as the identification of the corresponding shared quantum states between the system and the quantum observers. This leads to a (non-trivial) separation between the classical information broadcasted to the observers by the system and the (correlated) quantum noise they experience, clarifying the structure of the global correlations in the "observed system/many observers" quantum system.

PACS numbers:
Keywords: quantum Shannon theory, quantum observers, decoherence, classicality

Contents

I. INTRODUCTION AND MOTIVATION

A. Motivation

The emergence of the classical world from quantum theory and of a consensus among different observers is a long-standing problem for the foundations of quantum theory.

Already within classical physics, Einstein emphasized the crucial role played by the correspondance between the way different observers perceive and describe the same reality. Developing the theory of special relativity, he derived such correspondance, both for space-time coordinates and for tensor fields over space-time, from the invariance of the speed of light [20]. To extend special relativity to general relativity [22], Einstein went further and considered more general transformations between observers; and promoted the general covariance of physical laws to the status of a fundamental dynamical principle of Nature. Together with the equivalence principle [21], this led to the general theory of relativity. In this description, as well as in any consideration involving observers in classical physics, the act of observation does not perturb the system in a fundamental way: in principle, it is always possible to perform an infinite-precision measurement, and copy the "local physical reality" into the observer's memory. The correspondance between observers then consists of transition functions between the observers memories satisfying certain consistancy conditions, as sketched on Fig. 1. Local observers, their memory records, together with suitable transition functions between them, form what we will call a *relational description*, and is the kinematical framework of any classical relativistic model[1].

Remarkably, such relational description is sufficient to reconstruct a *geometrical* interpretation, according to which the observer's memory record "points" (or more general concepts) associated with an underlying geometrical object[2]. This geometrization of physics, which started with Minkowski [52], was one of the unexpected consequences of Einstein's work [24]. While this is often interpreted as the existence of an underlying absolute reality behind the *relational description* by all observers, as pointed out by Einstein himself [22], the absoluteness of the geometric object itself is however more subtle than one could naively expect.

It should be emphasized that Einstein's relativity, and more generally any classical model, posits a clear separation between the obsevers on the one hand, who can map the part of space-time accessible to them using clocks and rules, and the physical quantities on the other hand, such as local fields, which the observers can measure. As in any classical theory, the state of the system under study is not altered by an ideal measurement. This clear distinction between the object – here, the system – and the subject – here, the observers –, is not valid in quantum theory. As Heisenberg already pointed out, any act of measurement disturbs the system in a fundamental way, thereby constraining the observer's ability to infer information about the system [38]. As a consequence, understanding the process of measurement itself requires in principle a detailed modelization of the observer's instruments, which leads to a much more involved situation than in classical physics.

A workaround to this last problem was proposed by von Neumann [67], who formalized the so-called measurement postulate, which enables the extraction of predictions from the quantum formalism. This came at the price of cornering the observer away from the quantum formalism, under the form of a classical intruder whose back-action on the system's state is described by the measurement postulate.

But, as stressed out by Everett [27], if quantum theory is to be a complete description of Nature, the question of "relativity" in a quantum universe has to addressed in its full complexity. In particular, the observers themselves should be treated as physical systems. This led Everett to discuss a gedanken experiment [27] (known today as the Wigner's friend experiment [71]), and whose consequences are still highly debated [1, 32, 46]. To close this debate requires extending the kinematical framework of

[1] Notice that relativity is not necessarily Einsteinian: non-relativistic physics is described using models that have Gallilean invariance and therefore are "Gallilean relativistic" models.

[2] In classical field theory, the geometric object is a fiber bundle over a manifold, called space time, and a configuration of the classical field is a section of this fiber bundle.

classical relativity theories, to a quantum universe. This basic question is the main motivation behind the work presented in this report.

B. Towards quantum relativity

Trying to extend the ideas of relativity to the quantum realm immediately rises several questions. In this section, we formulate them in order of increasing complexity.

The first question concerns the definition of "quantum observers": in the quantum universe, what plays the role of Einstein's classical observers, who can reconstruct their own image of a physical reality? As pointed out by Everett, observers should be treated as subsystems within the quantum universe, acquiring information about the system through physical interaction. In the quantum formalism, this inevitably leads to quantum correlations between the system and all the observers.

In essence, this is reminiscent of Laudauer's insight that information only "exists" through the physical systems carrying it [48]. As explained in this report, this has led us to define quantum observer as quantum subsystems within the universe, able to perform general quantum measurements on the local degrees of freedom (DoF) they can access. These DoF are not directly the system's DoF they are trying to characterize. Instead, they correspond to certain "fragments" of the systems'environment, carrying incomplete information about the system itself. The operations perfomed by the observers result from physical interactions within the quantum universe, thereby promoting the observers to the status of active participants. This situation should be contrasted to classical physics, where ideal observers play only a passive role in acquiring information. As summarized in Fig. 1, the quantum observer is the embodiement of the classical observer's idea in a world where information can only be acquired via physical interactions.

Along the same lines, we shall also discuss the relationship among the different observers observations. The abstract transition functions of classical theory are replaced with physical subsystems, namely communication links. This perspective brings classical and quantum communication theory into the game.

After formulating this general framework, one must identify the circumstances under which this *quantum relational description* reduces to the classical one. Under which circumstances does a classical reality emerge from within a quantum universe? This grand-problem has been addressed since the very inception of quantum theory. Significant progress has been made with the theory of decoherence [72–74, 76], which explains the effective disappearance of quantum interference effects in open quantum systems. But this did not answer the question of the emergence of a unique and common classical reality. The idea of quantum Darwinism, originally introduced by Ollivier, Poulin and Zurek [55] recently opened a new perspective on this problem. This lead to promote the environment of a quantum system to the status of an active communication medium, and to discuss the question of a concensus among the observer using information-theory concepts. One of the main objectives of the present project was to explore the relevance of these two paradigm shifts to build an appropriate framework for "quantum relativity". These developments would enable to explore the *quantum geometric object* – if any – which lies behind the quantum relational description. The existence of such an object has been intuitively pushed forward following Everett's original work, and is indeed present in Everett's PhD thesis entitled *The theory of the universal wave function* [27].

In this report, we shed light on these problems using recent advances in multipartite quantum information theory. Besides their interest for fundamental physics, they are relevant to quantum sensing and metrology when considering a quantum network of stations equipped with quantum sensors. This situation indeed captures the essence of using quantum probes for space and universe exploration.

C. Main results

The main results obtained in this project are the following:

- Quantum observers have been defined. A dictionary between the observer approach to quantum Darwinism and quantum Shannon theory is proposed and a link to estimation theory discussed. A hierarchy of quantum observers is proposed, depending on their communication capabilities, and the degree of collaboration among them.

- The quantum Darwinism framework has been extended to account for different situations, based on the above-mentioned hierarchy of quantum observers: information-theorey criteria for the emergence of objectivity in each of the observer classes have been obtained, and the corresponding shared quantum state structure established (up to one class for which it is still a conjecture).

- Various "phases" of objectivity corresponding to "shared maximally entangled states", "information locking", "secret sharing" and finally "quantum Darwinian objectivity" are discussed. The role of scaling in this hierarchy of phases (each one appearing at a specific scale) has been explored at a preliminary level, opening a way to understand the emergent character of Everettian branches in a quantum multiverse.

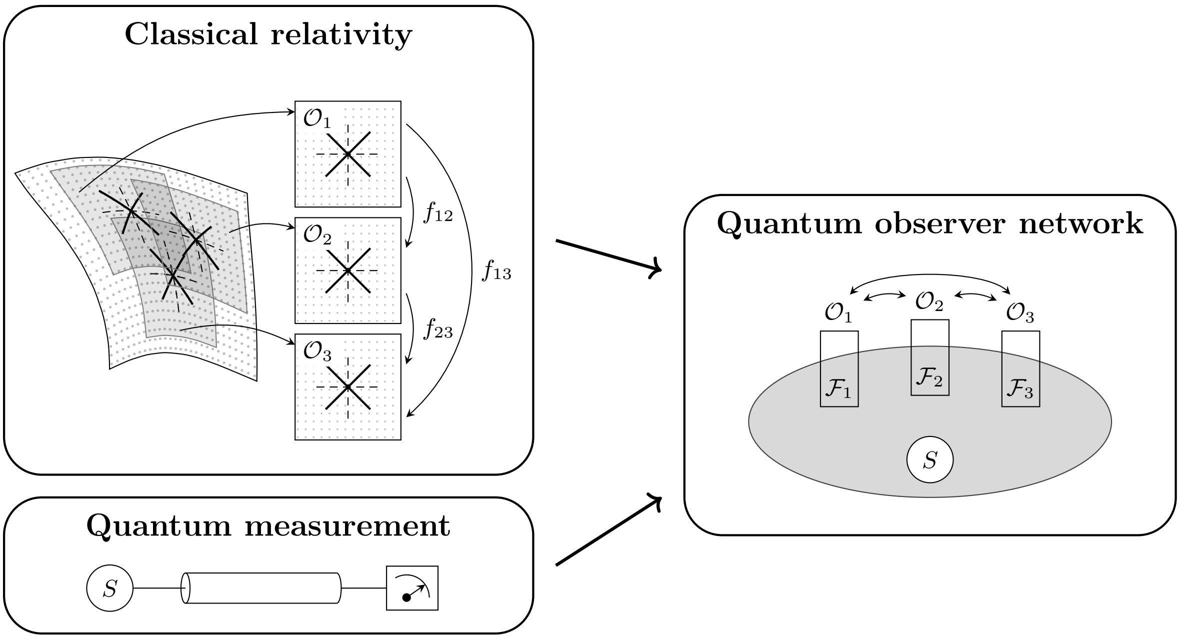

FIG. 1: Cartoon summary of the emergence of the quantum relativity framework from classical relativity and quantum measurements. *Classical relativity box*: Observers $\mathcal{O}_j$ pace space and time around them using clocks and rules (or radar signals) in order to map the various events in their vicinity into their local memory. The transition function f_{ij} maps the description for $\mathcal{O}_i$ into the one for $\mathcal{O}_j$. *Quantum measurement box*: a general quantum measurement consists of the interaction of the system and its environment and an ideal projective measurement performed in its environment. It can be modeled as a measurement performed on the output of a quantum communication channel. This doesn't need a caption

D. Organization of the report

This report is intended to be almost self-sufficient. It presents the results obtained during this project, but also reviews the basic concepts and existing works needed to recast them in a broader perspective. Technical details and proofs are presented in the Appendices.

In Section II, we introduce the notion of a quantum observer, by extending the classical (Einsteinian) notion of observer. We first consider ideal classical measurements, while imperfect classical measurements are discussed in Section II A. The role of communication and of information theory are already emphasized at the classical level. Quantum observers are then introduced in Section II B. The main questions addressed in this report are formulated in this framework in Section II C.

Then, Section III is devoted to the discussion of correlations. The important concept of quantum discord is introduced in Section III B, in order to quantify the amount of non-classical correlations between quantum subsystems. Quantum Darwinism is then reviewed in Section III C.

Section IV presents the main results obtained during the project. The distinction between *objectivity of observables* and *objectivity of observations* is discussed in Section IV A. A brief review of an important and general theorem (ref. [10]) is presented in Section IV B. In Section IV C, we introduce a hierarchy of objectivity levels, and discuss the information-theory criterion defining each level. We also unravel the structure of quantum states corresponding to each level of the hierarchy. Together with the framework introduced in Section II, these form the main results of our work.

Finally, Section V summarizes our results, and presents several perspectives for fundamental physics, quantum technologies, and potential space applications.

II. FROM CLASSICAL RELATIVITY TO QUANTUM OBSERVERS

A. Classical observers

In this section, we review measurement in classical physics, first by discussing in Section II A 1 the ideal measurement framework that underlies all classical theories[3]. Then, we will discuss noisy classical measurements by one classical observer in Section II A 2, connecting it to estimation problem and showing the relation to information theoretical quantities through the asymptotic equipartition theorem. Classical noisy measurements by several

[3] This framework was mainly introduced to discuss special [20] and then later general relativity [23] by A. Einstein but it can also be used to discuss any type of classical theory, even non-relativistic mechanics.

classical observers will then be discussed in Section II A 6.

1. *Ideal classical observation*

A classical observer $\mathcal{O}$ having complete access to the state X of a classical system records it in their own local memory as $X_{\mathcal{O}}$. If several observers $\mathcal{O}_i$ have access to the same system, one defines transition functions $F_{\mathcal{O}_i,\mathcal{O}_j}$ which maps $X_{\mathcal{O}_i}$ onto $X_{\mathcal{O}_j}$. Such transition functions satisfy the following compatibility condition:

$$F_{\mathcal{O}_1,\mathcal{O}_3} = F_{\mathcal{O}_2,\mathcal{O}_3} \circ F_{\mathcal{O}_1,\mathcal{O}_2} \,. \tag{1}$$

The transition functions satisfying Eq. (1) form what we call a classical *relational description*. In such a description, a transition function between two observers tells us which pairs of quantities are totally correlated.

Interestingly, such a classical relational description always define a geometric structure which is locally mapped by each observer in its own memory records. More specifically, in the case of classical field theory, specifying these compatibility conditions correspond to mathematically define space-time as a (differential) manifold and field space as a fiber bundle over space time as well as field configurations as (local) sections of this fiber bundle. In a measurement-oriented view, only the relational description is considered and, in classical physics, there is always an objectively-existing geometric object underlying this description (up to the appropriate isomorphism).

This is only a (very compact) review of the kinematic framework of observation in classical physics, for the special case of ideal observers having a complete access to the system's degree of freedom with infinite precision. In a more realistic framework, the observers do not have a perfect and complete access to the system: first of all, measurements are not perfect and one should also take into account the noise introduced by the measurement itself. Next, each observer could have access to only a part of the physical quantities, thereby giving them an incomplete reconstruction of physical reality. Hence, we shall now discuss more realistic (i.e. noisy) classical measurements.

2. *Noisy classical measurements*

As depicted on the left panel of Fig. 2, a noisy classical measurement can be viewed as an estimation problem. In this approach, the system S, characterized classically by its microstate s, broadcasts information on the value $X(s)$ of a certain observable X. The system is thus placed on the encoder end of a noisy classical communication channel $\mathcal{C}$, which represents both the experimental apparatus and noise induced by the environment. On the receiver end of the channel, an ideal measurement terminal together with a signal processor play the role of the decoder. The signal processing task consists in an estimation algorithm, which provides an estimate $X_{\text{est}}(m)$ for $x = X(s)$ from the measured values m.

The noisy communication channel plays the role of the experiment with its imperfections and the decoder plays the role of the classical observer. Notice that here, the encoder is fixed. The system S can be in various microstates that are characterized by a probability distribution $p_S(s)$ but the encoding stage of this system is the same, regardless of the microstate.

The noisy communication channel is characterized by the noise probability matrix formed by the conditional probabilities $p(m|x)$ for the measurement results, knowing the value of the observable X. This represents the imperfection of the measurement apparatus.

In full generality, the observer wishes to reconstruct the value of the observable from the measurements. The probability for guessing the correct result is defined as

$$p_{\text{guess}}(x|X_{\text{est}}, \mathcal{C}) = \mathbb{E}_m \left[\delta(X_{\text{est}}(m) - x)\right] \tag{2}$$

in which $m = M(x)$ is a random variable distributed according to $p(m|x)$. This quantity depends: 1) on the x we wish to estimate; 2) on the estimation algorithm X_{est}; and 3) on the noisy communication channel $\mathcal{C}$.

Various estimation scores can be defined. The most natural one for physicists is the average on all the possible values of x:

$$\mathrm{Q}[X_{\text{est}}|\mathcal{C}, p_X] = \mathbb{E}_x \left[p_{\text{guess}}(x|X_{\text{est}}, \mathcal{C})\right] \tag{3}$$

in which the average over x is taken according to the macro-state of S which induces a probability distribution $p_X(x)$ for the values of X. It quantifies the quality of the estimation algorithm X_{est} for a given imperfect measurement $\mathcal{C}$ and a given macro-state of the system[4]. In computer science as well as in all critical applications, one considers the worst case scenario score:

$$\mathrm{W}[X_{\text{est}}|\mathcal{C}, p_X] = \min_x \left[p_{\text{guess}}(x|X_{\text{est}}, \mathcal{C}, p_X)\right] \,. \tag{4}$$

It is easy to prove that the quality of the estimator $\mathrm{Q}[X_{\text{est}}|\mathcal{C}, p_X]$ is maximized by the maximum-likelihood estimator $X_{\text{est}}(m) = \operatorname{argmax}[p(x|m)]$ in which the likelihood of x for a given m is $p(x|m) = p(m|x)p(x)/(\sum_x p(x)p(m|x))$, defined using Bayes rule. For such a choice,

$$\mathrm{Q}[X_{\text{est}}|\mathcal{C}, p_X] = \sum_{m,x} p(x)p(m|x)\delta_{x,X_{\text{est}}(m)} \tag{5a}$$

$$= \sum_m p(m, X_{\text{est}}(m))$$

$$= \sum_m p(m)p(X_{\text{est}}(m)|m) \,. \tag{5b}$$

---

[4] This amounts to say that it depends on the joint probability distribution $p(m, x) = p(m|x)\, p_X(x)$.

This last quantity is maximized if, for all m, we choose $X_{\text{est.}}(m) = \text{argmax}[p(x|m)]$ which is the maximum likelihood estimator denoted by ML. Of course considering this optimal estimation algorithm gives the highest guessing probability for the noisy channel, or equivalently noisy measurement, considered. The guessing probability $p_{\text{guess}}(\mathcal{C}, p_X) = \text{Q}[\mathsf{ML}|\mathcal{C}, p_X]$ then only depends on: 1) the channel $p(m|x)$; and 2) on p_X. Equivalently, it is a function of the joint distribution $p(x, m)$ for the input x and output m of the channel. Note that, as it depends on p_X, this indicator is defined for a given macro-state of s. It is appropriate to assess the performance of the measurement in a given experimental situation. We shall come back on this point later, when discussing the notion of single-use capacity of $\mathcal{C}$.

The inverse of $p_{\text{guess}}(\mathcal{C}, p_X)$ represents the typical number of independent trials needed to obtain a correct estimate of the quantity $X(s)$, for a given distribution p_X on the system. We can then define the "min-entropy" as

$$H_{\min}(X|M) = -\log_2(p_{\text{guess}}(\mathcal{C}, p_X)) \tag{6}$$

which represents the size of the register needed to store the data associated with the typical number of independent trials to get a correct estimate. But before showing how this estimation-based entropy connects to information theoretical entropies, let us make things more explicit by discussing the simple example of the noisy estimation of a binary valued quantity.

We consider the bit-flip channel $\mathcal{C}_p$ in which a classical bit is flipped with probability p. We also assume that both values of x are equiprobable. The Bayesian likehood for having x given a result y is then $p_B(x|y) = p(y|x)$ which is $1 - p$ whenever $x = y$ and p whenever $x \neq y$. Consequently, for $p < 1/2$, the maximum likehood algorithm is $X_{\text{ML}}(0) = 0$ and $X_{\text{ML}}(1) = 1$. This enables us to estimate the guessing probabilities: $p_{\text{guess}}(0) = p_{\text{guess}}(1) = 1 - p$ which is the probability for obtaining $y = 0$ (resp. $y = 1$) for $x = 0$ (resp. $x = 1$). Consequently, averaging over x leads to $p_{\text{guess}}(\mathcal{C}_p) = 1 - p$ and thus $H_{\min}(\mathcal{C}_p) = -\log_2(1 - p)$ for $p < 1/2$. In the case $p > 1/2$, then the ML estimation algorithm gives $X_{\text{ML}}(0) = 1$ and $X_{\mathsf{ML}}(1) = 0$. When $x = 0$ is emitted, the probability for having the correct estimation thus corresponds to the case $y = 1$ which occurs with probabiloty p and thus $p_{\text{guess}}(0) = p_{\text{guess}}(1) = p$, which leads to $p_{\text{guess}}\mathcal{C}_p) = p$ or equlivalently $H_{\min}(\mathcal{C}_p) = -\log_2(p)$ for $p > 1/2$.

3. Classical noisy measurements: many-realizations

Note that, so far, we have defined an estimation problem which involves only one realization of the microstate s and therefore, a single value of X and thus only one measurement result. This can naturally be generalized to the case of an N-realizations estimation problem as follows:

- Instead of considering a single use of the noisy communication channel, one considers N uses of the communication channel, fed with the same value of x corresponding to N realizations of S prepared in the same microstate s.

- We can allow for independent measurements, which amounts to replacing M by N independent copies of M, or collective measurements, replacing $M^{\otimes N}$ by M_N acting on the N outputs of the physical noisy channels. This clearly makes sense in quantum theory as we shall see later.

  In the case of classical systems, the latter possibility seems less natural so we will restrict ourselves to $M^{\otimes N}$: starting from $(x, \cdots, x)$, we obtain $\mathbf{m} = (m_1, \cdots, m_N)$ where the m_j are independent random variables, each of them distributed according to $p(m|x)$.

- The estimation algorithm $X_{\text{est,N}}$ takes the N measurement results $(m_1, \cdots, m_N)$ and gives an estimate of x from the N realizations. It may consist of N parallel use of the single realization estimation algorithm and a decision process that computes the final estimation from these N estimation processes, for example through a majority vote. More generally, it may work directly on $(m_1, \cdots, m_N)$ and compute an estimate by exploiting all these values. This is the case of the maximum likelihood algorithm applied to $\mathbf{m}$.

From these, a guessing probability and an estimation score can also be defined which depend on Est_N and on $\mathcal{C}^{\otimes N}$ which describes the parallel use of the imperfect measurement apparatus $\mathcal{C}$.

It is expected that the estimation is much better in the case of a large number of realizations $N \gg 1$. Results in this directions are well known from estimation theory, reviewed in [15, Chapter 11], which focuses on the problem of distinguishing between two different values of x, say 0 and 1, each of them leading to a probability distribution $p_0(m) = p(m|0)$ and $p_1(m) = p(m|1)$ for the measurement results. The problem then becomes, starting from a string of N values $\mathbf{m} = (m_1, \cdots, m_N)$ independently distributed according to the same p_x ($x = 0$ or $x = 1$), to recognize from the statistics of the $\mathbf{m}$ whether they come from $x = 0$ or from $x = 1$.

The total probability of error $P_{\text{err}}^{(N)}$ for the N-realization measurement corresponds to the probability for estimating 0 when $x = 1$ and 1 when $x = 0$ weighted by the actual probabilities for $x = 0$ and $x = 1$. It can be shown that $P_{\text{err}}^{(N)}$ decays exponentially with $N \gg 1$:

$$\min(P_{\text{err}}^{(N)}) \sim 2^{-ND^*} \tag{7}$$

with an exponent given by the Chernoff bound:

$$D^* = -\min_{\lambda \in [0,1]} \left[\log_2 \left(\sum_m p_0(m)^\lambda p_1(m)^{1-\lambda} \right) \right]. \tag{8}$$

Remarkably, the exponent D^* only depends on the probability distributions $p_0(m)$ and $p_1(m)$ that we are trying to discriminate, namely, on the channel itself, and not on the probability $p(x)$ with which $x \in \{0,1\}$ is sent at the input of the channel.

The connexion to information theory quantities can be made through the following alternative expression

$$D^* = D[p_{\lambda^*} \| p_0] = D[p_{\lambda^*} \| p_1] \tag{9}$$

in terms of the Kullback-Leibler divergence [16, Section 2.3.4]

$$D[p\|q] = \sum_m p(m) \log_2 \left(\frac{p(m)}{q(m)} \right) \tag{10}$$

using the probability distribution

$$p_\lambda(m) = \frac{p_0(m)^\lambda p_1(m)^{1-\lambda}}{\sum_m p_0(m)^\lambda p_1(m)^{1-\lambda}} \tag{11}$$

evaluated for λ^* defined by the equation

$$D[p_{\lambda^*} \| p_0] = D[p_{\lambda^*} \| p_1] \, . \tag{12}$$

This optimum corresponds to the maximim likehood estimator for Bayesian probabilities $p(0|\mathbf{m})$ and $p(1|\mathbf{m})$ for the N-realizations experiment.

4. From estimation quality to communication theory

So far, we have defined quality estimators for noisy measurements which quantify the error in the reconstruction of the values of a given observable during a noisy measurement experiment.

But in the spirit of this report, we would like to quantify the amount of information that can be extracted from such a noisy measurement. Information represents the size (in bits) of memory registers that are needed to store some relevant information. For example, the Shannon entropy of a probability distribution [16, Section 2.2]

$$S[p] = -\sum_x p(x) \log_2(p(x)) \tag{13}$$

represents the size of such registers for a faithful compression of the source emitting $\mathbf{x} = (x_1, \cdots, x_M)$ where the x_i are distributed according to p in the limit $M \to +\infty$. In this limit, there exist a coding algorithm Enc sending the $\mathbf{x}$ onto $MS[p]$ bits and a decoding algorithm Est which will have a total guessing probability equal to unity. This is why, as we shall explain in the next section, the compression theorem of Shannon should be viewed as an asymptotic – *i.e* ($M \gg 1$)-shot – error-less result. The classical noisy channel capacity [16, Section 2.5] is a result of the same type: it tells us that the error-less capacity of a noisy communication channel $\mathcal{C}$ connecting an emitter X to a receptor R in the asymptotic limit is $C[\mathcal{C}] = \max_X(I[X,R])$ with I the mutual information.

However, asymptotic quantities such as Shannon entropies are not appropriate in our description of measurements: the experiment is indeed not repeated an infinite number of times; ultimately, it is only performed once. In principle, we should rather consider a single-shot framework, or a N-shot framework where N observers aim at inferring together the value of x emitted by the system.

Therefore, the proper notion of capacity in this context would quantify the number of bits of information which can be recovered from a noisy measurement process. As we have seen in Section II A 2, the estimation process introduces errors, as a result of the noise in the measurement. Formally, an error is a wrong estimation of x. The error probability is therefore $p_{\mathrm{err}}(\mathcal{C}, p_X) = 1 - p_{\mathrm{guess}}(\mathcal{C}, p_X)$ where $p_{\mathrm{guess}}(\mathcal{C}, p_X)$ has been defined in Section II A 2. The error probability depends both on the distribution p_X of input values, as well as on the channel $\mathcal{C}$.

We then define the ϵ-error single shot capacity of the channel as the maximum number of bits that can be retrieved from the output m, in a single shot experiment (up to error ϵ). We therefore imagine that x is drawn from a finite set Ω_X with uniform probability $p_{\Omega_X}(x) = 1/|\Omega_X|$, and consider the maximum of $|\Omega_X|$ such that the inferrence of x fails with error probability $p_{\mathrm{err}}(\mathcal{C}, p_{\Omega_X})$ at most ϵ [14]:

$$C_\epsilon^{(1-\mathrm{shot})}[\mathcal{C}] = \max_{(\mathrm{Est}, \Omega_X) \in A_\epsilon} (\log_2(|\Omega_X|)) \tag{14}$$

in which the maximum is taken over sets Ω_X, and over all estimation algorithms such that the error probability is bounded by ϵ:

$$A_\epsilon = \left\{ (\Omega_X, \mathrm{Est}) \text{ s.t. } p_{\mathrm{err}}(\mathrm{Est}_{|\Omega_X} | \mathcal{C}, p_{\Omega_X}) \le \epsilon \right\} \, . \tag{15}$$

5. Noisy classical measurements: $M > 1$ experimental runs

In the last two subsections, we have considered the case of a single experimental run, performed on N realizations of a given microstate of the system, which classically leads to N identical values x of the observable X we are considering. From a communication-theory point of view, only one message (emitted as x and received as $(m_1, \dots m_N)$ with probability $p(m_1|x) \dots p(m_N|x)$) is sent across the noisy communication channel. This explains why the capacity defined by Eq. (14) and (15) is called a “single-shot” or “single-use” capacity.

On the other hand, the quantities originally introduced by Shannon [64] are obtained in a so-called “asymptotic limit”: a stream of M messages $(x_1, \cdots, x_M)$ is sent into the communication channel and the coding/decoding stage is considered in the limit of $M \to +\infty$. Physically, this corresponds to performing M successive experimental runs: each of them leads to a different micro-state s_j $(1 \le j \le M)$ and thus to a different x_j. When a sin-

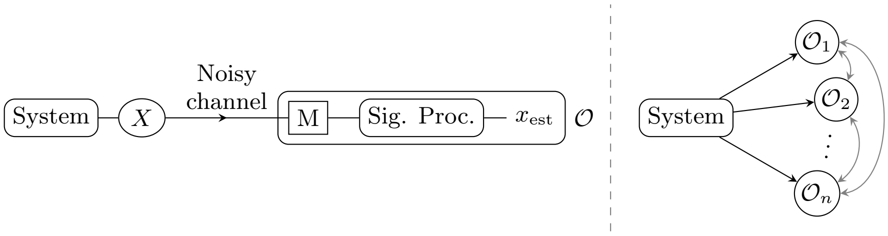

FIG. 2: A Shannon view at noisy classical measurements: (Left Panel) Recasting of measurement by a single observer in terms of a noisy communication channel. The system S is placed on the emitting side of a noisy communication channel into which it injects values of the observable X. The noise represent the experimental noise from the experimental system (not the system) to the imperfection of the measurement apparatus. The decoder combines an ideal noiseless measurement device M and a signal processor Sig. Proc. which together lead to an estimate $x_{\rm est} = {\rm Est}(m)$ for the value of X from the measurement results m. This end of the noisy channel is the classical observer $\mathcal{O}$. Right panel: when several classical observers $\mathcal{O}_1, \ldots, \mathcal{O}_N$ are present, the system is connected to several noisy communication channels (in black). Comparing the result of the post-processed data obtained by the observers is done by using ideal communication channels connecting them (in grey).

gle realization is available for each experimental run[5], we have $\mathbf{m} = (m_1, \cdots, m_M)$ measurement results. Assuming that the experiments are independent, and that the behavior of the channel is the same for all the M experimental runs, the measurement result $\mathbf{m}$ is obtained with probability:

$$p(\mathbf{m}|\mathbf{x}) = \prod_{j=1}^{M} p(m_j|x_j)\,. \tag{16}$$

This is called the $i.i.d$ case (independent and identically distributed) and corresponds to a stationary and memoryless channel. Exactly as in the single use of the channel, an estimation process at the output of the channel enables one to infer $\mathbf{x}_{\rm est}(\mathbf{m})$ from the measurements results $\mathbf{m}$; and to define an error probability as ${\rm Prob}(\mathbf{x}_{\rm est} \neq \mathbf{x})$, minimized over all possible choices of the estimation algorithm. This leads to define the average error probability for the M-use of the channel, which depends on: 1) the distribution $p_X^{(M)}(\mathbf{x})$ over the input messages $\mathbf{x}$; and 2) on the channel $\mathcal{C}$, defined by $p(m|x)$. Exactly as in the previous subsection, a M-shot (or M-use) channel capacity $C_\epsilon^{(M-{\rm shot})}[\mathcal{C}]$ can be defined as $1/M$ times the size, in bits, of the maximum set of $\mathbf{x}$ that can be estimated with error probability below ϵ.

While this inferrence framework (either in a 1-shot, or in a M-shot setting), seems the most natural one to discuss the emergence of a consensus among a collection of observers, it is traditional in studies of quantum Darwinism to focus on Shannon-theory quantities. Discussing the precise and quantitative link between inference theory (with estimators and $P_{\rm guess}$ as central objects), and Shannon theory, goes beyond the scope of this report. In brief, it can then be shown that, in the limit $M \to +\infty$, M-shot quantities can be related to Shannon-theory quantities. For instance, for any $\epsilon > 0$ the M-shot capacity of the noisy channel converges towards the Shannon capacity $C[\mathcal{C}]$. This result connects the finite-use, finite-error capacities, to Shannon's capacities which are error-less in the asymptotic regime. Together with other results such as [35] for example, it establishes Shannon's information theory as the asymptotic limit of M-shots information-theory notions.

To conclude this discussion, let us mention that there exists a vast area of work [35] which aims at studying the large M, finite $\varepsilon > 0$ limits of channel capacities. Reviewing this topic would go far beyond the scope of the present report but, from a physicist's point of view and to get a better understanding of the positioning of Shannon's theory in the landscape of noisy measurement presented here, it is useful to notice that Shannon's compression and noisy channel capacity's theorem are "thermodynamics theorems" ($M \to +\infty$) describing a behavior typical of a phase transition in the thermodynamic limit: the error probability jumps from zero to unity when one tries to compress or transmit at a higher rate than Shannon's rates of compression or transmission. The large but finite M limit deals with "finite size effects", taking into account that there is no threshold below which the error ratio would be strictly zero or above which it would be 1. A pioneer of these finite size effects is Strassen who has showed that the M-shot capacity of a channle is given by [66]:

$$C_\epsilon^{(M-{\rm shot})} \simeq C[\mathcal{C}] + \sqrt{\frac{V_\epsilon[\mathcal{C}]}{M}}\,\Phi^{-1}(\epsilon) + \mathcal{O}\left(\frac{\log_2(M)}{M}\right) \tag{17}$$

in which $\Phi^{-1}(\epsilon)$ is the inverse of the error function, and $V_\epsilon[\mathcal{C}]$ is called the ϵ-channel dispersion [59]. This expression shows how the large M-shot capacity converges, at fixed error rate ε, towards Shannon's noisy channel capacity.

On the other hand, in the $M = 1$ case, we are very far from the asymptotic limit of Shannon's theory. Strictly

[5] Considering M realizations corresponds to changing the dimensionality of the space of measurement results as well as the probability of these, replacing $p(m|x)$ by $p(\mathbf{m}|x) = p(m_1|x)\ldots p(m_N|x)$.

speaking, the proper way to discuss noisy measurements in the case of a single observation of a unique realization of a physical system should rely only on the single-shot version of information theory. As we hav not yet fully explored this path, in the forthcoming sections on quantum observers, we shall introduce Shannon-theory quantities.

6. Many classical observers

The case of several observers can be viewed as a generalization of the situation we have just considered. Here also a single source (the system) is connected to several noisy communication channels in a star topology, but they are allowed to measure different observables: instead of having $\mathcal{C}^{\otimes N}$, we have N different channels $(\mathcal{C}_1, \cdots, \mathcal{C}_N)$ as depicted on the right panel of Fig. 2. And of course, each of them has his own decoding stage, corresponding to a different observer, which we will now discuss.

Note that all the channels are fed with the same s which gives $(X_1(s), \cdots, X_N(s))$, the true values of the observables for the micro-state s. In this case, the goal of the observers is to recover the physicl quantity $x = X(s)$. One could imagine that the obserables X_j are multi-dimensional so that each of the estimation algorithms aim at reconstructing an estimate for x. Therefore, we are left with the question of consensus: can they agree on the same value of x, or, in estimation theory terms, can they have a collective estimation stage that will give them a consensual estimation whose quality can then be assessed along the lines sketched above ?

But one could also imagine that the various observers cannot reconstruct x individually because they don't access all the necessary observables. In such a case, they have to use a collective estimation algorithm. In both cases, they have to exchange information and that's why, in the present framework, we have to endow them with classical perfect communication channels as shown on Fig. 2. The collective estimation algorithm can be delegated to one of the observers, so we don't have to show it on the figure.

Exactly as in the single observer discussed in Section II A 4, the question of how much information can be retrieved via a single use of the whole measurement network with fidelity bounded by ϵ has to be asked. Of course, the information might then be spread across all the observers and then, the total number of bits that could be retrieved is certainly higher when the observers work collectively. Intuitively, this would lead to a single-shot version of a noisy channel Slepian-Wolf problem. Let us recall that the original Slepian-Wolf problem [65] deals with the distributed compression of a classical source and provides a zero-error result in the asymptotic limit [16, Section 2.5.4].

As of March 24, 2023, we haven't dug further along this line but this is certainly worth pursuing. Instead, we will now move forward to the extension of the present framework to the quantum domain, stressing the main differences with the classical case.

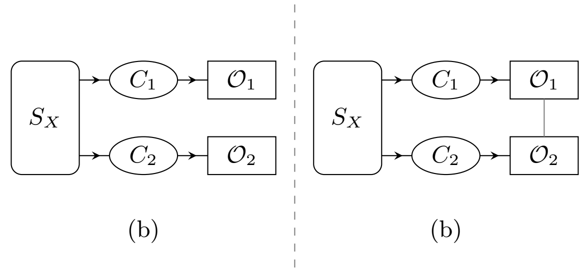

FIG. 3: Two classical observers $\mathcal{O}_1$ and $\mathcal{O}_2$ are measuring the same observable X of the system S through noisy channels C_1 and C_2: (a) the observers are not allowed to communicate (b) $\mathcal{O}_1$ and $\mathcal{O}_2$ are allowed to use a perfect classical communication channel (in grey) and exploit correlations between their imperfect measurements.

B. Quantum observers

1. Definition

A quantum observer is an agent capable of local quantum and classical operations as well as classical or quantum communications. It is realized as a physical subsystem that is macroscopic enough to be partly described as classical, thereby explaning why it has classical memory and computing capabilities as well as classical communication capabilities. But it has also quantum degrees of freedom that can be completly isolated from the rest on demand. The observer can then perform any unitary quantum operation on these degrees of freedom. Moreover, these degrees of freedom can be extended by adding more qubits or using the quantum communications links with the other observers.

Among these degrees of freedom, the quantum measurement apparatus of each observer is directly or indirectly coupled to the system and therefore, each observer can be thought of as having access to a part of the system's environment, exactly as in quantum Darwinism. Consequently, after the phase of pre-measurement, the system as well as the environmental's quantum degree of freedom are in a global entangled state $|\Psi_{\mathrm{tot}}\rangle$. Tracing over the degrees of freedom that are not the system's ones of the quantum observers ones lead to a reduced density operator ρ_{SF} describing the quantum state of the system and of all the quantum observers under consideration.

2. Measurements

Measurement denotes the process by which any observer extracts classical information from the quantum degrees of freedom it has access to.

Here, we will consider that the observer $\mathcal{O}_j$ has access to general measurements which are described by a set of Kraus operators $M_{j,m}$ where $m \in R_j$ denotes the possible results. These operators act only on $\mathcal{H}_j$ and satisfy:

$$\sum_{m\in R_j} M_{j,m}^\dagger M_{j,m} = \mathbb{1}_{\mathcal{H}_j} \,. \tag{18}$$

When the observer $\mathcal{O}_j$ performs such a generalized measurement and obtains the result $m \in R_j$, the state $|\Psi_{\text{tot}}\rangle$ of the system and all the observer's quantum degree of freedom jumps to the relative state with respect to this measurement result m:

$$|\Psi_{\text{tot}}[m]\rangle = \frac{M_{j,m}\,|\Psi_{\text{tot}}\rangle}{\sqrt{\langle M_{j,m}^\dagger M_{j,m}\rangle_{|\Psi_{\text{tot}}\rangle}}} \,, \tag{19}$$

with probability:

$$p\left(m|\,|\Psi_{\text{tot}}\rangle\right) = \langle M_{j,m}^\dagger M_{j,m}\rangle_{|\Psi_{\text{tot}}\rangle} \,. \tag{20}$$

Such a collection of at most one experimental result collected by each observer is called a measurement record. Note that the relative state depends on the full measurement record. The measurement records can be stored in the memories of the observers and then processed classically or communicated among the various observers. We thus have the notion of a quantum trajectory for the full quantum state associated to the experimental records collected by the observers. Note that this notion only depends on the data that the observer is using to retro-compute the quantum trajectory of the system with respect to the records he is considering. Considering all the data collected by all the observers gives access to the most refined quantum trajectory, which is then independent on the observer as soon as they have access to classical communications, and therefore to the same measurement records.

3. *Communication capabilities*

The observers are capable of classical communications but also of quantum communications. This means that they can establish ideal quantum communication channels between them and send parts of their quantum degrees of freedom along these ideal quantum channels. Of course, they have the same capabilities for classical communications channels. Observers can also introduce new quantum degrees of freedom into the game at will and therefore, can also generate on demand maximally entangled qubit pairs (called EPR pairs) between themselves.

Quantum and classical information quantities will be used to quantify the amount of quantum as well as classical communication resources that they can use. We will use the notations of resource calculus [17] (see [16, Chapter 7] for a pedagogical introduction) to describe protocols and relations between them.

4. *Classical image extraction by quantum observers*

For a quantum observer, extracting a classical image of the world consists in performing a generalized measurement on the quantum degrees of freedom it has access to. However, by doing so, the relative state of the other observers and of the system will depend on its measurement result. If all the observers perform such measurements, the relative state of the system and of the environmental degrees of freedom that are not included in any observer will depend on the measurement results. If a set of n quantum observers $\mathcal{O}_{1,\dots,n} = (\mathcal{O}_1,\dots,\mathcal{O}_n)$ perform generalized measurements, we are left with a relative state of the system and other untouched degrees of freedom – the environment of $(S,\mathcal{O}_{1,\dots,n})$ – with respect to the results of these n generalized measurements. Of course, the measurement results are random and therefore, we do not recover an ideal classical measurement. Moreover, generically, the measurement results cannot be described using a classical noise model, which amounts to assume that the correlations between the various measurements obey Bell like inequalities.

The question is precisely to determine under which conditions there is a classical reality to be recovered and, in such a case, how it can be recovered by the observers. The above discussion tells us that this is clearly a far more involved problem than in the classical world because here, it is first not obvious that there is such a classical reality, and, as will be discussed in the present record, it can be encoded in a more subtle way than in the classical world. An important point is also that the answer to at least the second question clearly depends on the choice of generalized measurements by the observers, on their communication capabilities. Obviously it also depends on the global entangled state generated by the quantum pre-measurement interaction.

To strengthen this intuition, let us first discuss more precisely what the quantum observers perspective introduced in the previous subsections can bring us. From this point of view, allowing only local operations and classical exchange of the measurements results (LO) corresponds to what the observers discussed in Section II A 2 were doing: they perform measurements and then compare their results. We shall call them LO-observers.

A more powerful set of capabilities is obtained by allowing the observers to condition the choice of measurements they are performing to the results of other observers. Of course, this does not matter in the classical world since measurement do not alter the state of the system and therefore commute but this leads to non-trivial possibilities in the quantum world because of the existence of incompatible quantum observables. This is the full power of local quantum operation and classical communication (LOCC) which enables them to perform a kind of collective adaptative measurement (see Fig. 4).

By contrast, sharing entangled quantum states on top of being able to perform all LOCC operations enables to do more than with LOCC only. For example, the quan-

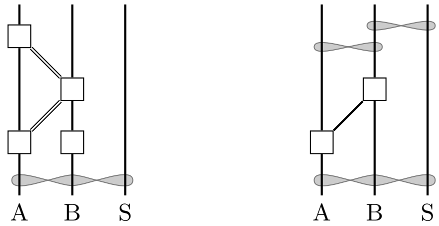

FIG. 4: Circuit representation of operations performed by observers: boxes represent classical measurements, double lines represent classical communication and oblique solid lines are perfect quantum channels. Left part: Alice and Bob who are two quantum observers entangled with a system S, are only using LOCC. Right part: by using quantum communication, Alice and Bob are performing a state transfer concentrating all their shared quantum information on Bob without affecting the system and possibly leaving some entangled pairs free for later use between them.

tum teleportation protocol expresses that sharing entangled pairs enables to simulate quantum communication between two observers. Therefore, a quantum observer entitled with LOCC and shared entanglement with the other observers can use quantum communication protocols (QCPs) to transfer his share of the global quantum state as well as its correlations with the system and with the other observers away from him [2] (see Fig. 4). The quantum observer which is the receiver in such a protocol can then performed generalized measurements on the global state he now has in his hands. This means that he potentially has access to all the correlations that were shared between him and the sender of the quantum state transfer protocol before it was executed. Consequently, by using quantum non-local resources, the observers can access quantum correlations between themselves and even can characterize, in the large number of realization limits, the quantum state they share by transferring correlations to one of them. The problem becomes non trivial in thez case of a finite number of realizations: this is the topic of *quantum estimation theory* which aims at bounding the quality of an estimation of a quantum state given a fixed number of realizations [16, Chapter 4].

Because it enables us to discuss the precise resources that are mobilized for performing certain communication tasks, the framework of quantum Shannon theory applied to quantum observers appears as the proper framework to understand the emergence of a classical picture for physical observers embedded within a quantum universe.

C. Questions to be addressed

1. *What is to be seen vs how it can be seen*

As explained before, the main questions addressed in these notes are: the question of (Q1) understanding which classical "image of the world" (if any) can be accessed to the observers and (Q2) how it can be retrieved. These questions are basically different : the first question involves the structure of the global quantum state for the system and the observers whereas answering the second one is relative to the operational capacities of the observers.

Concerning the first question (Q1) in the perspective of reconstructing a classical reality, the issue of objectivity can be formulated as the existence of a classical quantity, independent of the observers which could, in principle, be accessed by all the observers. This is the question of the emergence of an *objective observable*. Whether or not and how the observers have the capability to access it, to reach a consensus about it, heavily depends on the observers and, therefore, is part of the second question (Q2) we have formulated. This is the question of the *objectivity of observations*.

More precisely, questions about consensus between the observers often implicitly refer to a comparison between reconstructed classical images of the world by each observer. In this case, it implies a comparison between the observer's experimental data. As such it makes sense when considering LO-observers. In the case of LOCC-observers fully exploiting the possibility to design cooperative adaptative measurements, a consensus emerges automatically, almost by definition, from their collective work to reconstruct a classical image of the world. They can share it after their collaborative reconstruction process. This difference between LO and LOCC-observers shows that there is a hierarchy of notions of objectivity depending on the constraints we put on information exchange. As stated in the Introduction, these questions are discussed in Section IV.

In principle, one could also wish to formulate questions (Q1) and (Q2) in the perspective of reconstructing a quantum reality, that is the quantum state of the system. However, one should be careful that the state of the system is not absolute but relative to the observer [26, 27]. Several subtleties must therefore be addressed when considering questions (Q1) and (Q2): the role of the backaction of each measurement performed by each observer on the entangled state of the system must be taken into account and it is not obvious at all that a cooperative reconstruction of the reduced state of the system before any measurement takes place be always possible. Secondly, the notion of objectivity of the reconstructed quantum state has to be specified. Finally, since we expect these questions to be somehow connected to quantum estimation problems, this is directly a problem that deserves being studied for a finite number of realizations, which we think makes it technically not so simple. As of March 24, 2023, this is still left as a perspective and most of this report will address questions (Q1) and (Q2) in the perspective of reconstructing a classical reality.

To address these questions, three point of views will be used throughout these notes, each of them with a different focus:

- Focusing on the structure of the state of the system and the observers will enable us to understand which information they can access. This will be closely related to the nature of correlations between the system and the observers and between the observers themselves. Depending on their capacities, they will be able to extract all, part or maybe none of the information they are looking for on the system.

- We shall use information theoretical quantities to distinguish the various classes of system/observer states and we shall also use information theoretical measures of the abilities of the observers to recover information on the system given their capacities.

- How the observers extract an image of the world, either independently or collectively is an estimation problem. The answer to it depends on the observers capabilities. Of course, it would be interesting to quantify the minimal communication and computation resources needed to perform such a task.

Of course, since information theoretical quantities are notoriously difficult to compute, it might be interesting to address these questions within a more operational framework, that is by considering only specific set of observables accessible to the observers and focusing on correlations between these observables.

2. *Reconstructing an objective classical image*

Let us first discuss the problem of reconstructing an objective classical image. First of all, in quantum theory, classical configurations correspond to perfectly distinguishable quantum states (mutually orthogonal).

Let us assume that such states of the system S are denoted by $|s\rangle$, indexed by s and that the whole environment of S is denoted by E and contains the observers quantum degrees of freedom which are denoted by F. Then, assuming we are dealing with an observable X such that $x = X(s)$ is a coarse-graining of s, reconstructing a classical image means being able to distinguish between $x \neq x'$ by performing measurements on $\mathcal{F}$. In order for this to be possible, we need the state of F relative to the various values of x to be orthogonal. These relative or conditional states are defined by

$$\rho(F|x) = \mathrm{tr}_S\left(\Pi_x \rho_{SF}\right) \tag{21}$$

where Π_x is the projector onto the subspace spanned by all $|s\rangle$ such that $X(s) = x$, and ρ_{SF} denotes the joint state of S and F. Note that in general, $\rho_{SF} = \mathrm{Tr}E \setminus F(\rho SE)$ contains coherences between different values of x and x'. Assuming that no such coherences exist is an extra hypothesis which amount to saying that there are extra quantum degrees of freedom, not in S not in F which select an orthogonal basis for F. This *einselection* [58, 74, 76] is precisely what defines the objectivity of x of S for F. This extra assumption implies that:

$$\rho_{SF} = \sum_x p(x)\Pi_x \otimes \rho(F|x)\,. \tag{22}$$

in which $x \mapsto p(x)$ is a probability distribution and $\rho(F|x)$ denotes a conditional state of F relative to the value of x for S. The orthonormal projectors $\{\Pi_x\}_x$ then define an *objective observable* X for S: it is the same for all the observers in F.

The same definition can be given for a generalized measurement M on S with values x : in this case, there exist mutually orthonormal projectors Π_x, acting on a register space $\mathcal{H}_{E\setminus F}$, and non orthonormal projectors Π_{ψ_x} on $\mathcal{H}_S$ such that

$$\rho_{RSF} = \sum_x p(x)\Pi_x \otimes \Pi_{\psi_x} \otimes \rho(F|x) \tag{23}$$

where R denotes the space of register states (Π_x). Taking the trace over S, the quantum state of RF is then

$$\rho_{RF} = \sum_x p(x)\Pi_x \otimes \rho(F|x) \tag{24}$$

and consequently, we are in the same situation as before except that we have traded S for the register's degrees of freedom. States of the form given by Eqs. (22) and (24) are called *classical–quantum* states. As we shall see in Section III, not all bipartite (S, F) states are of this form.

A decomposition of the form (22) or (24) does not tell us if the various observers in F can identify unambiguously each value of x. For instance, it could be the case that for all x, $\rho(F|x) = \rho_0(F)$ with $\rho_0(F)$ some state independent of x; in this case, the global state would be of the form $\rho_{SF} = \rho_S \otimes \rho_0(F)$, so that the observers degrees of freedom are uncorrelated with the system – no information about x is available in F. But also in the other extreme case where all $\rho(F|x)$ are perfectly distinguishable for different values of x – so that the information about x is, in principle, available in F –, it does not tell us how the observers can, in practice, recover this information. In particular, it does not tell us whether this information about x is available, in principle, to every observer individually, or to all observers only if they share some information they extracted individually, or to all observers only if they perform a global collective measurement on all their degrees of freedom.

This precisely explains why, as suggested in Section II C 1, we first have to distinguish two different notions of objectivity:

- *Objectivity of observables*: different observers probing independent parts of the environment have access to only one observable of the system.

- *Objectivity of outcomes*: different observers probing independent parts of the environment have full access to the above observable and agree on the outcome.

The objectivity of observables states that the environment selects one specific observable of the system (one specific measurement) which may be accessible by performing a generalized measurement on F. In other words, this is the pointer observable of S induced by its environment. The second aspect of objectivity states that, given this pointer observable, all the observers agree on one specific outcome (one specific measurement result). This is the real consensus on memory records among many observers. Of course, this is a much stronger requirement.

As will be discussed in Section IV B, it has been shown [10] that the emergence of an objective observable is inherent to the structure of quantum theory while the emergence of one specific result is dependent on the global ρ_{SF} state and on the observer's capabilities. The first important point is that a single generalized measurement on F can unambiguously distinguish between x and x' if and only if $\rho(F|x) \perp \rho(F|x')$ (namely: $\mathrm{Tr}[\rho(F|x)\rho(F|x')] = 0$. This is the global orthogonality condition which states that quantum observers using fully quantum communication protocols between them can, in principle, cooperatively distinguish between s and $x' \neq x$. As we shall see now, even when the global orthogonality condition is satisfied, it is a non-trivial problem to assess whether more limited observers can distinguish the various values of x.

3. *Observer capabilities*

From a quantum communication point of view, we are dealing with a quantum communication channel that encodes the classical information x using quantum states $\rho(F|x)$ received by the observers. In cryptographic terms, we are encoding a secret. The ability to uncover this secret is a function of the decoder's capability at the reception's end of the quantum communication channel. As explained in the previsous section, introduce four different classes of observers corresponding to different decoding capabilities:

- Fully-Quantum observers (FQ-observers): each observer can apply any local quantum operation and is allowed to use any quantum non local communication resource.

- LOCC-observers: each observer can apply a local quantum operation and classical communication is allowed between the observers. Therefore, global adaptative protocols based on measurements and quantum operations conditioned to classical information is allowed. Equivalently, the outcome $\rho(F|x)$ is split over the F_i's which act as quantum decoders only allowed to use classical communication.

- LO-observers: observers are allowed to apply local operations but classical communication is only allowed for comparison of the results. No adaptative collective measurement is allowed. In cryptographics terms, this corresponds to 1-way decoding by a set of independant decoders allowed to use classical communication only.

Each of these observer classes therefore corresponds to a set of allowed generalized measurements.

- $1-$Obs: finally, we shall consider the ability of a single observer, not allowed to communicate with other observers, to recover x.

4. *Guessing probability and min entropy.*

In order to quantify the ability for the observers to infer x, we introduce the probability P_{guess} that the observer, receiving the quantum state $\rho(F|x)$, correctly "guesses" the actual value of x.

More generally, if the classical variable x is broadcasted through a quantum channel, the oberver receives the conditional or relative quantum state $\rho(F|x)$ with probability p_x as in Eq. (22). Here, the quantum state $\rho(F|x)$ is in general mixed, so that the classical case is recovered if all $\rho(F|x)$ are diagonal in the same basis $|m\rangle$, and can be viewed as classical probability distributions $p(m|x)$ over these states thanks to:

$$\rho_{SF} = \sum_x p(x)\Pi_x \otimes \sum_m p(m|x)\Pi_m \tag{25a}$$

$$= \sum_{m,x} p(m,x)\Pi_x \otimes \Pi_m \ , \tag{25b}$$

where $\Pi_m = |m\rangle\langle m|$. However, note that this is not generally the case: we are left with Eq. (22) in which the relative states $\rho(F|x)$ cannot be diagonalized in a common basis.

In the present setting, the guessing probability is defined by averaging over x a quantum guessing probability for each x, corresponding to the best choise of POVM $\{N_x\}_x$ by the observer:

$$P_{\text{guess}}[\{p_x, \rho(F|x)\}] = \max_{\{N_x\}} \left(\sum_x p_x \,\mathrm{tr}[N_x \rho(F|x)] \right) . \tag{26}$$

We then introduce the "min-entropy" as $H_{\min}(S|F) = -\log P_{\text{guess}}[\{p_x, \rho(F|x)\}]$. In general, as in the classical domain, we have the inequality:

$$H_{\min}(S|F) \leq S(S|F) \tag{27}$$

where the relative or conditional entropy is defined as $S(S|F) = S(\rho_{SF}) - S(\rho_F)$ with S the von Neumann entropy, and ρ_{SF} is the classical-quantum state defined in Eq. (22).

It is worth mentioning that quantum estimation theory is well developed: a quantum Chernoff bound is known [4] as well as the quantum Stein lemma [39, 54]. This enables us to discuss error rates and the notion of a single-shot classical capacity of a quantum channel has also been defined [68]. A forthcoming book by Mark Wide [43] should review this field.

Following the tracks of Section II A for the noisy classical measurements, we conjecture that single-shot quantum information theory provides the proper framework for discussing all the problematics of the emergence of a classical reality for observers conducting one run of general quantum measurements on a single realization of a quantum system. However since we haven't yet followed this track and explored all its consequences, we shall now introduce quantum Shannon's quantities which certainly appear in the limit of a large number M of experiments but which will nevereless be useful to discuss under which conditions the various classes of quantum observers can reconstruct a classical image of the system, or at least of one observable of the system (see Section IV).

5. *Accessible information.*

Anticipating on the following sections, one can define the accessible information for each measurement class X (LO, LOCC, etc.) by

$$I_{\text{acc}}(S, F; X) = \max_{M \in X} (I[S, M]) \tag{28}$$

in which $I[S, M]$ represents the von Neumann mutual information of the reduced density operator after the measurement M has been performed on the fragments F (see Section III). From an information theoretical perspective, this quantity is a kind of "constrained classical capacity" of the quantum channel which has S on its emitting side[6] and a receiver constrained to use a measurement in the X measurement class. The classical information transmitted along this quantum channel consists in the value x considered in Eq. (22). In particular, for the Fully-Quantum observer (FQ), $I_{\text{acc}}(S, F; \text{FQ})$ is nothing but the full accessible information on the F side of the quantum channel that encodes x into $\rho(F|x)$.

For a single observer, we consider any fragment F_i accessed by the observers. Then, the definitions imply that

$$\begin{aligned} I_{\text{acc}}(S, F; \text{FQ}) &\geq I_{\text{acc}}(S, F; \text{LOCC}) \\ &\geq I_{\text{acc}}(S, F; \text{LO}) \qquad (29) \\ &\geq I_{\text{acc}}(S, F_i; 1 - \text{Obs}) \qquad (30) \end{aligned}$$

These accessible informations correspond to the communication capacity of the full communication system, from the system S to the decoding by the F_i's.

[6] The encoding is fixed, hence the adjective *constrained*. This is precisely the definition of the accessible information: the encoding stage is fixed but one can choose the decoding state [16, Section 7.3]. Channel capacities are obtained by maximizing over the encoding stage.

The global orthogonality condition implies that $I_{\text{acc}}(S, F : \text{FQ}) = S[p(x)]$ –namely, there exists a global measurement on F whose outcome identifies a unique value of x–, but it does not imply that the information about x can be accessed through LOCC or LO protocols (see Section IV.

III. CORRELATIONS AND QUANTUM TO CLASSICAL TRANSITION

In the previous section, we have modelled the system as a source of classical information x, encoded into fragments of the environment, which are quantum degrees of freedom. We then considered the task to infer the value of the classical variable x by measurement on the fragments. By construction, in this situation, the system and the fragments only share *classical* correlations. However, in a general situation, also *quantum* correlations, such as quantum entanglement, may be present. The goal of this section is to clarify the distinction between classical and quantum correlations. In order to do so, we focus on information-theoretic quantities whose operational interpretation is rooted in Shannon theory, generalized to quantum systems. This point of view – traditional in the field of quantum Darwinism – is somewhat at variance with the inference point of view of the previous section; and clarifying the relation between these complementary perspectives is an important question for future studies.

In Section III A, we introduce the concept of quantum mutual information, and the concept of "information gained through a measurement". While these two quantities are identical in classical physics, they differ for quantum systems, which leads us to introduce the concept of quantum discord in Section III B, which is a quantitative measure of non-classical correlations between two systems. The vanishing of the quantum discord allows to identify classes of states which have the classical-quantum structure considered in the previous section. Finally, in Section III C, we discuss the role of the quantum discord in the quantum-to-classical transition.

A. Classical vs quantum correlations

We consider a bi-partite quantum system denoted AB (for instance, A is the system, and B is one specific fragment, the community of all fragments, or all the environement of the system). We will discuss the nature of correlations between subsystems A and B. In a classical situation, Alice and Bob are viewed as sources of classical information (encoded in classical variables denotes x for Alice and m for Bob), described by a joint probability distribution $p_{AB}(x, m)$. In the quantum case, they are sources of quantum information encoded in the joint quantum states ρ_{AB}.

1. The mutual information

In classical information theory, correlations between the A and B are quantified by the (classical) mutual information

$$I[A,B] = S[A] + S[B] - S[A,B]\ . \tag{31}$$

In this equation $S[A] = -\sum_x p_A(x)\log p_A(x)$ denotes the Shannon entropy of the marginal probability on A $[p_A(x) = \sum_m p_{AB}(x,m)$, and similarly for $S[B]$, with $p_B(m) = \sum_x p_{AB}(x,m)]$, whereas $S[A,B]$ denotes the Shannon entropy for the joint probability distribution $p_{AB}(x,m)$ for A and B subsystems. Its operational meaning in classical communication theory comes from the result of Slepian Wolf (see Appendix B): $I[A,B]$ corresponds the economy in classical information transmission rate that can be made by using correlations in a distributed compression protocol.

In the quantum case, the same formal definition can be used, replacing Shnnon entropy of the classical distributions p by von Neumann entropy of the quantum states ρ:

$$S[\rho] = -\operatorname{tr}(\rho\log\rho)\,. \tag{32}$$

The corresponding von Neumann mutual information will also be denoted by $I[A,B]$ (we keep the same notation for classical Shannon entropies and quantum von Neumann entropies; the context will make clear the appropriate quantity). The basic properties satisfied by this quantity are recalled in Appendix C. Crucially, although both the Shannon and the von Neumann mutual information have the same lower bound, equal to zero when the two systems are in a product state, they differ by their upper bound which is equal to $\min(S[A],S[B])$ for the Shannon entropy and $2\min(S[A],S[B])$ for the von Neumann entropy.

Figure 5 shows the region, in the $I[A,B]/S[A]$ and $S[B]/S[A]$ variables in which correlations cannot be described classically because the von Neumann mutual information is larger than the upper bound for the classical Shannon information. In this region, one of the two conditional entropies $S[A|B] = S[A,B] - S[A]$ or $S[B|A] = S[A,B] - S[A]$ has to be negative.

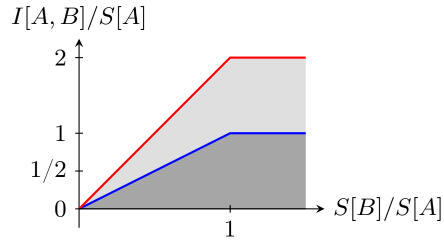

FIG. 5: Bounds for the Shannon and von Neumann mutual informations in units of $S[A]$ in terms of $S[B]/S[A]$. The dark grey zone corresponds to the classical bounds whereas the light grey zone corresponds to values of $I[A,B]/S[A]$ that are not allowed in a classically correlated system and even not in a separable state as shown in Appendix C 3.

2. Information gain through a measurement

Before considering the various types of bipartite states, let us revisit the problem of quantifying the information gained through a measurement. We consider both the classical and quantum situations.

Classical situation. Let us first consider the case of an imperfect measurement in a classical situation –, namely, the situation considered in Section II A 2. Alice prepares subsystem A in the state x with probability $p_A(x)$, and sends them to Bob through a noisy channel. Bob then measures the quantity M_B on the system B, obtained at the output of the channel, thereby obtaining results m which are randomly distributed according to $p_B(m|x)$. From this result m, using Bayes rule, Bob then infers a conditional probability distribution $p_A(x|m) = p_{AB}(x,m)/p_B(m)$ for the state of Alice. The Shannon entropy of this distribution is denoted $S[A|m]$:

$$S[A|m] = -\sum_x p_A(x|m)\log p_A(x|m)\ . \tag{33}$$

Averaging over the measurement results m, the entropy of A conditioned on the measurement M_B performed by Bob is:

$$\begin{aligned}S[A|M_B] &:= \sum_m p_B(m)S[A|m]\\ &= -\sum_{m,x} p_B(m)p_A(x|m)\log p_A(x|m)\\ &= -\sum_{m,x} p_{AB}(x,m)[\log p_{AB}(x,m) - \log p_B(m)]\\ &= S(A,B) - S(B)\end{aligned}$$

Performing the measurement has allowed Bob to reduce his ignorance about Alice's state x from $S[A]$ to $S[A|M_B]$, so that the information gain is:

$$S[A] - S[A|M_B] = S(A) + S(B) - S(A,B)\,. \tag{35}$$

This is exactly the mutual information defined in Eq. (31):

$$I_{\text{gain}}[A;M_B] \equiv I[A,B]\,. \tag{36}$$

If the channel does not introduce noise, then Bob can perfectly infer the value of x from his measurement outcome m. In this case, $S[A|M_B]$ vanishes, and the information gained by the measurement is, by definition, equal to the initial entropy of $p_A(x)$, namely $S(A)$.

Quantum situation. In the quantum situation Alice prepares a bipartite quantum state $\rho^{(0)}_{AB}$, and sends the B subsystem to Bob through the noisy channel. Before performing a measurement, the initial density matrix for Alice and Bob is ρ_{AB}. Bob then performs a generalized measurement (POVM) $M_B = \{M_m\}$, yielding the outcome m. As a result of the measurement, the reduced density operator for Alice is:

$$\rho(A|m) = \frac{\mathrm{Tr}_B(M_m \rho_{AB} M_m^\dagger)}{p_B(m)} . \tag{37}$$

The M_m operators act on $\mathcal{H}_B$, and satisfy $\sum_m M_m^\dagger M_m = \mathbb{1}_B$. The outcome m is obtained with probability $p_B(m) = \mathrm{Tr}[M_m^\dagger M_m \rho_{AB}]$. Notice that the conditional state $\rho(A|m)$ are not necessarily pairwise orthogonal. Similarly to the classical case, Bob's uncertainty about Alice's state is quantified by the conditional entropy:

$$S[A|M_B] = \sum_m p_B(m) S[\rho(A|m)] . \tag{38}$$

The information gained by Bob through the measurement M_B is therefore $I_{\mathrm{gain}}[A; M_B] = S[A] - S[A|M_B]$, where the (unconditioned) reduced state for Alice is $S[A] = S[\rho_A]$ with $\rho_A = \mathrm{Tr}_B(\rho_{AB})$. We have

$$\rho_A = \sum_m p_B(m) \rho(A|m) , \tag{39}$$

as a direct consequence of the expression $p_B(m)\rho(A|m) = \mathrm{Tr}_B(M_m \rho_{AB} M_m^\dagger)$, and of the property $\sum_m M_m^\dagger M_m = \mathbb{1}_B$. An equivalent expression for the information gain is therefore:

$$\begin{aligned} I_{\mathrm{gain}}[A; M_B] = \quad & S\left[\textstyle\sum_m p_B(m) \rho(A|m)\right] \\ & - \textstyle\sum_m p_B(m) S[\rho(A|m)] . \end{aligned} \tag{40}$$

This is exactly the Holevo quantity $\chi[(\rho(A|m), p_B(m))]$ which bounds the amount of information that can be recovered if Alice encodes the values of m distributed according to $p_B(m)$ using mixed states $\rho(A|m)$ [40]. Note that, by concavity of von Naumann entropy, this quantity is always non-negative. Moreover, when the relative states $\rho(A|m)$ are mutually orthogonal, it reduces to the Shannon entropy of the results, as would be expected, since we are dealing with a maximally efficient generalized measurement.

As shall be discussed in Section III B, the information on Alice's state gained through a measurement by Bob, as quantified by Eq. (40), is in general *smaller* than the von Neumann mutual information, while both quantities are equal in the classical case. This discrepancy leads to introduce the *quantum discord.*

3. Bipartite state typology

In order to clarify the origin of the discrepancy between the classical and quantum situations – leading to introduce the quantum discord in the next section –, it is useful to first clarify how classical correlations are described within the quantum framework. This leads us to introduce a hiearachy of four classes of bipartite quantum states: classical–classical, classical–quantum, separable, and entangled states.

Classical–classical states. A lassical–classical (CC) state is a density operator of the form

$$\rho_{AB}^{\mathrm{cc}} = \sum_{x,m} p_{AB}(x, m) \, |x\rangle\langle x| \otimes |m\rangle\langle m| \tag{41}$$

with $(|x\rangle)_x$ is an orthonormal basis for Alice's Hilbert space, and $(|m\rangle)_m$ an orthonormal basis for Bob's Hilbert space, and $p_{AB}(x, m)$ is a probability distribution. For such a state, the quantum mutual information coincides with the Shannon mutual information of the probability distribution $p_{AB}(x, m)$. Consequently, bipartite states with quantum mutual information exceeding the classical upper bound $\min(S[A], S[B])$ cannot be classical–classical. As is shown in Appendix C 3, they are in fact entangled.

Classical–quantum states. Classical–classical states are a special case of classical–quantum (CQ) states:

$$\rho_{AB}^{\mathrm{cq}} = \sum_x p_A(x) \, |x\rangle\langle x| \otimes \rho(B|x) \tag{42}$$

in which $|x\rangle$ an orthonormal basis and $p_A(x)$ defines a probability distribution. CQ states can be interpreted as the result of the following procedure: 1) a measurement M_A, diagonal in the basis $(|x\rangle)_x$, is performed on Alice's side, yielding outcome x with probability $p_A(x)$; 2) conditioned on the outcome x of the measurement, the state $\rho(B|x)$ is sent to Bob.

It is important to notice the asymmetry between Alice and Bob in this definition: a state could be classical–quantum but not quantum–classical. Importantly, the states $\rho(B|x)$ are not necessarily mutually orthogonal. To make manifest the difference with classical–classical states, let us diagonalize each $\rho(B|x) = \sum_m p_B(m;x) \, |m;x\rangle\langle m;x|$ with $p_B(m|x)$ defines a probability distribution for the m variable, and $(|m;x\rangle)_m$ form an orthonormal basis for Bob. Then, we have

$$\rho_{AB} = \sum_{x,m} p_A(x) p_B(x|m) \, |x\rangle\langle x| \otimes |m;x\rangle\langle m;x| \ . \tag{43}$$

This structure looks very similar to classical–classical states [Eq. (41)], except for the fact that the bases $|m;x\rangle)_m$ are *different* for each x, and the complete collection $(|m;x\rangle)_{m,x}$ forms a family of states which are *not* mutually orthoonal.

CC-states, CQ-states, and the accessible information. CC-states and CQ-states can be distinguished via information-theoretic measures of correlations. From the information gain, introduced in Eq. (40), we are led to introduce the notion of *accessible information* (namely, the information about Alice which can be accessed by making measurement on the Bob's subsystem):

$$I_{\text{acc}}(A,B) = \max_{M_B} \; I_{\text{gain}}(A; M_B) \; , \qquad (44)$$

where the max is over all (generalized) measurements M_B which can be performed by Bob. Clearly, the definition is asymmetric between Alice and Bob, and one can also introduce the information about Bob accessible by Alice:

$$I_{\text{acc}}(B,A) = \max_{M_A} \; I_{\text{gain}}(B; M_A) \; , \qquad (45)$$

where M_A are generalized measurements performed by Alice. One can finally introduce a symmetric notion of accessible information, which quantifies the correlations between measurement outcomes on Alice and Bob's subsystems. Indeed, if Alice measures M_A and Bob measures M_B, one can define a joint probability distribution $p_{M_A,M_B}(x,m)$ for obtaining the pair of outcomes (x,m). Therefore, to every choice of measurements (M_A, M_B) corresponds a classical mutual information $I(A,B;M_A,M_B)$ for the corresponding probability distribution $p_{M_A,M_B}(x,m)$ [Eq. (31)]. Maximizing this classical mutul information over all choices of measurements, we define the *classical–classical mutual information*:

$$I_{\text{cc}}(A,B) = \max_{M_A,M_B} \; I[A,B;M_A,M_B] \; . \qquad (46)$$

The accessible information and the classical–classical mutual information are non-increasing under local operations, and are non-negative. The accessible information represents the maximum information about Alice that can be obtained by performing a measurement on Bob's side (or vice-versa), whereas the classical–classical mutual information represents the maximal information which can be obtained by performing measurements on both sides.

These quantites can be used to distinguish classical–classical from classical–quantum states, as stated by the following theorems that we admit (see appendix D):

Theorem A.1. *The state ρ_{AB} is classical–classical if and only if $I(A,B) = I_{\text{cc}}(A,B)$.*

Theorem A.2. *The state ρ_{AB} is classical–quantum if and only if $I(A,B) = I_{\text{acc}}(B,A)$.*

Separable states. Before moving on to discuss quantum discord, let us mention that quantum correlations are not necessarily equivalent to quantum entanglement. General bipartite entangled states are defined as the opposite of separable states and a bipartite state is called separable if it can be written as a statistical mixture of product states:

$$\rho_{AB} = \sum_i p_i \, \sigma_i^A \otimes \sigma_i^B \qquad (47)$$

where the p_i's define a probability distribution. Separable states form a convex set but there is no easy way to determine whether or not a bipartite state is separable or not: in full generality it is an NP-hard problem [33].

Entangled states. Entangled states are exploited in quantum communication protocols such as quantum teleportation and superdense coding. Pure entangled states also exhibit non-local characteristics such as violations of Bell inequalities which express that the correlations they lead to cannot be explained by any model based on local hidden variables [11]. But in the case of statistical mixtures, entangled states are not necessarily showing non-local correlations. The recent review [3] presents a clear discussion of the different degrees of correlation between different systems and an in depth review of the various measures of the quantumness of correlations between two or more subsystems.

B. Quantum discord

1. *Basic definition*

In the previous section, we introduced two measures of correlations between Alice's and Bob's system: 1) the mutual information:

$$I(A,B) = S(A) + S(B) - S(A,B) \; , \qquad (48)$$

where $S = -\text{Tr}(\rho \log \rho)$ is the von Neumann entropy; and 2) the accessible information:

$$I_{\text{acc}}(A,B) = \max_{M_B}[S(A) - S(A|M_B)] \qquad (49)$$

where $S(A|M_B) = -\sum_m p_B(m) S[\rho(A|m)]$, with $\rho(A|m)$ the state of Alice conditioned on the measurement result m obtained by Bob (see Section III A 2). The quantity $I_{\text{gain}}(A;M_B) := S(A) - S(A|M_B)$ is the average information on the state of Alice gained by Bob as a result of his measurement. Maximizing I_{gain} over all possible measurements yields the accessible information. As discussed in Section III A 1, for classical probability distributions $p_{AB}(x,m)$, the accessible information and the mutual information coincide. However, this need not be the case for quantum systems, leading to the definition of the *quantum discord*:

$$D_B(\rho_{AB}) = I(A,B) - I_{\text{acc}}(A,B) \; . \qquad (50)$$

Equivalenlty, with $S(A|B) = S(A,B) - S(B)$ the relative entropy, the discord may be defined as:

$$D_B(\rho_{AB}) = \min_{M_B} S(A|M_B) - S(A|B) \qquad (51)$$

in which the B subscript recalls on which subsystem the measurements are performed. The discord is non-negative, vanishes only for quantum–classical states and represents a measure of quantum correlations.

We can also define the discord relative to a specific measurement M_B. This quantity is denoted

$D_{M_B}(\rho_{AB}) = S(A|M_B) - S(A|B)$, without the minimization over the measurements. It will be useful when discussing Quantum Darwinism in Section III C.

2. Simple examples and bounds

The two simplest examples of discord are obtained by considering a qubit Bell state where we can check that $D_B[\Psi_{AB}] = 1$ and a maximally mixed state of $|00\rangle$ and $|11\rangle$ for which $D_B[\rho_{AB}] = 0$. More generally, for a pure bipartite global state $|\Psi_{AB}\rangle$, then $S[A] = S[B]$ and the mutual information $I[A, B]$ reaches its maximum possible value $2S[A]$. But in this case, the Schmidt decomposition of $|\Psi_{AB}\rangle$ tells us that $S[A]$ is the minimum value $S[\rho_A|M_B]$ over all the possible measurements taken on B. Consequently, in this particular case $D_B[|\psi_{AB}\rangle] = S[B] + S[A] = 2S[A] = I[A, B]$. As expected, all correlations between A and B are of quantum origin.

However, as we shall see later, quantum discord goes beyond simple entanglement to characterize non classical features of quantum states since it can be non zero for separable states. These considerations raise two immediate questions: what are the lower and upper bounds on the quantum discord and what does it mean when they are saturated?

Because $I_{\text{gain}}(A; M_B) \geq 0$, we first obtain that the discord is bounded from above by $I(A, B)$: the quantum part of correlations cannot exceed the total correlation between A and B. Moreover, because of the upper bound $I_{\text{acc}}(A, B) \leq I(A, B)$, the quantum discord is non-negative. In summary:

$$0 \leq D_B(\rho_{AB}) \leq I[A, B] \,. \tag{52}$$

The upper bound is reached whenever all the correlations are of quantum origin as in the simple example of a maximally entangled Bell pair. The lower bound is reached in the case of a QC state defined in Section III A 3.

3. Operational meaning via quantum state merging

In order to be really meaningful in the context of information theory, a quantity should be given an operational meaning, usually related to an information processing task. The purpose of this section is to discuss such a meaning for the discord. We describe here two operational interpretations [12, 51] of the quantum discord using the quantum state merging protocol [41]. Most results here should be understood in the asymptotic regime of independently and identically distributed systems:

$$\overline{D}(\rho_{AB}) = \lim_{n\to+\infty} \frac{D(\rho_{AB}^{\otimes n})}{n} = I(A, B) - \overline{I}_{\text{acc}}(A, B) \,. \tag{53}$$

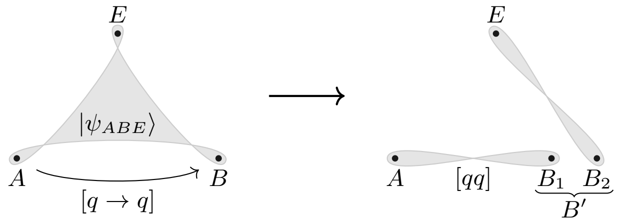

FIG. 6: State transfer protocol: Alice and Bob share a correlated statistical mixture that can be purified using the environment E into a pure state $|\Psi_{ABE}\rangle$. The state transfer protocol consists in transferring the quantum correlations shared between Alice and Bob (see left part), as well as between Alice and the environment in the hands of Bob using only classical and quantum communication between Alice and Bob and doing nothing on E. In the end (right part), $|\Psi_{ABE}\rangle$ is shared between the environment E and some of the qubits B_2 of Bob whereas the protocol can leave us with extra entangled pairs shared between Alice and Bob.

We first begin by reviewing the basics of the state merging protocol.

Classical state merging. We imagine that Alice and Bob have access to classical variables x and m, distributed according to $p_{AB}(x, m)$ (see Fig. 6). How many bits of information does Alice have to send so that Bob can recover the whole information content of x?

Naively, Alice could send $S(A)$ bits of information where S is the Shannon entropy of $p_A(x) = \sum_m p_{AB}(x, m)$. But this is not optimal: by performing a measurement M_B and discovering m, Bob acquires information about x; Bob's lack of information about x is now $S(A|m)$, and $S(A|M_B) == S(A, B) - S(B)$ on average. And indeed, this intuition is correct, as it is possible to use the correlation between x and m to lower the amount of bits Alice has to send: Slepian and Wolf have shown that Alice needs to send exactly $S(A|M_B)$ bits of information.

Quantum state merging. The quantum state merging protocol is the extension of this setup to the quantum case [41]. The quantum version of Slepian and Wolf states that, using a quantum communication channel, Alice can transfer all her share of correlations initially shared with Bob and the environment in Bob hands and that, in the process, Alice and Bob may be left with extra maximally entangled pairs shared together. Initially, Alice and Bob share a quantum state ρ_{AB}, which we purify by introducing the environment E and a tripartite pure state $|\psi_{ABE}\rangle$ shared between Alice, Bob and E. The corresponding ressources that are needed to perform the state transfert are given by

$$\langle\psi_{ABE}\rangle + \frac{1}{2} I[A, E]\,[q \to q] \geq \frac{1}{2} I[A, B]\,[qq] + \langle\psi_{B_2E}\rangle \,. \tag{54}$$

in which the state $\psi_{B_2E}^{\otimes n}$ carries the same information as $\psi_{ABE}^{\otimes n}$ in the large n limit. This inequality, which is a typical resource inequality from quantum Shannon theory [16, Sec. 7.2.4] should be read as follows: given the quantum state ψ shared between Alice, Bob and E and the possibility to use an ideal quantum channel $[q \to q]$ to transfer $I[A,E]/2$ qubits from Alice to Bob, we can arrive at a situation where a state almost identical to ψ_{ABE} is shared between qubits belonging to Bob – denoted here by B_2 – and the environment E and in which there remain $I[A,B]/2$ maximally entangled pairs $[qq]$ between Alice and Bob (using the qubits B_1). Note that these are asymptotic inequalities, meaning that they are realized in the limit of $n \to \infty$ realizations.

The protocol leads to another ressource inequality which quantifies the amount of classical and quantum communication resources that need to be used to achieve the desired result without extra entangled pairs:

$$\langle \psi_{ABE} \rangle + S[A|B]\,[q \to q] + I[A,B]\,[c \to c] \geq \langle \psi_{B_2 E} \rangle\,. \tag{55}$$

but using an ideal classical communication channel $[c \to c]$ to transfer $I[A,B]$ classical bits. This protocol, called the quantum Slepian-Wolf protocol expresses that, given the shared state $|\psi_{ABE}\rangle$, one has to transfer $I[A,B]$ classical bits and $S[A|B]$ quantum bits (when positive!) to Bob to achieve the transwer of this ABE-shared state in the hands of E and B only. If the qubit transfer is realized using quantum teleportation, it means that the state merging protocol consumes $S[A|B]$ maximally entangled pairs $[qq]$.

A conceptual subtlety arises in quantum theory since the von Neumann relative entropy $S[A|B]$ can become negative. In this case, its opposite is called the coherent information [50]. In this case, $S[A|B]\,[q \to q]$ can be put on the r.h.s. of the resource inequality (55) and the operational meaning of the quantum relative entropy is then that the protocol leaves us with $-S[A|B]$ possible future uses of a quantum channel. Since, by the quantum teleportation protocol, an ideal qubit transfer can be emulated by a shared EPR pair and two classical bit transfer, the state transfer protocol can be re-expressed as the state merging protocol resource inequality

$$\langle \psi_{ABE} \rangle + I[A,E]\,[c \to c] \geq -S[A|B]\,[qq] + \langle \psi_{B_2 E} \rangle \tag{56}$$

whenever $S[A|B] < 0$. In this case, the (positive) coherent information has an operational intepretation as the number of distilled maximally entangled pairs in the state merging protocol [41].

First operational meaning of the discord. In this perspective, the first operational definition of the quantum discord is given by the following theorem [51]:

Theorem B.1. *The quantum discord $D_B(\rho_{AB})$ is the minimum increase in the cost of quantum communication for state merging between A and B with a measurement performed on the receiving end B.*

Intuitively, a measurement on B destroys quantum correlations between A and B and consequently increases the cost for A to merge the post-measurement state with B. The rigourous proof is given in Appendix E 1.

This result sheds some light on the properties of the discord. First, measurement on B may lead to a loss of correlations, and therefore increases the price for state merging. This explains why the discord must be positive. Secondly, Bob can at most recover $S[B]$ qubits which is therefore an upper bound on the discord. Finally, the quantum discord is zero for a quantum–classical state of the form

$$\rho_{AB} = \sum_m p_m \rho(A|m) \otimes |m\rangle\langle m| \tag{57}$$

where the states $|m\rangle$ diagonalize ρ_B. Measuring on Bob's side ρ_{AB} in this basis and forgetting the result generates the same state state ρ_{AB}. This means that the measurement causes no loss of information and all the correlations between A and B are preserved in the measurement process.

Second operational meaning of the discord. The second operational meaning of the quantum discord is obtained by extending the standard state merging protocol by considering the preparation of the input state using LOCC and local ancilla which they forget once the preparation is completed. We then have the theorem [12]:

Theorem B.2. *The quantum discord is the total entanglement consumption in the extended state merging protocol:*

$$D(A|C) = S[A|B] + E_F[A,B]\,. \tag{58}$$

in which

$$E_F[A,B] = \min_{(p_i, |\psi_i^{AB}\rangle)} \left[\sum_i p_i S[\mathrm{tr}_A(\psi_i^{AB})] \right] \tag{59}$$

is the minimum amount of entanglement A and B have to use to create the state ρ_{AB} by LOCC. In Eq. (59), the minimum is taken on all the representation of ρ_{AB} as a mixture of the pure states $|\psi_i^{AB}\rangle$. The quantity $E_F(A,B)$ is therefore called the entanglement of formation of the state ρ_{AB}. Thus, $S[A|B]+E_F[A,B]$ quantifies the amount of entanglement the two agents have to consume to first prepare the state ρ_{AB} and then perform the state merging protocol. This-two stage protocol is called the extended state merging protocol. The rigourous proof of this result is given in Appendix E 2.

This formula for the discord illustrates directly its asymmetry. A difference between this characterization and the previous one is that this one refers to one state ρ_{AB} while the previous one referred to two different states, ρ_{AB} and the one measured by Bob. Here, no measurement by Bob explicitly appears.

As a general remark, those two operational characterization of the quantum discord, while fulfilling their purpose, do not really go beyond its mere definition

$D_B(\rho_{AB}) = \min_{M_B}(S[A|M_B]) - S[A|B]$. Indeed, they essentially compare Bob viewed as a purely quantum system with Bob viewed as performing measurements and thus being a source of classical information. In the first vision, we allow for a quantum conditionning of A by B and therefore the proper conditional entropy is $S[A|B]$ whereas in the latter case, only conditioning by classical measurement results is allowed and we have to use $S[\rho_A|M_B]$. In the next section, we will derive a different operational characterization of the quantum discord that does not rely at all from the start on a classical measurement.

C. Quantum Darwinism

In the Quantum Darwinism approach to the emergence of classicality that has been initiated by Zurek [9, 55, 56], the environment is considered as a set of complex systems collecting information on the state of the system. Each observer has then access to a specific fragment or sets of fragments of this environment. The quantum Darwinism approach aims at establishing information theoretical criteria for the emergence of a consensual classical view of the state of the system by the observers. Thus, quantum Darwinism appears to be very close, in the spirit, to the discussion of classical measurement by many observers presented in Section II A 2. We shall therefore review it briefly here and discuss more specifically the quantumness of the correlations that are considered in this approach.

1. *A short review*

Quantum Darwinism considers situations in which a quantum system S is probed by multiple observers having access to independent sets of data through parts of the total environment E of the system as shown on Figure 7 represents schematically the shift on how to model the environment.

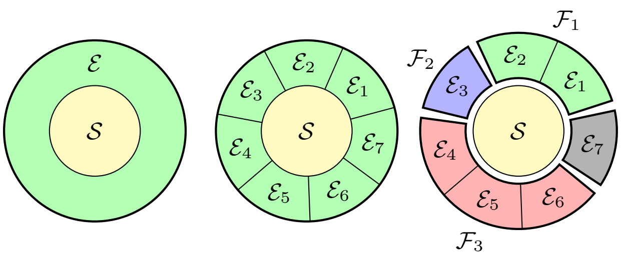

FIG. 7: Different structures for the environment. While standard decoherence models consider the environment as a huge monolithic set of degrees of freedoms, the quantum Darwinism approach focuses on the physics of fragments $\mathcal{F}_i$ of the environment and their relations to the systems. Figure extracted from [75].

This framework is very simular to the one considered in Section II A 2 except that here, the system is also quantum as well as the environmental degrees of freedom each observer can access.

To quantify the ability of one observer monotoring one specific fragment F of the environment, Zurek *et al* consider the von Neumann mutual information $I[S, F]$. In the same way, discussing the consensus between two observers respectively monitoring F_i and F_j is based on the the mutual information $I[F_i, F_j]$.

Let us begin by considering the ability of a given observer to access information on the state of the system. To quantify the classicality of a state , we focus on the region where the mutual information defined by Eq. (31) is close the von Neuman entropy of the system:

$$I[S, F] \geq (1 - \delta) S[S]\,, \tag{60}$$

where the small δ parameter quantifies a possible information deficit that happens in every realistic setup. Thi intuition is that the observer has enough information to reconstruct classical information about the state of the system[7].

An average version of this quantity over the set of fragments of the same size $0 < f \leq 1$ with respect to E is more convenient to have a general view of the behavior of the mutual information. Zurek *et al* thus consider

$$I[S, F_f] = \langle I[S, F] \rangle_{|F|/|E|=f} \tag{61}$$

in which the average is taken over all fragments size f.

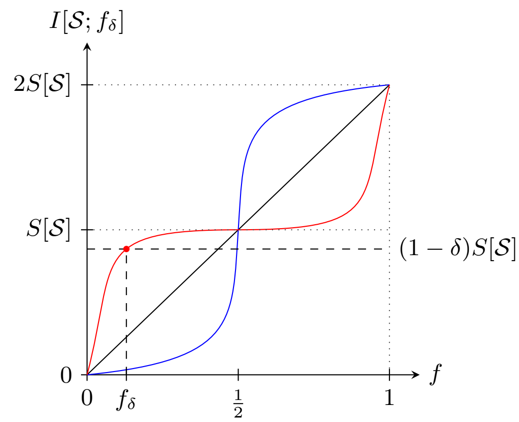

FIG. 8: Two of the typical behaviors of the mutual information $I[S, F_f]$ as a function of the size f of the fragment. The blue curve is the expected behavior for a randomly chosen state in $\mathcal{H}_S \otimes \mathcal{H}_E$ while the red one with its plateau corresponds to a Darwinian situation.

Figure 8 shows the two limiting cases for the behavior of the mutual information. The most interesting case for

[7] The distinction between quantum and classical correlations will be discussed later using the quantum discord.

the problem of the emergence of a consensus is when a plateau around the Von Neumann entropy of the system forms for a very small sized fragment of the environment. Indeed, it means that, by probing a small set of degrees of freedom of the environment, an observer has access to all the classical information about the state of the system. Moreover, other observers can do the same by probing other degrees of freedom. Information accessible to *observers* is information redundantly stored in the environment. This is in this sense that the information is said to be *objective*. The redundancy R_δ is then defined as the inverse of the minimum size reach the inequality (60):

$$R_\delta = \frac{1}{f_\delta} \,. \tag{62}$$

What kind of states maximize this redundancy? Intuitively, pointer states of the system will maximize this since by definition pointer states are states which are robust to the interaction with the environment and can thus "live on" to proliferate their information into many environment channels.

Before coming back on this question, let us spend some time on discussing the nature of the correlations between the system on a fragment F.

2. Discord in quantum Darwinism

As a preliminary, let us recall how to quantify the maximum amount of classical information that can be transmitted though a quantum channel established betwee, say Alice and Bob. Alice encodes the classical messages a emitted with probabilities p_a within not necessarily orthogonal quantum states $\rho(A|a)$, thereby leading to the statistical mixture

$$\rho_A = \sum_a p_a \, \rho(A|a) \,. \tag{63}$$

The amount of classical information that can be retrieved by Bob using classical measurements is bounded from above by the Holevo quantity [40]:

$$\chi[(\rho(A|a), p_a)] = S\left[\sum_a p_a \rho(A|a)\right] - \sum_a p_a S[\rho(A|a)] \,. \tag{64}$$

In quantum Darwinism, the quantity of interest is the total correlation, measured by mutual information, between the system S and a fragment of the environment F, assuming that S is totally decohered by its environment. It is then natural to decompose the mutual information into a "classical component" given by the Holevo quantity and a "quantum component".

This can be done as follows. Let us assume that the system has been decohered by the environment on a basis associated to to a measurement M_S selected by the environment . Consequently, the joint (S, F) state is a classical-quantum state (see Section III A 3):

$$\rho_{SF} = \sum_s p_s \, |s\rangle\langle s| \otimes \rho(F|s) \tag{65}$$

where p_s denote the probability for obtaining the result s. Note that this does not imply that the $\rho(F|s)$ are mutually orthogonal but only that the $\rho(E|s)$ are: information on s can be spread all over E. The non-orthogonality of the $\rho(F|s)$ is precisely what prevents an observer accessing only F to recover all the information about the values s.

The mutual information of this decohered state between S and F can then be obtained as[8]:

$$I[M_S, F] = S\left[\sum_s p_s \rho(F|s)\right] - \sum_s p_s S[\rho(F|s)] \tag{66}$$

which is precisely the Holevo quantity for F prepared in the mixture of the relative states $\rho(F|s)$ with probabilities p_s (see Eq. (64)). We will denote it by $\chi(M_S, F)$. The difference between the initial mutual information $I[S, F]$ and the post-decoherence information $\chi(M_S, F)$ thus coincides with the quantity appearing in Eq. (50) but without optimizing on the choice of the measurement on S. This is the quantum discord [57] relative to the measurement M_S selected by the environment. It obeys [77]:

$$I[S, F] = \chi(M_S, F) + D_{M_S}(\rho_{SF}) \,. \tag{67}$$

The left hand side of (67) doesn't depend on the measurement M_S chosen by the observer while the decomposition does. The information characterized by $\chi(M_S, F)$ is a locally classical accessible information while $D_{M_S}(\rho_{SF})$ corresponds to the global quantum correlations. This can be viewed as a kind of "conservation law" or alternatively as the expression of the objectivity of the total correlation $I[S, F]$ between S and F with respect to the choice of the measurement M_S performed on the system.

We can have a refined understanding of the general form of the mutual information as we vary the size of the fragment when we make the assumption that the global state ρ_{SE} is pure and that the measurement made on the system is projective (a rank one POVM). In this case, we have $D(M_S, E \setminus F) = H_S - \chi(M_S, F)$. Indeed:

$$\begin{aligned} D(M_S, E \setminus F) &= I[S, E \setminus F] - \chi(M_S, E \setminus F) \\ &= S[S] - S[SE \setminus F] + \sum_s p(s) S[\rho(E \setminus F|s)] \\ &= S[S] - S[F] + \sum_s p(s) S[\rho(F|s)] \\ &= S[S] - \chi(M_S, F) \,, \end{aligned}$$

[8] See discussion of [16, Sec.7.5.1].

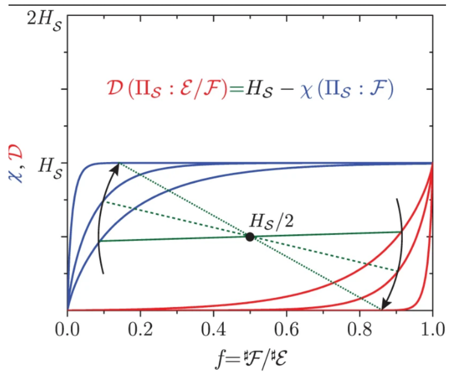

FIG. 9: Anti-symmetry between the quantum discord and the Holevo quantity around $(1/2, S[S]/2)$ for a global pure state and a projective measurement on S. Note: on this figure $H_S = S[S]$.

where we have used, for the second and last lines, the definitions of the mutual information and of the Holevo quantity and, for the third line, that the global state is pure so that $S[SE\backslash F] = S[F]$ as well as the fact that, for a rank one POVM, the relative state is also pure. This equality can also be written in the more symmetrical form

$$\chi(M_S, F) - S[S]/2 = S[S]/2 - D(M_S, E \setminus F) . \quad (68)$$

which geometrically means that we have a symmetry around $S[S]/2$. It shows that the larger the classical information is stored in the fragment F, the less quantum information is stored in the larger fragment $E \setminus F$.

For an arbitrary state ρ_{SE}, we have instead the inequality:

$$D(M_S, E \setminus F) \leq S[S] - \chi(M_S, F) . \quad (69)$$

Indeed, by considering a purification of the state from an additional ancilla R, the previous computation gives the equality $D(M_S, E \setminus F) = S[S] - \chi(M_S, FR)$. Now, from the data processing inequality Eq. (C7), we know that the Holevo bound contracts when we trace over R giving $\chi(M_S, F) \leq \chi(M_S, FR)$, from which we obtain the inequality.

If $\chi(M_S, F) \geq (1 - \delta)H_S$ then $D(M_S, E \setminus F) \leq \delta S[S]$. Thus again the more classical information we have access in F, the less quantum information we have in $E \setminus F$. A world where classical information is present is a world where quantum information is only accessible to global observers.

The first important result is that the Holevo quantity is largest when computed using pointer states and the associated observable than when using another observable. To show this, let us consider the so-called "branching state" which it is said to have the spectrum broadcast structure [42]:

$$|\Psi_{SE}\rangle = \sum_s \sqrt{p_s}\, |s\rangle \otimes \left(\bigotimes_k |\psi(k|s)\rangle \right) , \quad (70)$$

where s indicates the outcomes of a pointer measurment M_S and k denotes the various environmental channels within the environment E. Each fragment F corresponds to a certain set of values of k. Because of the factorized and pure form of $\rho(E|s)$, the state of any fragment of the environment relative to the value s of the pointer observable is pure. Consequently, when considering the Holevo quantity for the statistical mixture of the $\rho(F|s)$ with probability p_s, the Holevo quantity is equal to the entropy of the reduced density operator $\rho_F = \sum_s p_s \rho(F|s)$ because we substract zero in the r.h.s. of (64) since all the $\rho(F|s)$ are pure. By contrast, when considering another observable X of the system and decomposing the global entangled state $|\Psi_{SE}\rangle$ relatively to another observable for the system, the state $\rho(F|x)$ is no longer pure. Consequently $\chi[(\rho(F|x), p_x)] \leq \chi[(\rho(F|s), p_s)]$.

From the conservation law (67), we therefore expect the discord to be minimal when computed relatively to the pointer observable. More precisely, Eq. (67) tells us that the discord has the form

$$D_{M_X}(\rho_{SF}) = D_{M_S}(\rho_{SF}) + \sum_x p_x S[\rho(F|x)] . \quad (71)$$

The second term of the r.h.s. is always positive and vanishes for the pointer basis M_S showing that, for such a choice, the discord is minimum. In the Darwinian case, the $E \setminus F$ part of the environment is large enough to completely decohere (S, F) thereby leading to a classical-quantum state of the form (65) for ρ_{FS}. On the Darwinian plateau, $I[S, F] = S[S]$, thereby leading to $I[S, F] - S[S] = S[F] - S[SF] = 0$. For all practical purposes, in the darwinian case, the discord vanishes and therefore, Eq. (67) leads to:

$$I[S, F] = \chi(M_S, F) = S[F] . \quad (72)$$

In the end, the information relative to a pointer state obtained from a small fragment F quickly saturates the Holevo bound, meaning that only classical information can be obtained from this fragment, even when we enlarge it. Of course, things change at the end of the Darwinian plateau: a global measurement encompassing almost the whole environment can provide more information coming from the hidden quantum correlations characterized by the discord. But in the Darwinian case, one has to go to the end of the plateau – *id est* to very large scales – to see quantum correlations ($I[S, F]$ exceeding $S[S]$).

Note however the very important fact that both equalities in Eq. (72) are necessary and not equivalent. Indeed, the original plateau condition of quantum Darwinism $I[S, F] = S[F]$ is not sufficient to select classically accessible information [42]. Indeed, there exists states forming a Darwinian plateau composed in its majority of quantum discord which means that the information contained in a fragment cannot be classically accessed which is what we should require for a natural notion of independent objectivity. We shall discuss this subtlety in more details in Section IV C. Hence, not only do we have for all practical purposes both equalites in Eq. (72), but we have to impose them both in the ideal situation where each observer can recover with classical means all the classical information about the system.

3. Perfect coding and decoding

Let us assume here that the reduced density operator of the system corresponds to a statistical mixture of mutually orthogonal states $|s\rangle$ with von Neumann entropy $S[S]$. What does it mean if the capacity $I_{\mathrm{qc}}[\rho_{SF}] = \max_{M_F}(I[S, M_F|\rho_{SF}])$ is equal to $S[S]$? The answer is given by the following theorem [56]:

Theorem C.1. *$I_{qc}[\rho_{SF}] = S[S]$ if and only if there exists an observable M_F on F such that $S[S|M_F] = 0$. Moreover, in this case, $S[F|X_S] = 0$ where X_S denotes any observable of S that commutes with ρ_S.*

Proof. We assume that $I_{\mathrm{qc}}[\rho_{SF}] = S[S]$. Consequently, there exists a generalized measurement M_F on F such that $S[\rho_{SF}|M_F] = 0$ since, by definition, $I[S, M_F|\rho_{SF}] = S[S] - S[S|M_F]$. Conversely, if there exist a generalized measurement M_F on F such that $S[S|M_F] = 0$, then $\min_{M'_F}(S[S|M'_F]) \leq 0$. But $S[S|M'_F] \geq 0$ for any generalized measurement on F as we have seen in Section III A 2. Consequently we have $\min_{M'_F}(S[S|M'_F]) = 0$. Therefore $I[\rho_{SF}, M_F] = S[S] - \min_{M'_F}(S[S|M'_F]) = S[S]$.

Let us now turn to the second part of the theorem. Assuming now that $I_{\mathrm{qc}}[\rho_{SF}] = S[S]$, we will show that $S[SF|X_S] = 0$ for any observable of S that commutes with ρ_S. Because measurements performed on S and F commute, we can always specify the action of the generalized measurement on F by its action on each $\rho(S|s)$. The post-measurement reduced density operator for F is thus of the form:

$$\rho_{S,F}^{(\text{post } M_F)}(F|s) = \sum_f p(f|s)\, \rho(F|f,s) \otimes |f\rangle \langle f| \;, \quad (73)$$

in which the f are the possible results of the measurement M_F. Here, $|f\rangle$ denote the auxiliary mutually orthogonal ancillary quantum states that keep the record of the measured value f. Indeed, in the following, we shall include these degrees of freedom within F. Therefore, the total post-measurement density operator is:

$$\rho_{S,F}^{(\text{post } M_F)} = \sum_{(f,s)} p(s)\, p(f|s)\, \rho_S(s) \otimes \rho(F|f,s) \,, \quad (74)$$

in which the $\rho(F|f,s)$ and $\rho(F|f',s')$ are mutually orthogonal as soon as $f \neq f'$. The only way $S[S|M_F]$ could be zero is by assuming that the set of possible values of f is partitioned into disjoint sets J_s, each of them associated with a value of s therfore defining a function that associates to f a given s so that indeed, conditionning to a value of M_F leads to only one possible $\rho_S(s) = |s\rangle \langle s|$. We can then coarse grain the measurement performed on F so that it associates a unique value $\tilde{f}_s$ to each s. This coarse-graining defines a generalized measurement on F which we denote by $\tilde{M}_F$. After this coarse-graining, the correspondence between s and the $\tilde{f}_s$ is a bijection. Moreover, the state ρ_{FS} is indeed classical/classical:

$$\rho_{SF} = \sum_s p(s)\rho_S(s) \otimes \rho(F|\tilde{f}_s) \quad (75)$$

in which the $\rho(S|s)$ are mutually orthogonal as well as the $\rho(F|\tilde{f})$ which are coarsed grained of reduced density operators $\rho(f,s)$, involving non-intersecting sets of values of f for different s, and therefore are mutually orthogonal for $\tilde{f}_s \neq \tilde{f}_{s'}$. Note that for the (S,F) system, measuring S and finding s is then equivalent to measuring on F and finding $\tilde{f}_s$:

$$\rho(FS|s) = \rho(FS|\tilde{f}_s) = \rho(S|s) \otimes \rho(F|\tilde{f}_s)\,. \quad (76)$$

We thus have a perfect classical channel connecting the classical values s and $\tilde{f}_s$. Thus, given any observable X_S on S that admits $|s\rangle$ as eigenvectors, $S[F|X_S] = 0$. We thus have $I[F, X_S|\rho_{SF}] = I[S, \tilde{M}_F|\rho_{SF}] = S[S]$. □

To summarize, in terms of communication theory, when $I_{\mathrm{qc}}[\rho_{SF}] = S[S]$, the system sends some classical information s into its environment and the fragment F is such that there exist a way to decode it without any ambiguity by choosing the appropriate generalized measurement on F's side.

IV. MANY-OBSERVER STRUCTURES

A. Statement of the problems

In this section, we consider the many-observers situation: instead of focusing on one specific fragment of the environment, we consider all of them and try to understand the nature of correlations between the fragments along a given decomposition in parts.

As mentioned in Section II B 4, the question of the emergence of objectivity can be addressed at two different levels:

- *Objectivity of observables*: different observers probing independent parts of the environment have access to only one observable of the system.
- *Objectivity of outcomes*: different observers probing independent parts of the environment have full access to the above observable and agree on the outcome.

The objectivity of observables states that the environment selects one specific observable of the state (one specific measurement) which is accessible to almost all observers. In other words, this is the pointer observable of

S induced by its environment. It can be shown [10] that the emergence of an objective observable is indeed quite generic and inherent to the structure of quantum theory. We will review this work in Section IV B.

The second aspect of objectivity concerns the observers and states that, given this pointer observable, all the observers agree on one specific outcome (one specific measurement result). This is the real consensus on memory records among many observers. Whether or not and how the observers can achieve this depends on the observer's capabilities as explained in Section II C 3.

First, we wan assume that a decomposition into observers is given and ask whether or not, and how the observers can reconstruct the values of the observable. Elaborating on the discussion of Section II B 4, we will introduce a hierarchy of objectivity notions. This amounts to understanding the nature of correlations when reconstruction is possible. It will be discussed in Section IV C.

B. Objectivity of observables

Let us consider a set of subsystems and chose one as our system, calling it S. The rest of them forming its environment are denoted as $F_1, \dots, F_n$. We suppose that S is finite dimensional. We can model the dynamics of S as a completely positive trace preserving (CPTP) map denoted Λ. The fundamental result of Ref. [10] subsequently improved in Ref. [60] is the following[9]:

Theorem B.1. *Let $\Lambda : D(S) \to D(F_1 \otimes \cdots \otimes F_n)$ be a quantum channel and Λ_j the reduced channel to the subsystem F_j. Given an integer $q \geq 1$, there exists a POVM $\{M_x\}_x$ and a set $Q \subseteq \{1, \dots, n\}$ of size q such that for all $j \in \{1, \dots, n\} - Q$:*

$$\|\Lambda_j - \mathcal{E}_j\|_\diamond \leq d_S^3 \sqrt{\frac{2 \ln d_S}{q}} \,, \tag{77}$$

with $\mathcal{E}_j$ defined by

$$\mathcal{E}_j(X) = \sum_x \mathrm{tr}(M_x X) \sigma_{j,x} \,, \tag{78}$$

for states $\sigma_{j,x} \in D(F_j)$ and d_S the dimension of the space S.[10]

The operation $\mathcal{E}_j$ is called a measure-and-prepare map since it can be obtained by F_j first by measuring M_x and then preparing any state $\sigma_{j,x}$ conditioned on the result x. The system F_j can at most recover the information about the measurement $\{M_x\}$.

The remarkable point of the theorem is that the measurement M_x does not depend on the observers F_j. It can be thought as the pointer basis of the interaction Λ. Thus almost all the observers have a dynamics very close to a measure-and-prepare dynamics with the measurement M_x independent of them.

Another way of thinking about it is that the dynamics is close to a situation where the system S is first measured by M_x and only then the classical information is broadcasted and locally degraded into the state $\sigma_{j,x}$ for all the different observers F_j. Notice that the theorem does not say that the quantum states $\sigma_{j,x}$ broadcasted to observer i are perfectly distinguishable ($\mathrm{tr}(\sigma_{j,x}\sigma_{j,x'}) = 0$ if $x \neq x'$). Therefore, in general, the observer j, who only receives the statistical mixture $\mathcal{E}_j(X)$, has only a partial information about the actual value of x.

As a remark, keep in mind that this result states the existence of a least one pointer observable but does not say anything about its uniqueness. Depending on the dynamics, it is still allowed to have different possible pointer observables.

An important point here is that even if all the observers F_j have access to the same measurement M_x, the theorem does not say anything about a consensus around the actual outcome of the measurement. Indeed, we see from Eq. (77) that for each j, the evolution is only close to a mixture of the states $\sigma_{j,x}$. The possibility still exists that F_j can obtain the result x_j while F_i with $i \neq j$ can obtain $x_i \neq x_j$. The objectivity of outcomes is not yet quantified by this result.

Still we can go in this direction and obtain a slightly more refined theorem:

Theorem B.2. *Let $\Lambda : D(S) \to D(F_1 \otimes \cdots \otimes F_n)$ be a quantum channel. Given integers $q, t \geq 1$, there exists a POVM $\{M_x\}$ and a subset $Q \subseteq \{1, \dots, n\}$ of size q such that for any subset $T \subseteq \{1, \dots, n\}$ of size t that is disjoint from Q, we have*

$$\|\Lambda_T - \mathcal{E}_T\|_\diamond \leq d_S^3 \sqrt{\frac{2(\ln d_S)t}{q}} \,, \tag{79}$$

with:

$$\mathcal{E}_T(X) = \sum_x \mathrm{tr}(M_x X) \sigma_{T,x} \,, \tag{80}$$

for states $\sigma_{T,x} \in D(\otimes_{j \in T} F_j)$ and d_S the dimension of the space S.

Note the difference of this result compared to the previous one. Here we can state that the evolution of a collection of systems of size t is close to a measure-and-prepare dynamics with measurement M_x still independent of the subsystems. However, now, we know that after a measurement, the joint state of the ensemble of systems T is one of the $\sigma_{T,x}$ with one outcome x: if the states $\sigma_{T,x}$

[9] Here $D(S)$ denotes the set of density operators on $\mathcal{H}_S$ and Λ is the positive super-operator that defines the quantum quantum channel [16, Secs. 7.1 & 9.5].

[10] The norm $\| \cdot \|_\diamond$ is the widely used "diamond" norm between quantum channels. As its precise definition is not important for our discussion, we refer to [69].

are distinguishable from each other for different values of x, then considering the community T as a whole there is only one value of x that is consistent with the global state. This means that in principle an internal consensus can be reached between the members of the community T. However, this consensus strongly depends on the allowed operations within the community. If we consider individual systems F_i for $i \in T$, they might not contain information about the outcome x and as a result the different observers F_i for $i \in T$ do not necessarily agree on the outcome x.

C. Hierarchy of objectivity

In this section, we focus on the objectivity of outcomes. We assume that our observers $F_1 \dots F_n$ can in principle collectively reach a consensus on the value of X (which we can think of as a measurement outcome). Mathematically, this can be written as $P_{\text{guess}}(X|F_1 \dots F_n) = 1$, i.e., the optimal probability of correctly guessing X by performing a measurement on $F_1 \dots F_n$ is 1. We also recall the definition $H_{\min}(X|F) = -\log P_{\text{guess}}(X|F)$. Note that the procedure for determining X will in general involve all the systems $F_1 \dots F_n$ together. We are interested in determining under what conditions, the observers can locally determine this objective value X?

Following [10], the states we consider in this section have the following form:

$$\rho_{XF_1\dots F_n} = \sum_x p(x)|x\rangle\langle x| \otimes \rho^{(x)}_{F_1\dots F_n} \ . \tag{81}$$

1. Independent objectivity

Here, we consider the setting where the observers are completely independent. We determine the conditions under which each observer can completely determine X. In mathematical terms, this means that for all $i \in \{1,\dots,n\}$, $P_{\text{guess}}(X|F_i) = 1$.

Theorem C.1. *Let $\rho_{XF_1\dots F_n}$ be a state as in* (81)*. Then the following conditions are equivalent:*

1. *$P_{guess}(X|F_i) = 1$ for all $i \in \{1,\dots,n\}$*
2. *$I_{acc}(X,F_i) = S(X)$ for all $i \in \{1,\dots,n\}$*
3. *For all $i \in \{1,\dots,n\}$, there exists an isometry W_i (i.e., $W_i^\dagger W_i = I$) that maps the space F_i to $\bar{X}_i \otimes N_i$, where $\bar{X}_i$ is isomorphic to X, such that*

$$\begin{aligned}&(\bigotimes_{i=1}^n W_i)\rho_{XF_1\dots F_n}(\bigotimes_{i=1}^n W_i^\dagger)\\&= \sum_x p(x)|x\rangle\langle x|_X \otimes \left(\bigotimes_{i=1}^n |x\rangle\langle x|_{\bar{X}_i}\right) \otimes \rho^{(x)}_{N_1\dots N_n} \ .\end{aligned}$$

Proof. We will prove that (3) ⇒ (2) ⇒ (1) ⇒ (3).

For (3) ⇒ (2), the measurement on F_i can be taken to apply W_i followed by a measurement of the register $\bar{X}_i$ in the basis $\{|x\rangle\}$.

For (2) ⇒ (1), we have

$$\begin{aligned}&I_{\text{acc}}(X,F_i)\\&= S(X) - \min_{\substack{\text{measurement } \{M_z\}_z \\ \text{with outcome } Z}} S(X|Z)\\&\leq S(X) - H_{\min}(X|F_i) \ .\end{aligned}$$

So $I_{\text{acc}}(X,F_i) = S(X)$ implies that $H_{\min}(X|F_i) = 0$ which is the same as $P_{\text{guess}}(X|F_i) = 1$.

For (1) ⇒ (3), $P_{\text{guess}}(X|F_i) = 1$ implies that there exists a measurement $\{M_x^i\}_x$ on F_i such that $\mathrm{tr}(M_x^i \rho^{(x)}_{F_i}) = 1$ for all x. This implies that $M^i_{x'}\rho^{(x)}_{F_i} = 0$ for $x \neq x'$ and thus $M^i_{x'}\rho^{(x)}_{F_1\dots F_n} = 0$ for $x \neq x'$. Now we define the isometry $W_i = \sum_x |x\rangle \otimes \sqrt{M_x^i}$. Thus, for any x, we have

$$\begin{aligned}&(\bigotimes_{i=1}^n W_i)\rho^{(x)}_{F_1\dots F_n}(\bigotimes_{i=1}^n W_i^\dagger)\\&= \left(\bigotimes_{i=1}^n |x\rangle \otimes \sqrt{M_x^i}\right)\rho^{(x)}_{F_1\dots F_n}\left(\bigotimes_{i=1}^n \langle x| \otimes \sqrt{M_x^i}\right)\\&= \left(\bigotimes_{i=1}^n |x\rangle\langle x|\right) \otimes \left(\bigotimes_{i=1}^n \sqrt{M_x^i}\right)\rho^{(x)}_{F_1\dots F_n}\left(\bigotimes_{i=1}^n \sqrt{M_x^i}\right) \ .\end{aligned}$$

This proves the desired result. □

The structure of this state has *in fine* a very natural structure and transparently shows its objective nature: each observer has access to the complete information about x and can do it independently of what the others are doing. Independent objectivity does not in itself forbid the existence of a strongly correlated noise, potentially relative to x but requires that each observer can filter x out of it. Note that nothing can *a priori* be said on the hardness of this filtering process, it is just possible in principle.

2. Shared objectivity

It is simple to construct examples of states where independent objectivity does not hold, i.e., information about X is stored in the correlations between the different fragments. Let $X \in \{0,1\}$ be uniformly distributed, and let $F_1,\dots,F_n$ be classical systems with $F_1,\dots,F_{n-1}$ be independent and uniform bits, while $F_n = X \oplus \bigoplus_{i=1}^{n-1} F_i$. Then for $n \geq 2$, we have $I(X,F_i) = 0$ for any $i \in \{1,\dots,n\}$ whereas $I(X,F_1\dots F_n) = S(X) = 1$. In fact, this example shows an even more extreme setting where information is stored in the correlations: any group of at most $n-1$ fragments does not get any information about X. This is a simple example of a well-studied topic in

TABLE I: Hierarchy of objectivity levels for a classical data. The examples illustrate that the objectivity levels we define are all different. The letter ϵ denotes the existence of a family of states (in general with growing dimension) such that the corresponding quantity can be made arbitrarily close to 0.

Assumptions	State	Recovery of X	Objectivity Level
$I_{\text{acc}}(X, F_1 \ldots F_n) = S(X)$	$\forall x \neq x'$, $\rho^{(x)}_{F_1 \ldots F_n} \perp \rho^{(x')}_{F_1 \ldots F_n}$	collectively from $F_1 \ldots F_n$	Collective
$\forall i$, $I_{\text{acc}}(X, F_i) = S(X)$	see Th. C.1	individually from each F_i	$1 -$ Obs
$I_{\text{acc,LO}}(X, F_1 \cdots F_n) = S(X)$	see Th. C.2	via local meas. outcomes	LO
$I_{\text{acc,LOCC}}(X, F_1 \cdots F_n) = S(X)$		via local adaptive meas. outcomes	LOCC

$I_{\text{acc}}(F_1 \ldots F_n) \geq$ Global meas.	$I_{\text{acc,LOCC}}(X, F_1 \cdots F_n) \geq$ Adaptive meas.	$I_{\text{acc,LO}}(X, F_1 \cdots F_n) \geq$ Sharing meas.	$I_{\text{acc}}(F_i)$ Indep. estimation	Example states
$S[X]$	$S[X]$	$S[X]$	$S[X]$	F_i have a copy of X
$S[X]$	$S[X]$	$S[X]$	0	(secret sharing)
$S[X]$	$S[X]$	ϵ	ϵ	(information locking)
$S[X]$	ϵ	ϵ	ϵ	(data hiding)

information theory and cryptography called secret sharing [7].

Considering quantum systems $F_1, \ldots, F_n$, it is natural to consider limited classical communication abilities and ask whether the consensus on the value X may be recovered by the fragments. The first setting that we consider is when each fragment i can perform a measurement $\{M^i_z\}$ obtaining an outcome Z_i. When $Z_1 \ldots Z_n$ are sufficient to recover X, we say that we have local operations (LO) objectivity. Observe that the secret sharing example that we just described does satisfy LO objectivity, and in fact for classical states LO objectivity is equivalent to collective objectivity. But for general quantum states, this is not the case. We now state a result to characterize the states satisfying LO objectivity in the following theorem.

Theorem C.2 (Characterizations of LO-objectivity). *Let $\rho_{XF_1 \ldots F_n}$ be a state as in* (81)*. Then the following conditions are equivalent:*

1. *For all $i \in \{1, \ldots, n\}$, there exist measurements $\{M^i_z\}$ with outcome Z_i such that $P_{guess}(X|Z_1 \ldots Z_n) = 1$*
2. *$I_{acc,LO}(X, F_1 \ldots F_n) = S(X)$*
3. *For all $i \in \{1, \ldots, n\}$, there exists an isometry W_i (i.e., $W_i^\dagger W_i = I$) that maps the space F_i to $\bar{Z}_i \otimes N_i$ such that*

$$
(\bigotimes_{i=1}^n W_i)\rho_{XF_1 \ldots F_n}(\bigotimes_{i=1}^n W_i^\dagger) = \sum_{\substack{z_1 \ldots z_n, z'_1 \ldots z'_n \\ f(z_1, \ldots, z_n) = x \\ f(z'_1, \ldots, z'_n) = x}} p(x)|x\rangle\langle x|_X \otimes \left(\bigotimes_{i=1}^n |z_i\rangle\langle z'_i|_{\bar{Z}_i} \right) \otimes \rho^{(z,z')}_{N_1 \ldots N_n} ,
$$

for some function f.

Proof. We will prove that (3) $\Rightarrow$ (2) $\Rightarrow$ (1) $\Rightarrow$ (3).

For (3) $\Rightarrow$ (2), the measurement on F_i can be taken to apply W_i followed by a measurement of the register $\bar{Z}_i$ in the basis $\{|z\rangle\}$. Given the outcomes $z_1, \ldots, z_n$, the only compatible value of X is $f(z_1, \ldots, z_n)$ and hence $I(X, Z_1 \ldots Z_n) = S(X)$.

For (2) $\Rightarrow$ (1), there exists measurements $\{M^i_z\}_z$ with outcomes Z_i such that

$$
\begin{aligned}
&I_{\text{acc,LO}}(X, F_1 \ldots F_n) \\
&= I(X, Z_1 \ldots Z_n) \\
&= S(X) - S(X|Z_1 \ldots Z_n) \\
&\leq S(X) - H_{\min}(X|Z_1 \ldots Z_n) \ .
\end{aligned}
$$

So $I_{\text{acc,LO}}(X, F_1 \ldots F_n) = S(X)$ implies that $H_{\min}(X|Z_1 \ldots Z_n) = 0$ which implies that the outcomes of measurements $\{M^i_z\}_z$ can be used to guess X perfectly.

For (1) $\Rightarrow$ (3), given the POVMs $\{M^i_z\}_z$, we define the isometry $W_i = \sum_{z_i} |z_i\rangle_{\bar{Z}_i} \otimes \sqrt{M^i_{z_i}}$. As $Z_1, \ldots, Z_n$ can be used to guess X perfectly, let us call f the guessing function that maps $z_1, \ldots, z_n$ to the corresponding guess x. With this notation, we have for any x, $\sum_{z_1 \ldots z_n : f(z_1 \ldots z_n) = x} \text{tr}\left((\otimes_{i=1}^n M^i_{z_i}) \rho^{(x)}_{F_1 \ldots F_n} \right) = 1$. This implies that $(\otimes_{i=1}^n M^i_{z_i}) \rho^{(x)}_{F_1 \ldots F_n} = 0$ whenever $f(z_1 \ldots z_n) \neq x$. As a result,

$$
\begin{aligned}
&(\bigotimes_{i=1}^n W_i)\rho^{(x)}_{F_1 \ldots F_n}(\bigotimes_{i=1}^n W_i^\dagger) \\
&= \left(\bigotimes_{i=1}^n \sum_{z_i} |z_i\rangle \otimes \sqrt{M^i_{z_i}} \right) \rho^{(x)}_{F_1 \ldots F_n} \left(\bigotimes_{i=1}^n \sum_{z_i} \langle z_i| \otimes \sqrt{M^i_{z_i}} \right) \\
&= \sum_{\substack{z_1 \ldots z_n, z'_1 \ldots z'_n \\ f(z_1, \ldots, z_n) = x \\ f(z'_1, \ldots, z'_n) = x}} \bigotimes_{i=1}^n |z_i\rangle\langle z'_i| \otimes \sqrt{M_z} \rho^{(x)}_{F_1 \ldots F_n} \sqrt{M_{z'}} \ ,
\end{aligned}
$$

where we introduced the notation $z = z_1 \ldots z_n$ and $M_z = \bigotimes_{i=1}^n M^i_{z_i}$. By defining $\rho^{(z,z')}_{N_1 \ldots N_n} = \sqrt{M_z} \rho^{(x)}_{F_1 \ldots F_n} \sqrt{M_{z'}}$ we get the desired result. □

Not all states that satisfy collective objectivity satisfy LO-objectivity. In fact, consider the state where F_1 and

F_2 are qubit systems.

$$\rho_{XF_1F_2} = \frac{1}{4} \sum_{\substack{x\in\{0,1\}\\ b\in\{0,1\}}} |x\rangle\langle x|_X \otimes |b\rangle\langle b|_{F_1} \otimes H^b |x\rangle\langle x|_{F_2} H^b \ , \tag{82}$$

where H is the Hadamard transform. Note that from the point of view of F_2, the state of system X is encoded in a basis b chosen at random that F_1 has access to. Because we are encoding with two complementary bases, without knowing the basis, the observer F_2 cannot fully recover X. This is sometimes called conjugate coding [70] and can be quantified by uncertainty relations. Conjugate coding is widely used in quantum information, it is at the basis of quantum key distribution protocols [8]. Encoding with different incompatible basis is also called information locking [18, 29].

Observe that for the state in (82), if F_1 and F_2 are allowed to communicate before performing the measurement, then X can be recovered: F_1 could send the value of b to F_2 who would then perform a measurement in the basis b to extract X. This motivates our second shared objectivity setting: the observers are allowed to communicate classically, and then perform measurements that depend on this communication, obtaining outcomes Z_i and collectively $Z_1 \dots Z_n$ are sufficient to recover X. We say in this case that we have local operations and classical communication (LOCC) objectivity. As the example in (82) shows, LO objectivity can be a strictly stronger requirement than LOCC objectivity. Being adaptive, LOCC operations are much more difficult to characterize, so we do not have for the moment a result about the structure of LOCC objective states, similar to Theorem C.2.

Note that LOCC objective states do not correspond to all states with collective objectivity. In fact, considering

$$\rho_{XF_1F_2} = \frac{1}{2} \sum_{x\in\{0,1\}} |x\rangle\langle x|_X \otimes \rho^{(x)}_{F_1F_2} \ , \tag{83}$$

where $\rho^{(0)}_{F_1F_2} = \frac{\Pi_{\mathrm{sym}}}{\mathrm{tr}(\Pi_{\mathrm{sym}})}$ and $\rho^{(1)}_{F_1F_2} = \frac{\Pi_{\mathrm{antisym}}}{\mathrm{tr}(\Pi_{\mathrm{antisym}})}$, where Π_{sym} is the projector onto the symmetric subspace of $\mathbb{C}^d \otimes \mathbb{C}^d$ (i.e., the span of the vectors of the form $|a\rangle \otimes |b\rangle + |b\rangle \otimes |a\rangle$) and Π_{antisym} is the projector onto the antisymmetric subspace of $\mathbb{C}^d \otimes \mathbb{C}^d$ (i.e., the span of the vectors of the form $|a\rangle \otimes |b\rangle - |b\rangle \otimes |a\rangle$ for $a \neq b$). This is the typical example of a data hiding state that are well-studied in quantum information theory, see [19, 47] and the references therein for more details.

V. CONCLUSION

A. Summary

In this report, we have introduced the notion of a quantum observer network that generalizes the ideal observer network underlying classical relativity. This definition naturally embedded the paradigm shift introduced by quantum Darwinism which is the promotion of the system's environment to an active quantum medium through which the system braodcasts information about its state. This also lead us us to propose an agent-based hierarchy, where different levels of objectivity are defined by the observer's communication capabilities. We have then used this framework to discuss how such a network could reconstruct a classical image in a quantum world. While, as shown by Brandao *et al* [10], the existence of objective classical observable X is quite generic, the conditions under which a consensus can emerge among the observers remained unclear.

We discussed extreme cases where the information about x (the value of the objective observable X) can be perfectly extracted from either: (1) each observer individually (1–Obs level); (2) all observers after sharing their measurement outcomes (LO level); (3) all observers by exchanging classical information and feedback, in order to optimize the reconstruction of x (LOCC level) and, finally, (4) all observers by exchanging quantum information (FQ level). We have discussed information-theory criteria ensuring that such a reconstruction is possible at these various levels, and discussed the structure of the associated correlated quantum state shared by the observers and the system. These results, which are summarized in Table IV C, show that the levels of objectivity form a strict hierarchy:

$$\text{1-Obs-Obj} \subsetneq \text{LO-Obj} \subsetneq \text{LOCC-Obj} \subsetneq \text{FQ-Obj} \, . \tag{84}$$

This should be contrasted to the classical situation, where only two different levels of objectivy exist (namely, the 1-Obs and LO levels).

Several questions remain open: first of all, the description of the shared quantum state for the LOCC class has not been obtained. But most importantly, while he levels of objectivity we have defined correspond to ideal cases, a *measure* of objectivity should also be defined, in order to quantify the "proximity" of a given situation to these levels. Here, different inequivalent approaches should be explored. First, one could, in the spirit of the historical studies on quantum Darwinism by Zurek and coworkers, focus on information-theory quantities such as mutual informations. Operationally, such information-theory quantities are meaningful in situations where one operates on a very large number of identically-prepared systems. Our work points however towards an alternative approach – and probably a more appropriate one; namely: to focus on estimation-theory quantities such as the probability of guessing the state of the system by the observers. This approach has not been explored in the context of quantum Darwinism, but seems appropriate in situations where one operates on a *single* copy of the system. Intuitively, quantum Darwinism aims at clarifying why a fundamentally-quantum world appears classical; and the world is only given in a single copy, not as a collection of infinitely-many identically-prepared systems. In this context, the natural quantification of how objective a given situation is, should be achieved via such estimation-theory concepts.

B. Perspectives: from quantum data hiding to quantum signatures of space-time

On the longer term, many perspectives are suggested by the work done during this Ariadna project. We shall discuss perspectives for different domains: 1) quantum Darwinism; 2) quantum data hiding; 3) statistical inferrence; 4) space applications.

Selection of a unique quantum observable by the environment. To build our hierarchy of objectivity criteria, we considered the system as a source of classical information x encoded into quantum degrees of freedom (or fragments F_i), broadcasted through a medium (formally a quantum channel), and finally measured by different observers. The task for the observers was to infer, collectively or individually, the value of x from their measurement results obtained by measuring the fragments F_i. The root for this framework was found in a theorem by Brandao *et al* [10], who showed that this structure generically emerges from the formalism of quantum mechanics itself, as a result of broadcasting the state of the system to many observers.

A question left open by this theorem is the *uniqueness of the observable* being reconstructed by the observers. It is *a priori* not excluded that the observers can infer the outcomes of different and incompatible quantum observables of the system x, y from their measurements. The precise ability of the environment to single out one observable is left as an open question.

Hiding classical information into quantum correlations. This "classical state broadcasting" framework may be interpreted as a "data-hiding" situation, where x (a classical variable) is encoded into multipartite quantum states $\rho(F|x)$ shared among all the fragments. On a fundamental level, the possibility to sharply distinguish the accessible information within each of the frameworks, from 1-Obs to FQ, is not clear. For instance, it is expected that if both the number of fragments and the local Hilbert space dimensions are finite, it is not possible to hide the information about x only at the FQ level; *incomplete*, yet non-zero, information will always be extractable from, *e.g.* LOCC observers. Increasing the number of observers and the Hilbert space dimension of the fragments could allow one to come arbitrarily close to this ideal situation, but the precise way to achieve this is an open question.

Furthermore, if the cost of exchanging classical or quantum information among the fragments is not free, a trade-off exists between the quality of the reconstruction of x and the resources used to achieve it. Quantifying this trade-off is also an important open question. In the same spirit, the computational cost of the reconstruction in the LO, LOCC and FQ cases remains to be estimated.

Inferring incompatible quantum observables. To further clarify the structure of the many-body states emerging from the Quantum Darwinism framework, an interesting and fundamental side question must be addressed. Indeed, it is *a priori* not excluded that different groups of observers, measuring different groups of fragments, infer the value of incompatible quantum observables of the system (e.g. the position $\hat{X}$ and the velocity $\hat{P}$ of the system). Since $\hat{X}$ and $\hat{P}$ do not commute, Heisenberg inequalities must constrain the possible accuracy of this inferrence process.

Clarifying the role of Heisenberg inequalities in such an inferrence framework is an important open question for future studies. As a matter of fact, this is already relevant for experiments: in [31], it has been demonstrated that a full tomography of a super-conducting qubit can be performed by the monitoring of its fluorescence radiation as well as of the scattering of off-resonant radiation by averaging over all the measurement outcomes of many realizations. The incomptability of the three basic observables of the qubit manifests itself through the statistical properties of the whole set of trajectories (see [30, Chapter 6]). See also [34] for another experiment accesing two incompatible observables of a qubit using two QND measurements and [25] in which two qubits are used to probe a multimode quantum electromagnetic field in the time/frequency domain.

Space applications. We would like to conclude this technical report with a few remarks concerning the potential space applications of our investigations. Clearly, the framework of Quantum Darwinism, where a single source (e.g. a star) brodcasts information through a noisy channel (e.g. the interstellar medium) towards several observers (e.g. telescopes), is precisely the situation of space observations. First, it is clear that exploiting the correlations between different detectors is already a key tool in order to improve the resolution of space observations. At a classical level, this corresponds to the LO-situation we have described.

It is already interesting to notice that the exchange of classical information, such as timestamps of photon detection in coincidence detection experiments, can already reveal interesting quantum properties of the source. For example, Hanbury Brown Twiss (HBT) interferemetry, which was used to measure stellar diameters in the 50s [36], was then originally interpreted using classical wave interferometry [37] but, in the quantum regime, is indeed an example of two-particle quantum interference effect [28]. It was used to demontrate differences between classical and quantum field-theoretical descriptions for the photoelectric effect [13] as well as the quantum nature of the fluorescence light emitted by a single atom [44] thereby revealing the quantum nature of the correlations present in the light emitted by these sources.

However, *quantum* correlations between the detectors could also contain useful information about the source – quantifying this information at a theoretical level is already an interesting perspective for future research. Elaborating over the previous discussion on HBT experiments, this suggests to revisit ideas of *quantum signal processing* [61], put forward in electron quantum optics, within the present framework (clearly, space applications would require exploring these ideas in quantum photonic interferometry).

In a second step, quantifying the resources (in terms of quantum-communication channels) necessary to extract this putative information, will be important. Finally, beyond exploiting quantum correlations to improve the inferrence about a far-away body *viewed as source of classical information*, on may speculate on observing *quantum signatures of the source itself*. Clarifying the role of the interstellar medium in altering such quantum signatures of the source will be a key challenge, together with establishing the limits for our detectors to reconstruct them. Such considerations are especially important concerning black holes, which are considered as the most appropriate objects to probe the interplay of Einstein's gravity and quantum physics, potentially revealing quantum signatures of space-time itself.

Appendix A: Quantum measurements

1. General measurements and POVMs

In quantum theory, the most general form of measurement is described by positive valued operator measurements (POVMs). These involve a sequence of positive semi-definite operators E_m ($\langle\psi|\,|E_m|\,|\psi\rangle \geq O$ for any state vector $|\psi\rangle$) summing to $\mathbf{1}$. The probability for obtaining the result m when the system is prepared in the statistical ensemble described by the density operator ρ is

$$p(m|\rho) = \mathrm{Tr}(\rho E_m)\,. \tag{A1}$$

Because of the conditions imposed on the E_m operators, these probabilities are positive and sum to unity. The key point is that the POVM only gives the probability of obtaining a certain result but does not give the state of the system conditionned to this result: it does not describe the quantum backaction. We need to specify it.

In the most general case, an initial (pre-measurement) state described a density operator ρ is sent onto a relative state described by a conditional density operator

$$\rho(S|m) = \frac{1}{p_m}\sum_k M_{m,k}\rho M_{m,k}^\dagger \tag{A2}$$

in which the $M_{m,k}$ operators obey

$$\sum_k M_{m,k}^\dagger M_{m,k} = E_m\,. \tag{A3}$$

and $p_m = p(m|\rho) = \mathrm{Tr}(\rho E_m)$ denotes the probability for obtaining the result m.

These fully general measurement can be viewed as obtained from the so-called generalized measurement, which are described by Kraus operators M_m such that $\sum_m M_m^\dagger M_m = \mathbf{1}$, by a process of coarse graining. Of course, given a POVM, one can always build Kraus operators by taking $U_m\sqrt{E_m}$ in which U_m is a unitary operator that corresponds to a feedback action on the system determined by the measurement result. It is straightforward to check that $M_m^\dagger M_m = E_m$. But as explained in the previous paragraph, this is not the only possibility.

To understand it, let us come back to the construction of Kraus operators in a purified picture: they are associated with a projective measurmeent in the environment. More precisely, given $|\psi\rangle \in \mathcal{H}_S$, we first consider that the environment is prepared in a state $|\mathcal{E}_0\rangle$ and make the two systems through a unitary evolution operator acting on $\mathcal{H}_S \otimes \mathcal{H}_E$ thereby obtaining $U(|\psi\rangle \otimes |\mathcal{E}_0\rangle)$. We then perform a projective measurement on the environment. In full generality, it is described by a set of orthorgonal projectors Π_m summing to $\mathbf{1}$ in $\mathcal{H}_E$. This leads to a POVM through

$$E_m = \langle\mathcal{E}_0|\, U^\dagger \Pi_m U(|\psi\rangle \otimes |\mathcal{E}_0\rangle)\,. \tag{A4}$$

If the projector Π_m has rank strictly greater than one, we can naturally write down a decomposition for the post-measurement relative state of the form Eq. (A2) by using $|x_{m,k}\rangle$ an orthonormal basis of the projection space of Π_m and defining

$$M_{k,m}\,|\psi\rangle = \langle x_{m,k}|\, U(|\psi\rangle \otimes |\mathcal{E}_0\rangle) \tag{A5}$$

A straigtforward calculation shows that

$$M_{m,k}^\dagger M_{m,k}\,|\psi\rangle = \langle\mathcal{E}_0|\, U^\dagger \Pi_{m,k} U(|\psi\rangle \otimes |\mathcal{E}_0\rangle) \tag{A6}$$

thereby implying that $\sum_k M_{m,k}^\dagger M_{m,k} = E_m$. Defining the post-measurmement state by Eq. (A2) amounts to using a much finer generalized measurement defined by all the $M_{m,k}$ operators and then coarse-graining the (m,k) through the forgetting of k.

In the case where the initial state is pure, this lead to a relative state which is a statistical mixture whereas considering $M_m = U_m\sqrt{E_m}$ can be viewed as the most efficient measurement: no information is forgotten in the result and therefore, the post-measurement relative state is still pure. The maximally efficient general quantum measurements correspond to the ones introduced in Section II B 2.

2. Ideal measurements

Ideal measurmeents are characterized by their reproductibility: performing twice the same measurement without letting the system evolve between them will lead to the same result. In classical physics, because there is no a essentiel back-action on the system, ideal measurements are noiseless measurement: the result is a function of the microscopic state of the system. In quantum physics, reproductible measurements are the projective ones, characterized by a set of orthogonal projectors which sum up $\mathbf{1}$.

Here, we will recall the known results on the information gain through an ideal measurement process. As we shall see, the information gain is always positive but the balance equation for information satisfied in the classical domain is not valid in the quantum realm because of the essential backaction of the measurement on the system [5].

Let us first consider the case of an ideal measurement in the classical domain. A system is prepared in a statistical ensemble characterized by a probability distribution $a \mapsto p_a$ of microstates. We then measure a quantity X, thereby obtaining the values x which are of course randomly distributed. The pre-measurement entropy is the Shannon entropy of this ensemble $S[A] = -\sum_a p_a \ln(p_a)$. After measurement, the average entropy is

$$S[A|X] = \sum_x p_x S[A|x] \tag{A7}$$

where $S[A|x] = S[p_A(\cdot|x)]$ is the Shannon entropy for the conditional probability distribution $p(a|x)$ for As microstates conditionned to the results of the measurement. Consequently, the information gain is

$$\Delta I = S[A] - S[A|X] \tag{A8}$$

which is nothing but the mutual information $I[A, X]$. Since an ideal (noiseless) measurement performs a coarse graining of microstates according to the values of the measured quantity, we have

$$S[A] - S[A|X] = -\sum_x p(x) \ln(p(x)) \tag{A9}$$

which expresses that the information gain is exactly equal to the entropy of the results. This can be viewed as a conservation law for information: the results carry an entropy which corresponds to the information gained on the system.

Let us now discuss quantum ideal measurements, which are projective measurements defined by a set of orthogonal projection operators Π_m summing to $\mathbf{1}$. Starting from an initial density operator ρ, the projective measurement generates the "unread" post-measurement density operator

$$\rho^{(\mathrm{unread})} = \sum_m \Pi_m \rho \, \Pi_m \tag{A10}$$

which corresponds to summing, over all the possible results, the relative density operator conditionned to a specific measurement result

$$\rho(S|m) = \frac{\Pi_m \rho \, \Pi_m}{p(m|\rho)} \tag{A11}$$

weighted by the probability $p(m|\rho)$ of obtaining m:

$$p(m|\rho) = \mathrm{Tr}(\Pi_m \rho) \tag{A12}$$

The post-measurement entropy is then the average of the entropies of these relative density operator:

$$S^{(\mathrm{post})} = \sum_m p(m|\rho) \, S[\rho(S|m)] \tag{A13}$$

which plays the analogue of Eq. (A7). Of course, the Shannon entropy of the measurement results is

$$S^{(\mathrm{results})} = -\sum_m p(m|\rho) \, \ln\left(p(m|\rho)\right) . \tag{A14}$$

The entropy conservation law (A9) is then replaced by three distinct results which are proved and discussed in Ref. [5].

First of all, entropy conservation

$$S\left[\rho^{(\mathrm{unread})}\right] - S^{(\mathrm{post})} = S^{(\mathrm{results})} \tag{A15}$$

which is almost identical to Eq. (A9) except that the initial entropy has been replaced by the post-measurement entropy without reading the results. Consequently, the entropy of the results is not anymore the information gained by through the measurement process.

The second result is that performing a projective meaurement will necessarily kills some coherences and therefore increases the entropy

$$S\left[\rho\right] \geq S\left[\rho^{(\mathrm{unread})}\right] \tag{A16}$$

whihc shows that the information gain $S\left[\rho\right] - S^{(\mathrm{post})}$ is smaller than the entropy of the results. Compared to the classical case, this comes from the erasure of all the information associated with the observables incomatibles with the one that is measured [5].

The third result is that this price is never high enough to make the information gain negative. Finally, we obtain the general bounds

$$0 \leq S[\rho] - S^{(\mathrm{post})} \leq S^{(\mathrm{results})} \tag{A17}$$

Note that, for projective measurements, the information gain through the measurement can be expressed as

$$\Delta I = \chi\left[(p(m|\rho), \rho(S|m))_m\right] - (\Delta S)_{\mathrm{dec}} \tag{A18}$$

in which the first term is the Holevo quantity for the $(p(m|\rho), \rho(S|m))_m$ and $(\Delta S)_{\mathrm{dec}}$ is the entropy increased induced by the decoherence associated with the projective measurement process. This is the price to pay for not reaching the maximal possible information gain, which is here the Shannon entropy of the results.

Appendix B: Operational meaning of the mutual information

The operational meaning of the Shannon mutual information appears in the classical Slepian-Wolf protocol [65]. In this protocol, we are considering two correlated sources of information and we want to find the mainimal information rate that they need to use to transmit their joint content. Ignoring their correlations would require $R_A \geq S[A]$ and $R_B \geq S[B]$ bits of information whereas one can achieve full transmission with $R_A \geq S[A|B]$, $R_B \geq S[B|A]$ and $R_A + R_B \geq S[A,B]$ by taking correlations into account. The gain is therefore $S[A] + S[B] - S[A,B] = I[A,B]$.

In the quantum context, the analogous interpretation is provided with the quantum Slepian-Wolf or state merging protocol[41]: we want to compress a bipartite quantum source (A,B). Although the general solution to this problem is not known, it is known when we allow free classical communication between the parties. The end result is that the full state can be compressed using $S_{\rm vn}[A,B]$ qubits, which is the Schumacher compression rate [63] even in this situation where the source is splitted in two. This can be done by first transfering the information shared between A and B to Bob who already holds the B source. This requires the consumption of $S_{\rm vn}[A|B]$ qubits (when the conditional entropy is positive) and $I_{\rm vn}[A,B]$ classical bits. Then, once Bob has the equivalent of the full composite source in its hands, it can use $S[A,B]$ to transfer the full quantum information to the third partner.

Appendix C: Classical and quantum bounds on entropies

1. Sub-additivity and lower bounds

The conditional information $S[A|B]$ defined as $S[A,B] - S[B]$ satisfies the following inequalities. First of all, both the Shannon and the von Neumann entropies are sub-addidive:

$$S[A,B] \leq S[A] + S[B]\,. \tag{C1}$$

with equality if and only off the two sources A and B are uncorrelated.

The Shannon and von Neumann entropies of a composed system differ by the lower bounds they obey. The Shannon entropy of a composed system is always larger than the Shannon entropy of each of the subsystems

$$S[A,B] \geq \max(S[A], S[B])\,. \tag{C2}$$

whereas the von Neumann entropy obeys the Araki-Lieb inequality

$$S_{\rm vn}[A,B] \geq \left|S_{\rm vn}[A] - S_{\rm vn}[B]\right|\,. \tag{C3}$$

which enables the full system's entropy to be lower than the one of each component.

This also leads to the following bounds on the mutual information $I[A;B] = S[A] + S[B] - S[A,B]$. Both the classical (Shannon) and quantum (von Neumann) mutual informations are positive as a consequence of Eq. (C1) but their upper bounds differ. For the classical case, the mutual information cannot exceed the Shannon entropy of any subcomponent

$$I[A;B] \leq \min(S[A], S[B]) \tag{C4}$$

whereas, it can reach twice that limit in the quantum case

$$I_{\rm vn}[A;B] \leq 2\min(S_{\rm vn}[A], S_{\rm vn}[B])\,. \tag{C5}$$

2. Strong subadditivity

Strong sub-additivity (SSA) is a property for tripartite systems. It is obeyed by the Shannon entropy but also by the von Neumann entropy, for which it is a far less trivial result [49] although simpler proofs have been given recently [53, 62]. It states that

$$S[A,B,C] + S[B] \leq S[A,B] + S[B,C] \tag{C6}$$

which is equivalent to a statement on conditional entropies called the data processing inequality[11]:

$$S[A|B,C] \leq S[A|B]\,. \tag{C7}$$

It thus expresses that forgetting a conditionning can only increase the entropy. This bound translates into a bound on mutual information

$$I[A;(B,C)] \geq I[A;B] \tag{C8}$$

stating that correlations between a smaller set of subsystems can only be smaller than correlations between bigger sets.

The strong-subadditivity can finally be recasted as a statement on conditional mutual information. We first need to define the conditional mutual information by generalizing $I[A;B] = S[A] - S[A|B] = S[B] - S[B|A]$:

$$\begin{aligned} I[A;B|C] &= S[A|C] - S[A|B,C] && \text{(C9a)} \\ &= S[A,C] + S[B,C] - S[A,B,C] - S[C] && \text{(C9b)} \\ &= S[A|C] - S[A|B,C] = I[B,A|C] && \text{(C9c)} \end{aligned}$$

Then the strong-subadditivity expresses nothing but the posititivity of the conditional mutual information: $I[A,B|C] \geq 0$.

[11] See [6] for generalizations.

The strong-subaddivity for the von Neumann entropy can be shown to be equivalent to other important properties. First it is equivalent to the decrease of relative entropy or Kullback-Leibler divergence under a completely positive trace preserving (CPTP) operation. This property, called Monotonicity of quantum relative entropy states that for any CPTP $\mathcal{N}$ and any pair of density operators ρ_1 and ρ_2, we have

$$D[\mathcal{N}(\rho_1)\|\mathcal{N}(\rho_2)] \leq D[\rho_1\|\rho_2] \tag{C10}$$

This is also equivalent to the monotonicity of quantum relative entropy under partial trace. Considering two density operators ρ_1 and ρ_2 for a composed system shared by Alice and Bob, we have:

$$D[\mathrm{Tr}_B(\rho_1))\|\mathrm{Tr}_B(\rho_2)] \leq D[\rho_1\|\rho_2] \tag{C11}$$

Finally, the strong-subadditivity is equivalent to the joint convexity of the quantum relative entropy.

3. Sufficient criterion for entanglement

By definition, bipartite entangled states are states which are not separable, the latter being states whose density operator can be written as (C12). Let us show that for separable states, we have $S[A|B] \geq 0$ and $S[B|A] \geq 0$. This shows that whenever $S[A|B]$ or $S[B|A]$ is negative, the state is entangled.

This statement follows from the concavity of $\rho_{AB} \mapsto S[\rho_{AB}] - S[\mathrm{Tr}_A(\rho_{AB})]$ since, starting from the separable state

$$\rho_{AB} = \sum_i p_i \rho_A(i) \otimes \rho_B(i) \tag{C12}$$

the concavity properties implies that

$$\begin{aligned} S[A|B] &\geq \sum_i p_i S[\rho_A(i) \otimes \rho_B(i) | \rho_B(i)] \\ &\geq \sum_i p_i S[\rho_A(i)] \geq 0\,. \end{aligned} \tag{C13a}$$

Concavity of the quantum conditional entropy directly follows from strong subadditivity through thefollowing argument. Let us introduce $0 \leq \lambda \leq 1$ and ρ_{AB} and ρ'_{AB} two density operators. The statistical mixture $\rho_{AB}(\lambda) = \lambda\rho_{AB} + (1-\lambda)\rho'_{AB}$ is obtained by tracing over an arbitrary qubit, which we call C, the density operator

$$\rho_{ABC}(\lambda) = \lambda\rho_{AB} \otimes \Pi_0 + (1-\lambda)\rho'_{AB} \otimes \Pi_1 \tag{C14}$$

where $\Pi_0 = |0\rangle\langle 0|$ and $\Pi_1 = |1\rangle\langle 1|$. Applying the strong subadditivity property (C6) to $\rho_{ABC}(\lambda)$ directly leads to the concavity of $\rho_{AB} \mapsto S[A|B]$.

Note that the positivity of both conditional entropies is what leads to $I[A,B] \leq \min(S[A], S[B])$ since $I[A,B] = S[A] - S[A|B] = S[B] - S[B|A]$. Whenever $I[A,B]$ exceeds this classical bounds means that $S[A|B]$ or $S[B|A] \geq 0$ and therefore implies that the state is entangled.

However, note that this is not necessarily a necessary condition for entanglement.

4. Behavior under temporal evolution

The quantum mutual information being nothing but the relative quantum entropy or Kullback-Leibler divergence between the state of the full system ρ_{AB} and its marginals ρ_A and ρ_B, property (C10) implies that it decreases under temporal evolution when completely positive trace preserving maps are separately applied ot each part. It also decreases (see Eq. (C8)) when one discards part of one of the subsystem. These results can be viewed as the leak of correlations between two systems into all the degrees of freedom they are connected to.

Appendix D: Proof the Theorem A.2

Proof. The proof relies on two lemma that we admit:

Lemma 0.1. *Given a classical–quantum state* $\rho_{AB} = \sum_x p_x |x\rangle\langle x| \otimes \rho(B|x)$*, we have* $I[A,B] = I_{acc}(A,B) = \chi(p_x, \rho(B|x)$ *where* χ *is the Holevo quantity. Moreover,* $I_{\mathrm{cc}}[A,B] = I[A,B]$ *if and only if the states* $\rho(B|x)$ *are mutually commuting, therefore* ρ_{AB} *being classical–classical.*

Lemma 0.2. *Given a state* ρ_{AB} *and two local maps* Λ_A *and* Λ_B*, if* $I[(\Lambda_A \otimes \Lambda_B)(\rho_{AB})] = I[A,B]$*, then there exists inverse maps* Λ_A^* *and* Λ_B^* *such that* $(\Lambda_A^* \otimes \Lambda_B^*)(\Lambda_A \otimes \Lambda_B)(\rho_{AB}) = \rho_{AB}$*.*

Let us now prove the theorem. First, if the state is classical–classical, the result is straightforward by simply performing ideal measurements diagonal in the same basis as the state. Conversely, suppose that $I[A,B] = I_{\mathrm{cc}}(A,B)$, and consider M_A and M_B optimal measurements, s.t. $I[A,B] = I(A,B; M_A, M_B)$. We define $\rho_{AB}^{\mathrm{cc}} = \rho_{AB}(M_A \otimes M_B)$. Then by the second lemma, we have two inverse measurements M_A^* and M_B^*. Now we can consider the quantum-classical state $\rho_{AB}^{\mathrm{qc}} = \rho_{AB}^{\mathrm{cc}}(M_A^* \otimes \mathbf{1}_B)$. This is thus a QC state such that $I_{acc}(B,A) = I_{\mathrm{cc}}[\rho_{AB}^{\mathrm{qc}}] = I_{\mathrm{cc}}[\rho_{AB}] = I[A,B]$. Thus by the second part of the first lemma we conclude that ρ_{AB} is a CQ state (commutation of the marginals) and again with the first lemma that ρ_{AB} is CC. □

Appendix E: Alternative expressions of quantum discord

1. Proof of theorem B.1

The argument starts with the inequality:

$$S[A|B,C] \leq S[A|B]\,. \tag{E1}$$

This version of strong additivity states that by disregarding the system C, Alice and Bob have to share more information to apply the state merging protocol. The goal of the argument is to relate this increase of the cost of the protocol to the discord.

Extend the Hilbert space so that measurement, that we will denote M_i, can be modeled by a coupling to C. Suppose C to be initially in a pure state $|0\rangle$. In the purified setting, M_i is represented by a unitary operator U acting on C and B. For the moment, we initially have have $S(A,B) = S(A,BC)$. After the evolution, we have $I(A,BC) = I(A',B'C')$. Discarding C', we have $I(A',B) \leq I(A',B'C')$. Apply the state merging protocol. $S(A|B) = S(A|BC) = S(A'|B'C')$. Thus acting on B with a unitary operator though an initially factorized ancilla C doesn't change the cost of the protocol. After forgetting C', we have $I(A',B') \leq I(A,B)$ or $S(A|B) \geq S(A'|B')$.

So the cost on state merging induced by the measurement M_i is given by $D(\rho_{AB}|M_i) = I(A,B) - I(A',B')$. This quantity, which still depends on the measurement performed by B, becomes the quantum discord when our operation is a measurement maximizing $I(A',B')$. To see this, let's compute explicitly $I(A',B')$. We have $\rho'_{AB} = \sum_j p_j \rho_{A|j} \otimes \pi_j$ where π_j is the projector on the subspace of the result j. The unconditioned reduced states for A and B are respectively $\rho'_A = \sum_j p_j \rho_{A|j} = \rho_A$ and $\rho'_B = \sum_j p_j \pi_j$. Then, we have:

$$\begin{aligned} I(A',B') &= S(A') + S(B') - S(A',B') \\ &= S(A') + H(p) - (H(p) + \sum_j p_j S(\rho_{A|j})) \\ &= S(A') - \sum_j p_j S(\rho_{A|j}) \,, \end{aligned} \tag{E2}$$

where $H(p)$ is the Shannon entropy of the probability distribution p_j and the second line is obtained using the mixing property of the von Neumann entropy. After maximization over the measurement M_S, we obtain exactly $J[\rho'_{AB}]$, the asymmetric mutual information Eq. (49), ending the proof of the theorem by recalling the definition Eq. (50) of the discord.

2. Proof of theorem B.2

The monogamy relation [45] for a tripartite pure state $|\psi_{ABC}\rangle$ reads:

$$S[B] = E_F[A,B] + I[B,C_c] \,, \tag{E3}$$

where we used the notation C_c to mean that a measurement has been performed on the subsystem C. We then have:

$$E_F[A,B] = S[B|C_c] = S[A|C_c] \,. \tag{E4}$$

From the definition of the discord in terms of relative entropies, we have:

$$D(A|C) = E_F[A,B] - S[A|C] \,. \tag{E5}$$

Since we prepared a pure state, $S[A|C) = S[A,C] - S[C] = S[B] - S[A,B] = -S[A|B]$. Thus $D(A|C) = S[A|B] + E_F[A,B]$, proving the result.

[1] Scott Aasonson, *It's hard to think when someone hadamards your brain.*

[2] A. Abeyesinghe, I. Devetak, P. Hayden, and A. Winter, *The mother of all protocols: restructuring the quantum information's family tree*, Proc. Roy. Soc. A **465** (2009), 2537.

[3] G. Adesso, T. R. Bromley, and M. Cianciaruso, *Measures and application of quantum correlations*, J. Phys. A: Math. Theor. **49** (2016), 473001.

[4] K. M. R. Audenaert, J. Calsamiglia, R. Muñoz Tapia, E. Bagan, Ll. Masanes, A. Acin, and F. Verstraete, *Discriminating states: The quantum chernoff bound*, Phys. Rev. Lett. **98** (2007), 160501.

[5] R. Balian, *Gain of information in a quantum measurement*, Eur. J. Phys. **10** (1989), 208 – 212.

[6] N.J. Beaudry and R. Renner, *An intuitive proof of the data processing inequality*, Quantum Information and Computation **12** (2012), 432.

[7] Amos Beimel, *Secret-sharing schemes: a survey*, International Conference on Coding and Cryptology, Springer, 2011, pp. 11–46.

[8] Charles H Bennett and Gilles Brassard, *Quantum cryptography: Public key distribution and coin tossing*, arXiv preprint arXiv:2003.06557 (2020).

[9] Robin Blume-Kohout and Wojciech H. Zurek, *Quantum darwinism: Entanglement, branches, and the emergent classicality of redundantly stored quantum information*, Phys. Rev. A **73** (2006), 062310.

[10] Fernando G. S. L. Brandao, Marco Piani, and Paweł Horodecki, *Generic emergence of classical features in quantum darwinism*, Nature Communications **6** (2015), 7908.

[11] Nicolas Brunner, Daniel Cavalcanti, Stefano Pironio, Valerio Scarani, and Stephanie Wehner, *Bell nonlocality*, Rev. Mod. Phys. **86** (2014), no. 2, 419–478.

[12] D. Cavalcanti, L. Aolita, S. Boixo, K. Modi, M. Piani, and A. Winter, *Operational interpretations of quantum discord*, Phys. Rev. A **83** (2011), 032324.

[13] John F. Clauser, *Experimental distinction between the quantum and classical field-theoretic predictions for the photoelectric effect*, Phys. Rev. D **9** (1974), no. 4, 853–860.

[14] R. A. Costa, M. Langberg, and J. Barros, *One-shot capacity of discrete channels*, 2010 IEEE International Symposium on Information Theory, 2010, pp. 211–215.

[15] T. M. Cover and J. A. Thomas, *Elements of information theory*, 2nd edition ed., John Wiley and Sons, Inc., 2006.

[16] P. Degiovanni, N. Portier, C. Cabart, A. Feller, and B. Roussel, *Physique quantique, information et calcul : des concepts aux applications*, Savoirs Actuels, EDP Sciences, January 2020.

[17] I. Devetak, Aram W. Harrow, and A. Winter, *A resource*

framework for quantum shannon theory, IEEE Trans. Inf. Th. **54** (2008), 4587.

[18] David P DiVincenzo, Michał Horodecki, Debbie W Leung, John A Smolin, and Barbara M Terhal, *Locking classical correlations in quantum states*, Physical Review Letters **92** (2004), no. 6, 067902.

[19] David P DiVincenzo, Debbie W Leung, and Barbara M Terhal, *Quantum data hiding*, IEEE Transactions on Information Theory **48** (2002), no. 3, 580–598.

[20] A. Einstein, *Zur elektrodynamik bewegter körper*, Annalen der Physik **322** (1905), 891–921.

[21] ———, *On the relativity principle and the conclusions drawn from it*, Jahrbuch der Radioaktivitat und Elektronik **4** (1907), 411–462.

[22] ———, *Die grundlage der allgemeinen relativitätstheorie*, Annalen der Physik **354** (1916), 769–822.

[23] A. Einstein, *Strahlungs-Emission und Absorption nach der Quantentheorie*, Deutsche Physikalische Gesellschaft **18** (1916).

[24] A. Einstein, *Quatre conférences sur la théorie de la relativité*, 2nd ed., Gauthier-Villars, 1980.

[25] A. Essig, Q. Ficheux, T. Peronnin, N. Cottet, R. Lescanne, A. Sarlette, P. Rouchon, Z. Leghtas, and B. Huard, *Multiplexed photon number measurement*, ArXiv:2001:03217, 2020.

[26] H. Everett, *Relative state formulation of quantum mechanics*, Rev. Mod. Phys. **29** (1957), 454.

[27] Hugh Everett, *The many worlds interpretation of quantum mechanics*, Ph.D. thesis, Princeton university, 1957.

[28] U. Fano, *Quantum theory of interference effects in the mixing of light from phase-independent sources*, Am. J. Phys. **29** (1961), 539.

[29] Omar Fawzi, *Uncertainty relations for multiple measurements with applications*, arXiv preprint arXiv:1208.5918 (2012).

[30] Q. Ficheux, *Quantum trajectories with incompatible decoherence channels*, Ph.D. thesis, Ecole Normale Supérieure (Paris), 2018.

[31] Q. Ficheux, S. Jezouin, Z. Leghtas, and B. Huard, *Dynamics of a qubit while simultaneously monitoring its relaxation and dephasing*, Nature Communications **9** (2018), no. 1, 1926.

[32] Daniela Frauchiger and Renato Renner, *Quantum theory cannot consistently describe the use of itself*, Nature communications **9** (2018), no. 1, 1–10.

[33] S. Guaribian, *Strong np-hardness of the quantum separability problem*, Quantum Information and Computation **10** (2010), 343.

[34] Shay Hacohen-Gourgy, Leigh S. Martin, Emmanuel Flurin, Vinay V. Ramasesh, K. Birgitta Whaley, and Irfan Siddiqi, *Quantum dynamics of simultaneously measured non-commuting observables*, Nature **538** (2016), no. 7626, 491–494.

[35] V.Y.F. Han, *Asymptotic estimates in information theory with non-vanishing error probabilities*, Foundations and trends in theoretical computer science, Now Publishers Inc., 2014.

[36] R. Hanbury Brown and R.Q. Twiss, *A test of a new type of stellar interferometer on sirius*, Nature **178** (1956), 1046.

[37] ———, *Interferometry of the intensity fluctuations in light - i. basic theory: the correlation between photons in coherent beams of radiation*, Proc. Roy. Soc. A **242** (1957), 300.

[38] W. Heisenberg, *Über den anschaulichen inhalt der quantentheoretischen kinematik und mechanik*, Zeitschrift für Physik **43** (1927), 172–198.

[39] F. Hiai and D. Petz, *The proper formula for relative entropy and its asymptotics in quantum probability*, Comm. Math. Phys. **143** (1991), 99.

[40] A. S. Holevo, *Bounds for the quantity of information transmitted by a quantum communication channel*, Problems of Information Transmission **9** (1973), 177.

[41] Michał Horodecki, Jonathan Oppenheim, and Andreas Winter, *Partial quantum information*, Nature **436** (2005), 673 EP –.

[42] R. Horodecki, J. K. Korbicz, and P. Horodecki, *Quantum origins of objectivity*, Phys. Rev. A **91** (2015), 032122.

[43] S. Khatri and M.M. Wilde, *Principle of quantum communication theory: a modern approach*, ArXiv:2011.04672, 2020.

[44] H. J. Kimble, M. Dagenais, and L. Mandel, *Photon antibunching in resonance fluorescence*, Phys. Rev. Lett. **39** (1977), 691.

[45] Masato Koashi and Andreas Winter, *Monogamy of quantum entanglement and other correlations*, Phys. Rev. A **69** (2004), 022309.

[46] F. Laloë, *Can quantum mechanics be considered consistent ? a discussion of frauchinger and renner's argument*, ArXiv:1802.06396, 2018.

[47] Ludovico Lami, Carlos Palazuelos, and Andreas Winter, *Ultimate data hiding in quantum mechanics and beyond*, Communications in Mathematical Physics **361** (2018), no. 2, 661–708.

[48] R. Landauer, *Information is physical*, Physics Today (1991), 23.

[49] Elliott H. Lieb and Mary Beth Ruskai, *Proof of the strong subadditivity of quantum-mechanical entropy*, Journal of Mathematical Physics **14** (1973), 1938–1941.

[50] S. Lloyd, *Quantum mechanical maxwell demon*, Phys. Rev. **A 56** (1997), 3374.

[51] Vaibhav Madhok and Animesh Datta, *Interpreting quantum discord through quantum state merging*, Physical Review A **83** (2011), no. 3, 032323.

[52] H. Minkowski, *Die grundgleichungen für die elektromagnetischen vorgänge in bewegten körpern*, Nachrichten von der Gesellschaft der Wissenschaften zu Göttingen, Mathematisch-Physikalische Klasse **1908** (1908), 53–111.

[53] M.A. Nielsen and D. Petz, *A simple proof of the strong subadditivity inequality*, Quantum Information and Computation **5** (2005), 507.

[54] T. Ogawa and H. Nagaoka, *Strong converse and stein's lemma in quantum hypothesis testing*, IEEE Trans. Inform. Theory **46** (2000), 2428.

[55] Harold Ollivier, David Poulin, and Wojciech H. Zurek, *Objective properties from subjective quantum states: Environment as a witness*, Phys. Rev. Lett. **93** (2004), 220401.

[56] ———, *Environment as a witness: Selective proliferation of information and emergence of objectivity in a quantum universe*, Phys. Rev. A **72** (2005), 042113.

[57] Harold Ollivier and Wojciech H. Zurek, *Quantum discord: A measure of the quantumness of correlations*, Phys. Rev. Lett. **88** (2001), 017901.

[58] J.P. Paz, S. Habib, and Wojciech Hubert Zurek, *Reduction of the wave packet: Preferred observable and decoherence time scale*, Phys. Rev. **D 47** (1993), 488–501.

[59] Y. Polyanskiy, H.V. Poor, and S. Verdu, *Channel coding*

rate in the finite blocklength regime, IEEE Transactions on Information Theory **56** (2010), 2307–2359.
[60] Xiaoliang Qi and Daniel Ranard, *Emergent classicality in general multipartite states and channels*, arXiv preprint arXiv:2001.01507 (2020).
[61] B. Roussel, C. Cabart, G. Fève, E. Thibierge, and P. Degiovanni, *Electron quantum optics as quantum signal processing*, Physica Status Solidi B **254** (2017), 16000621.
[62] M.B. Ruskai, *Another short and elementary proof of strond subadditivity of quantum entropy*, Reports on Mathematical Physics **60** (2007), no. 1.
[63] Benjamin Schumacher, *Quantum coding*, Phys. Rev. A **51** (1995), 2738–2747.
[64] Claude E Shannon and W. Weaver, *Théorie mathématique de la communication*, Retz - CEPL, 1975.
[65] D. Slepian and J. Wolf, *Noiseless coding of correlated information sources*, IEEE Transactions on Information Theory **19** (1973), no. 4, 471–480.
[66] V. Strassen, *Asymptotische abschätzungen in shannons informationstheorie*, Trans. Third Prague Conf. Inf. Theory, 1962, pp. 689–723.
[67] J. von Neumann, *Mathematical foundations of quantum mechanics*, Princeton University Press, 1955.
[68] Ligong Wang and Renato Renner, *One-shot classical-quantum capacity and hypothesis testing*, Phys. Rev. Lett. **108** (2012), 200501.
[69] J. Watrous, *The theory of quantum information*, Cambridge University Press, 2018.
[70] Stephen Wiesner, *Conjugate coding*, ACM Sigact News **15** (1983), no. 1, 78–88.
[71] E.P. Wigner, *Remarks on the mind-body question*, Heinemann (London), 1961.
[72] H. D. Zeh, *On the interpretation of measurement in quantum theory*, Foundations of Physics **1** (1970), 69–76.
[73] Wojciech Hubert Zurek, *Decoherence and the transition from quantum to classical*, Physics Today **44** (1991), 36 – 44.
[74] ——, *Decoherence, einselection and the quantum origins of the classical*, Rev. Mod. Phys. **75** (2003), 715.
[75] ——, *Quantum darwinism*, Nature Physics **5** (2009), 181–188.
[76] Wojciech Hubert Zurek, S. Habib, and J.P. Paz, *Coherent states via decoherence*, Phys. Rev. Lett. **70** (1993), 1187–1190.
[77] Michael Zwolak and Wojciech H. Zurek, *Complementarity of quantum discord and classically accessible information*, Scientific Reports **3** (2013), 1729.