Quantum Enhancement of Randomness Distribution

Raul Garcia-Patron[1], William Matthews[2], and Andreas Winter[3]

1 **Quantum Information and Communication**
Ecole Polytechnique de Bruxelles, CP 165, Université Libre de Bruxelles, 1050 Bruxelles, Belgium
rgarciap@ulb.ac.be

2 **Department of Applied Mathematics and Theoretical Physics**
University of Cambridge, Cambridge CB3 0WA, U.K.
wm266@statslab.cam.ac.uk

3 **ICREA & Física Teòrica: Informació i Fenòmens Quàntics,**
Universitat Autònoma de Barcelona, ES-08193 Bellaterra (Barcelona), Spain
andreas.winter@uab.cat

Abstract

The capability of a given channel to transmit information is, a priori, distinct from its capability to distribute random correlations. Despite that, for classical channels, the capacity to distribute information and randomness turns out to be the same, even with the assistance of auxiliary communication. In this work we show that this is no longer true for quantum channels when feedback is allowed. We prove this by constructing a channel that is noisy for the transmission of information but behaves as a virtual noiseless channel for randomness distribution when assisted by feedback communication. Our result can be seen as a way of unlocking quantum randomness internal to the channel.

1998 ACM Subject Classification E.4 Coding and Information Theory

Keywords and phrases Quantum Shannon theory, noisy channels, capacity, randomness

Digital Object Identifier 10.4230/LIPIcs.TQC.2015.180

1 Summary

Randomness and information are different concepts. We think of information of as that which is sent as a specific message to another person or machine. On the other hand, randomness can be intuitively understood as the outcome of a noisy process. Information and randomness being different concepts, the capability to distribute them over a channel could, a priori, be inequivalent resources. More precisely, the capability to distribute a bit of randomness is a weaker resource than the potential to communicate a bit of information over a channel, because if Alice is capable of distributing a bit of information to Bob over a noisy channel she can also locally generate a pair of correlated bits and transmit one to Bob, generating a bit of shared randomness. Therefore, the capacity $R(\mathcal{E})$ of randomness distribution of a noisy channel $\mathcal{E}$ is in principle higher than that of information communication $C(\mathcal{E})$, i.e., $C(\mathcal{E}) \leq R(\mathcal{E})$.

We may also ask about the capacity of a channel to communicate or distribute randomness when auxiliary classical communication is allowed. For communication, we thus have the capacity of the channel assisted by feedback $C_{\leftarrow}$, the capacity assisted by auxiliary forward communication $C_{\rightarrow}$ and the capacity assisted by two-way classical communication $C_{\leftrightarrow}$. Since

10th Conference on the Theory of Quantum Computation, Communication and Cryptography (TQC 2015).
Editors: Salman Beigi and Robert König; pp. 180–190

Leibniz International Proceedings in Informatics
Schloss Dagstuhl – Leibniz-Zentrum für Informatik, Dagstuhl Publishing, Germany

the auxiliary *forward* communication can be used to communicate by itself, we must subtract the amount of auxiliary forward communication from the gross communication rates in the definitions of the later two quantities. For the distribution of shared randomness we can similarly define rates $R_{\leftarrow}$, $R_{\rightarrow}$, $R_{\leftrightarrow}$, but in this case we must subtract both forward and backward auxiliary communication, as both of these may be used to establish shared randomness by themselves.

In the setting with feedback assistance, the tradeoff between the gross rate of randomness distribution and the rate of feedback allowed was characterised (among many other things) by Ahlswede and Csiszár in [2]. A corollary of their result is that $R_{\leftarrow}(\mathcal{E}) = C(\mathcal{E})$ for classical channels. To our knowledge the only previous work studying the generation of shared randomness in a quantum scenario was the work of Devetak and Winter [6] on the distillation of common randomness from bipartite quantum states. That work considered a static scenario of distillation of randomness from a quantum state already shared between Alice and Bob, where in this manuscript we are interested on a dynamic scenario of randomness distribution over quantum channels.

In section 3 we show that, for general quantum channels $\mathcal{E}$, the entanglement-assisted capacity [12] of $\mathcal{E}$, $C_E(\mathcal{E})$, is an upper bound on the largest of the randomness distribution capacities, $R_{\leftrightarrow}(\mathcal{E})$. Since $C_E(\mathcal{E})$ is equal to $C(\mathcal{E})$ for classical-quantum channels (which include classical channels), this establishes the equality

$$R(\mathcal{E}) = R_{\leftarrow}(\mathcal{E}) = R_{\rightarrow}(\mathcal{E}) = R_{\leftrightarrow}(\mathcal{E}) = C(\mathcal{E})$$

for such channels. A simple argument can be used to show that we also have

$$C(\mathcal{E}) = C_{\leftarrow}(\mathcal{E}) = C_{\rightarrow}(\mathcal{E}) = C_{\leftrightarrow}(\mathcal{E})$$

for such classical-quantum channels, so in this case all eight quantities are then same. When the channel is classical

$$C(\mathcal{E}) = \max_{P_X} I(X:Y), \tag{1}$$

where X and Y are the input and output to a single use of the channel $\mathcal{E}$ with X distributed according to P_X [1].

As opposed to the classical regime, where all capacities turn out to be equal, in the quantum scenario randomness distribution and communication remain equivalent only when we consider unassisted or forward assisted classical communication. That is, for general channels $\mathcal{E}$, we have $C(\mathcal{E}) = R(\mathcal{E}) = C_{\rightarrow}(\mathcal{E}) = R_{\rightarrow}(\mathcal{E})$, as shown in subsection 4.1.

In section 4.2 we show that, for quantum-classical channels $\mathcal{E}$, feedback allows the upper bound in terms of $C_E(\mathcal{E})$ to be achieved, and therefore

$$R_{\leftarrow}(\mathcal{E}) = R_{\leftrightarrow}(\mathcal{E}) = C_E(\mathcal{E}).$$

On the other hand, since quantum-classical channels are entanglement-breaking, a result of Bowen and Nagarajan [3] tells us that $C_{\leftarrow}(\mathcal{E}) = C(\mathcal{E})$, so any quantum-classical channel with $C(\mathcal{E}) < C_E(\mathcal{E})$ also demonstrates a separation $C_{\leftarrow}(\mathcal{E}) < R_{\leftarrow}(\mathcal{E})$. Holevo has shown that there are many such channels [4], and we give an explicit example, where the randomness distribution protocol is noiseless, in subsection 4.3.

2 Definitions

Our definitions in this section are based on those used by Ahlswede and Csiszár in [2], and Devetak and Winter [6].

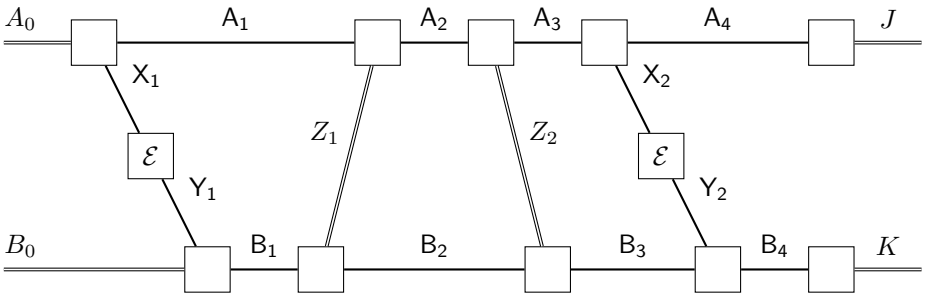

Figure 1 An example of a two-way assisted randomness distillation protocol which makes two uses of the channel $\mathcal{E}$. Time runs left to right. Classical systems are shown as double lines, quantum systems as solid lines. Empty boxes represent local processing. We denote by A_j Alice's system, and by B_j Bob's system, immediately after step j of the protocol. The communication is either forward communication via one use of the noisy channel $\mathcal{E}$, where Alice inputs X_i and Bob receives the output Y_i, or forward/backward auxiliary noiseless classical communication Z_i.

A **two-way assisted** randomness distribution protocol for a channel $\mathcal{E}$ consists of local generation of random variables A_0 and B_0 followed by a finite number of steps, each consisting of communication followed by local processing. The communication is either (i) forward communication via one use of the noisy channel $\mathcal{E}$, where Alice makes an input X_i and Bob receives the output Y_i; (ii) forward auxiliary noiseless classical communication; (iii) backward auxiliary noiseless classical communication.

Suppose we have a protocol of $n+m$ steps where n of the steps are of type (i) and the other m steps are of type (ii) or (iii). We denote by A_j Alice's system, and by B_j Bob's system, immediately after step j of the protocol. At the end of the protocol, Alice must produce random variable J and Bob must produce K, both of which take values in the same alphabet $\mathcal{A}_K$, by local processing of their respective final systems A_{m+n} and A_{m+n}. An example of such a protocol with $n = m = 2$ is illustrated in Figure 1.

We require that

$$\log |\mathcal{A}_K| \leq \exp(cn) \tag{2}$$

for some constant c independent of n (but depending on the channel $\mathcal{E}$). We say that the protocol is ϵ**-good** if $\Pr(J \neq K) \leq \epsilon$. By Fano's inequality and (2), an ϵ-good protocol has

$$H(K|J) \leq \epsilon cn + 1 \tag{3}$$

We denote the data transmitted in each instance of auxiliary communication (regardless of whether it is forward or backward) by Z_k, where $k \in \{1, \ldots, m\}$, in temporal order.

If the total auxiliary communication $Z := Z^{(m)} := (Z_1, \ldots, Z_m)$ has $|\mathcal{A}_Z|$ possible values (we require this number to be finite for any given protocol), then this alone would allow the parties to establish $\log |\mathcal{A}_Z|$ bits of perfect common randomness without using $\mathcal{E}$ at all! We therefore subtract $\log |\mathcal{A}_Z|$ from the final amount of common randomness established and hence define the **net rate** of the protocol is

$$\frac{1}{n}(H(K) - \log |\mathcal{A}_Z|).$$

A **forward-assisted** randomness distribution protocol is one in which all steps are of type (i) or (ii). A **back-assisted** randomness distribution protocol is one in which all steps are of type (i) or (iii). An **unassisted** randomness distribution protocol is one in which all steps are of type (i).

▶ **Definition 1.** We say a net rate R is achieved by two-way protocols for channel $\mathcal{E}$ if for all $\epsilon > 0$ and all sufficiently large n, there is an ϵ-good protocol for n noisy channel uses with net rate no less than R. We define $R_{\leftrightarrow}(\mathcal{E})$ to be the supremum of net rates achieved by two-way protocols; $R_{\rightarrow}(\mathcal{E})$ to be the supremum of net rates achieved by forward-assisted protocols; $R_{\leftarrow}(\mathcal{E})$ to be the supremum of net rates achieved by back-assisted protocols; and $R(\mathcal{E})$ to be the supremum of net rates achieved by unassisted protocols;

It follows immediately from the definitions that

$$R(\mathcal{E}) \leq R_{\rightarrow}(\mathcal{E}) \leq R_{\leftrightarrow}(\mathcal{E}) \text{ and } R(\mathcal{E}) \leq R_{\leftarrow}(\mathcal{E}) \leq R_{\leftrightarrow}(\mathcal{E}). \tag{4}$$

3 Classical equality between information and randomness distribution

In this section we will show that $R_{\leftrightarrow}(\mathcal{E})$ can be no larger than the entanglement-assisted capacity of $\mathcal{E}$, $C_E(\mathcal{E})$. Since $C_E(\mathcal{E}) = C(\mathcal{E})$ for classical channels (and, more generally, for classical-quantum channels), this establishes that

$$C(\mathcal{E}) = R_{\leftarrow}(\mathcal{E}) = R_{\rightarrow}(\mathcal{E}) = R_{\leftrightarrow}(\mathcal{E})$$

for such channels. We note that common randomness distribution via a classical channel $\mathcal{E}$ and noiseless feedback was considered by Ahslwede and Csiszar in [2], and that the equality $C(\mathcal{E}) = R_{\leftarrow}(\mathcal{E})$ is a corollary of their Theorem 4.3.

It was shown by Bennett, Shor, Smolin and Thapliyal [12], that the entanglement-assisted classical capacity of a channel $\mathcal{E}_{\mathsf{Y}\leftarrow\mathsf{X}}$ is given by

$$C_E(\mathcal{E}) = \max_{\rho_{\mathsf{RX}}} I(\mathsf{R}:\mathsf{Y})_{\mathcal{E}_{\mathsf{Y}\leftarrow\mathsf{X}}\rho_{\mathsf{RX}}}. \tag{5}$$

We will show that the same formula is an upper bound on $R_{\leftrightarrow}(\mathcal{E})$.

▶ **Theorem 2.** *For any channel $\mathcal{E}$, $R_{\leftrightarrow}(\mathcal{E}) \leq C_E(\mathcal{E})$.*

Proof. Let us consider a protocol which makes n uses of the channel $\mathcal{E}$ and m auxiliary communication steps. For $k \in \{1, \ldots, n\}$, let X_k denote the input system, and Y_k the output system, for the k-th use of the noisy channel.

Initially, Alice and Bob have systems A_0 and B_0 which are uncorrelated in that $I(\mathsf{A}_0 : \mathsf{B}_0) = 0$. We may assume without loss of generality that any local randomness used in the protocol is already present in the state of these systems. We may assume without loss of generality that at each step Alice and Bob have retained a full record of all auxiliary communication up to that step.

Suppose that at step j of the protocol, Bob sends Alice Z_k by auxiliary back communication. Then we may bound

$$\begin{aligned} I(\mathsf{A}_j : \mathsf{B}_j) &\overset{(a)}{\leq} I(\mathsf{A}_{j-1}Z_k : \mathsf{B}_j) \overset{(b)}{\leq} I(\mathsf{A}_{j-1}Z_k : \mathsf{B}_{j-1}) \\ &= I(\mathsf{A}_{j-1} : \mathsf{B}_{j-1}) + H(Z_k|\mathsf{A}_{j-1}) - H(Z_k|\mathsf{A}_{j-1}\mathsf{B}_{j-1}) \\ &\overset{(c)}{\leq} I(\mathsf{A}_{j-1} : \mathsf{B}_{j-1}) + H(Z_k|\mathsf{A}_{j-1}) \overset{(d)}{\leq} I(\mathsf{A}_{j-1} : \mathsf{B}_{j-1}) + H(Z_k|Z^{(k-1)}) \end{aligned} \tag{6}$$

where (a) and (b) are data processing, (c) is because, since Z_k is classical, $H(Z_k|\mathsf{A}_{j-1}\mathsf{B}_{j-1}) \geq 0$ and (d) is because A_{j-1} includes $Z^{(k-1)}$. A similar argument establishes the same inequality when Alice sends Bob Z_k by auxiliary forward communication, instead.

Now consider the case where Alice makes an input X_k to the noisy channel $\mathcal{E}$ at step j, with Bob receiving output Y_k. Then

$$\begin{aligned} I(\mathsf{A}_j : \mathsf{B}_j) &\overset{(a)}{\leq} I(\mathsf{A}_j : \mathsf{B}_{j-1}\mathsf{Y}_k) \\ &= I(\mathsf{A}_j : \mathsf{Y}_k) + I(\mathsf{A}_j : \mathsf{B}_{j-1}|\mathsf{Y}_k) \\ &= I(\mathsf{A}_j : \mathsf{Y}_k) + I(\mathsf{A}_j\mathsf{Y}_k : \mathsf{B}_{j-1}) - I(\mathsf{Y}_k : \mathsf{B}_{j-1}) \\ &\overset{(b)}{\leq} I(\mathsf{A}_j : \mathsf{Y}_k) + I(\mathsf{A}_j\mathsf{Y}_k : \mathsf{B}_{j-1}) \\ &\overset{(c)}{\leq} I(\mathsf{A}_j : \mathsf{Y}_k) + I(\mathsf{A}_{j-1} : \mathsf{B}_{j-1}) \\ &\overset{(d)}{\leq} C_E(\mathcal{E}) + I(\mathsf{A}_{j-1} : \mathsf{B}_{j-1}). \end{aligned} \tag{7}$$

Here, (a) and (c) are by data processing, (b) is positivity of mutual information, and (d) is by the result of Bennett, Shor, Smolin and Thapliyal.

Recall that $Z := Z^{(m)}$ is the total record of auxiliary communication. Starting with $I(\mathsf{A}_{n+m} : \mathsf{B}_{n+m})$, and repeatedly invoking the inequality (6) or (7) depending on the type of step, we obtain

$$\begin{aligned} I(\mathsf{A}_{n+m} : \mathsf{B}_{n+m}) &\leq I(\mathsf{B}_0 : \mathsf{A}_0) + nC_E(\mathcal{E}) + \sum_{k=1}^{m} H(Z_k|Z^{(k-1)}) \\ &= nC_E(\mathcal{E}) + H(Z) \\ &\leq nC_E(\mathcal{E}) + \log|\mathcal{A}_Z|, \end{aligned} \tag{8}$$

where the equality is by the chain rule and $I(\mathsf{B}_0 : \mathsf{A}_0) = 0$. Finally, we bound the net rate R of the protocol by

$$\begin{aligned} R &= \frac{1}{n}(H(K) - \log|\mathcal{A}_Z|) = \frac{1}{n}(I(K:J) + H(K|J) - \log|\mathcal{A}_Z|) \\ &\overset{(a)}{\leq} \frac{1}{n}(I(\mathsf{A}_{n+m} : \mathsf{B}_{n+m}) + H(K|J) - \log|\mathcal{A}_Z|) \\ &\overset{(b)}{\leq} \frac{1}{n}(nC_E(\mathcal{E}) + \log|\mathcal{A}_Z| + nc\epsilon + 1 - \log|\mathcal{A}_Z|) \\ &\overset{(c)}{=} C_E(\mathcal{E}) + c\epsilon + 1/n \end{aligned}$$

where (a) is data processing, (b) is by inequalities (8) and (3), and (c) by is Shannon's noisy channel coding theorem. Recalling the definition of $R_{\leftrightarrow}$, we have established that

$$R_{\leftrightarrow}(\mathcal{E}) \leq C_E(\mathcal{E}). \tag{9}$$

◀

We also claimed that $C(\mathcal{E}) = C_{\leftarrow}(\mathcal{E}) = C_{\rightarrow}(\mathcal{E}) = C_{\leftrightarrow}(\mathcal{E})$ for classical-quantum channels. In fact, we can show that this is true of any entanglement-breaking channel. The general equality $C(\mathcal{E}) = C_{\leftarrow}(\mathcal{E})$ is established in the next section. Now, note that we can write

$$C_{\leftrightarrow}(\mathcal{E}) = \sup_m \{C_{\leftarrow}(\mathcal{E} \otimes \mathcal{A}_m) - \log m\}$$

where $\mathcal{A}_m$ is a classical identity channel with m input symbols. Since $\mathcal{E}$ and $\mathcal{A}_m$ are both entanglement-breaking, we have

$$C_{\leftarrow}(\mathcal{E} \otimes \mathcal{A}_m) = C(\mathcal{E} \otimes \mathcal{A}_m) = C(\mathcal{E}) + C(\mathcal{A}_m) = C(\mathcal{E}) + \log m$$

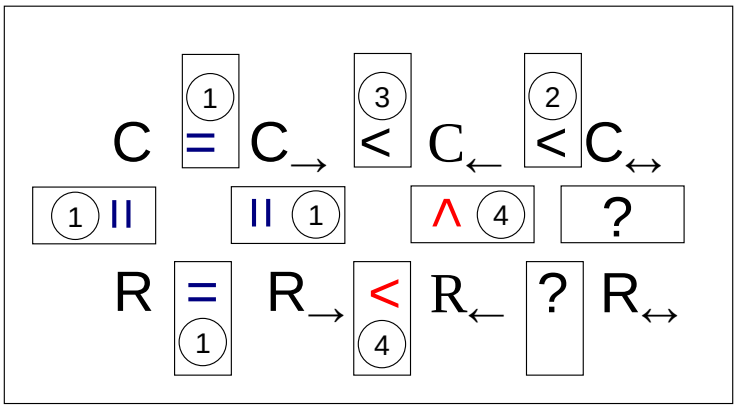

Figure 2 Relations between the communication (C) and randomness distribution (R) capacities. Note that an equality means that both capacities are equal for all channels; On the other hand, an inequality means that we know of at least one channel where one is strictly higher, which does not preclude the possibility that for other channels they may be equal. (1) It is easy to prove $C = R = C_{\rightarrow} = R_{\rightarrow}$, see Section 4.1 below. (2) Corollary of [11], using echo-correctable channels. (3) Corollary of [9], using random-phase coupling channels. (4) Our result in subsection 4.2. The relations between $R_{\leftarrow}(\mathcal{E})$, $R_{\leftrightarrow}(\mathcal{E})$ and $C_{\leftrightarrow}(\mathcal{E})$, i.e., whether they are equal for all channels or there are some examples of strict separation between them, remains an open question.

by Bowen-Nagarajan [3], the HSW theorem [7, 8], and the fact that the Holevo information is additive for entanglement breaking channels [10]. Therefore,

$$C_{\leftarrow}(\mathcal{E}) = C_{\leftrightarrow}(\mathcal{E}) = C(\mathcal{E})$$

for entanglement-breaking $\mathcal{E}$.

4 Quantum scenario

As opposed to the classical scenario, where all capacities of randomness distribution and information transmission, collapse into a single quantity given by Shannon's capacity, quantum channels have a richer behaviour depicted in Figure 2. The only similarity between the quantum and classical scenario is restricted to the unassisted and forward assisted capacities where, as shown in subsection 4.1 below, one can prove the equality

$$C(\mathcal{E}) = R(\mathcal{E}) = C_{\rightarrow}(\mathcal{E}) = R_{\rightarrow}(\mathcal{E}). \tag{10}$$

The situation changes radically for feedback and two-way assisted capacities. It was shown in [11] that by concatenating an *echo-correctable channel* and a depolarizing channel one can obtain an entanglement-breaking channel exhibiting a strict separation $C_{\leftarrow}(\mathcal{E}) < C_{\leftrightarrow}(\mathcal{E})$. Subsequently, in [9], the possibility of a strict separation $C_{\rightarrow}(\mathcal{E}) < C_{\leftarrow}(\mathcal{E})$ was shown using random-phase coupling channels (also informally called *rocket* channels). In subsection 4.2 below we show that there can be a separation $C_{\leftarrow}(\mathcal{E}) < R_{\leftarrow}(\mathcal{E})$. As a corollary we also obtain the separation $R_{\rightarrow}(\mathcal{E}) < R_{\leftarrow}(\mathcal{E})$.

4.1 Equality between unassisted and forward-assisted capacities

It is straightforward to see that $C(\mathcal{E}) \leq C_{\rightarrow}(\mathcal{E})$ and $R(\mathcal{E}) \leq R_{\rightarrow}(\mathcal{E})$ (assistance can only increase the rate), $C(\mathcal{E}) \leq R(\mathcal{E})$ (if you can send a bit of information you can also distribute

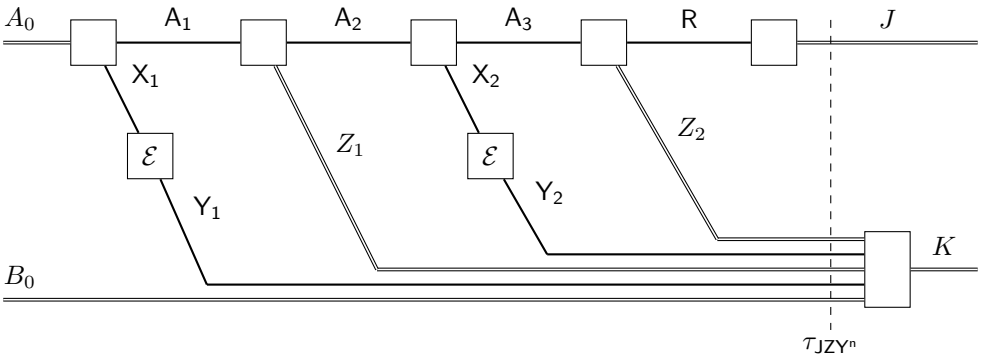

Figure 3 An example of a forward assisted randomness distillation protocol which makes two uses of the channel $\mathcal{E}$. Without loss of generality, Bob waits until receiving all communication from Alice to perform his local processing, and obtain K.

a bit of shared randomness) and similarly $C_{\rightarrow}(\mathcal{E}) \leq R_{\rightarrow}(\mathcal{E})$, because in both cases we only subtract the assisting forward communicaiton. In order to prove the equality between unassisted and forward-assisted capacities in eq. (10), it is sufficient to prove that the highest of the four capacities, $R_{\rightarrow}(\mathcal{E})$, is upper-bounded by the lowest of them, i.e. that $R_{\rightarrow}(\mathcal{E}) \leq C(\mathcal{E})$.

Since Bob does not send anything back to Alice during a forward-assisted protocol, there is no loss of generality if Alice makes all n uses of the noisy channel, sends all auxiliary classical communication and computes her share of the common randomness, J, before Bob does anything, as illustrated in Figure 3. We denote by R all systems retained by Alice, from which she computes her share of the common randomness.

Let X^n be the n input systems, and Y^n the n output systems, for the n uses of the noisy channel $\mathcal{E}^{\otimes n}$. We introduce a register Z which stores the value of the auxiliary forward communication Z, which can take one of $|\mathcal{A}_Z|$ values. After Alice has made all her communication to Bob, the state of the $\mathsf{ZY}^n\mathsf{R}$ system is

$$\sigma_{\mathsf{ZY}^n\mathsf{R}} = \sum_z p(z)|z\rangle\langle z|_{\mathsf{Z}} \otimes \mathcal{E}^{\otimes n}_{\mathsf{Y}^n \leftarrow \mathsf{X}^n} \rho^{(z)}_{\mathsf{X}^n\mathsf{R}} \tag{11}$$

where $\rho^{(z)}_{\mathsf{X}^n\mathsf{R}}$ is the state of the $\mathsf{X}^n\mathsf{R}$, conditioned on $Z = z$. Now, Alice performs a measurement $E(j)$ (POVM of outcome j) on the system R to obtain her share J of the common randomness, which is stored in register J. At this point the state of the system is

$$\tau_{\mathsf{JZY}^n} = \sum_z q(j|z)p(z)|j\rangle\langle j|_{\mathsf{J}} \otimes |z\rangle\langle z|_{\mathsf{Z}} \otimes \mathcal{E}^{\otimes n}_{\mathsf{Y}^n \leftarrow \mathsf{X}^n} \rho^{(z,j)}_{\mathsf{X}^n}, \tag{12}$$

where

$$q(j|z)\rho^{(z,j)}_{\mathsf{X}^n} := \mathrm{Tr}_{\mathsf{R}} E(j)_{\mathsf{R}} \rho^{(z)}_{\mathsf{X}^n\mathsf{R}}$$

defines the states $\rho^{(z,j)}_{\mathsf{X}^n}$ and conditional distribution $q(j|z)$.

Then Bob performs a measurement on the ZY^n system to obtain his share of randomness K. We can bound the mutual information between the shares by

$$\begin{aligned} I(J:K) &\overset{(a)}{\leq} I(\mathsf{J}:\mathsf{ZY}^n)_\tau = I(\mathsf{J}:\mathsf{Y}^n)_\tau + I(\mathsf{J}:\mathsf{Z}|\mathsf{Y}^n)_\tau \\ &= I(\mathsf{J}:\mathsf{Y}^n)_\tau + H(\mathsf{Z})_\tau - I(\mathsf{Z}:\mathsf{Y}^n)_\tau - H(\mathsf{Z}|\mathsf{J},\mathsf{Y}^n)_\tau \\ &\overset{(b)}{\leq} I(\mathsf{J}:\mathsf{Y}^n)_\tau + H(\mathsf{Z})_\tau \overset{(c)}{\leq} \chi(\mathcal{E}^{\otimes n}) + \log|\mathcal{A}_Z| \end{aligned} \tag{13}$$

where (a) is data processing, (b) is because τ is separable with respect to the Z/JY^n bipartition so $H(\mathsf{Z}|\mathsf{JY}^n) \geq 0$, and by positivity of mutual information, and (c) is because $I(\mathsf{J}:\mathsf{Y}^n) \leq \chi(\mathcal{E}^{\otimes n})$. We use this to bound the net rate R of the protocol thus

$$\begin{aligned} R =& \frac{1}{n}(H(K) - \log|\mathcal{A}_Z|) = \frac{1}{n}(I(K:J) + H(K|J) - \log|\mathcal{A}_Z|) \\ \leq& \frac{1}{n}(\chi(\mathcal{E}^{\otimes n}) + \log|\mathcal{A}_Z| + H(K|J) - \log|\mathcal{A}_Z|) \leq \frac{1}{n}\chi(\mathcal{E}^{\otimes n}) + c\epsilon + 1/n, \end{aligned}$$

and therefore $R_{\rightarrow}(\mathcal{E}) \leq \lim_{n\to\infty} \frac{1}{n}\chi(\mathcal{E}^{\otimes n}) = C(\mathcal{E})$, where the equality is the Holevo-Schumacher-Westmoreland theorem [7, 8].

4.2 Quantum-classical channels; separation $C_{\leftarrow}(\mathcal{E}) < R_{\leftarrow}(\mathcal{E})$

Suppose that $\mathcal{E}_{Y\leftarrow\mathsf{X}}$ is a **quantum-classical** channel. That is, a channel of the form

$$\mathcal{E}_{Y\leftarrow\mathsf{X}} : \rho_{\mathsf{X}} \mapsto \sum_{y\in\mathcal{A}_Y} |y\rangle\langle y|_Y \mathrm{tr} E(y)_{\mathsf{X}}\rho_{\mathsf{X}} \tag{14}$$

where $\{E(y)_{\mathsf{X}} : y \in \mathcal{A}_Y\}$ is a POVM on X. In this case we can show that there is a back-assisted randomness distribution which achieves the upper-bound $C_E(\mathcal{E})$ for two-way assisted protocols, and therefore:

▶ **Theorem 3.** *For quantum-classical channels $\mathcal{E}_{Y\leftarrow\mathsf{X}}$, $R_{\leftarrow}(\mathcal{E}) = R_{\leftrightarrow}(\mathcal{E}) = C_E(\mathcal{E})$.*

We just need to show achievability: One way that n uses of a quantum-classical channel can be used to produce randomness with auxiliary back communication is as follows. Alice locally prepares n copies of a state ψ_{RX} and applies the n uses of the channel to X^n. This results in n copies of a quantum-classical state

$$\sum_y p(y)\rho(y)_{\mathsf{R}} \otimes |y\rangle\langle y|_Y = \mathcal{E}_{Y\leftarrow\mathsf{X}}\psi_{\mathsf{RX}} \tag{15}$$

being shared between Alice and Bob, with Bob holding the classical register Y, and $p(y) := \mathrm{tr}_{\mathsf{RX}} E(y)_{\mathsf{X}}\psi_{\mathsf{RX}}$ and $\rho(y)_{\mathsf{R}} := \mathrm{tr}_{\mathsf{X}} E(y)_{\mathsf{X}}\psi_{\mathsf{RX}}/p(y)$. Now, in the proof of the classical-quantum Slepian-Wolf theorem of Devetak and Winter [5] it was shown that, for any $0 < \epsilon < 1/2$ and $\delta > 0$, and all sufficiently large n, we can find $|\mathcal{A}_Z|$ disjoint subsets $\{C_z : z \in \mathcal{A}_Z\}$ of $\mathcal{A}_Y^n$ such that

- **(i)** the probability that Y^n fails to belong to one of the subsets is not more than 2ϵ,
- **(ii)** given the knowledge $Y^n \in C_z$, Alice can perform a measurement on R^n which identifies Y^n with probability of error no more than ϵ,
- **(iii)** $\frac{1}{n}\log|\mathcal{A}_Z| \leq H(Y|\mathsf{R}) + 2\delta$.

This suggests the following protocol: Bob takes $K = Y^n$ as his share of the common randomness (so $H(K) = nH(Y)$) and sends Alice the identity Z of a subset C_Z containing Y^n (if such exists) whereupon Alice measures R^n to obtain an estimate J of Y^n. This protocol has $\Pr(K \neq J) \leq 3\epsilon$ and net rate

$$\frac{1}{n}(H(K) - \log|\mathcal{A}_Z|) \geq H(Y) - H(Y|\mathsf{R}) - 2\delta = I(Y:\mathsf{R}) - 2\delta.$$

Therefore, by optimising over the choice of ψ_{XR} in the protocol, we have established that

$$R_{\leftarrow}(\mathcal{E}) \geq \max_{\psi_{\mathsf{XR}}} I(Y:\mathsf{R})_{\mathcal{E}_{Y\leftarrow\mathsf{X}}\psi_{\mathsf{XR}}} = C_E(\mathcal{E}), \tag{16}$$

where $C_E(\mathcal{E})$ is the entanglement-assisted capacity of $\mathcal{E}$, and the equality is the theorem of Bennett, Shor, Smolin and Thapliyal [12].

Now, quantum-classical channels are entanglement breaking. It was shown by Bowen and Nagarajan [3] that classical feedback cannot increase the classical capacity of entanglement breaking channels, so we have $C_{\leftarrow}(\mathcal{E}) = C(\mathcal{E})$. Meanwhile, in [4], Holevo has given examples of quantum-classical channels with $C_E(\mathcal{E}) > C(\mathcal{E})$. By Theorem 3 and Bowen-Nagarajan, these channels also exhibit a separation $R_{\leftarrow}(\mathcal{E}) > C_{\leftarrow}(\mathcal{E})$. To be more specific, consider the case where the POVM elements determining $\mathcal{E}$ are rank-one projectors onto pair-wise linearly independent subspaces. Then $C(\mathcal{E}) \leq C_E(\mathcal{E}) = \log d$, and Holevo shows that the inequality is strict *unless* the the POVM is a orthonormal basis measurement [4].

4.3 Specific example

Given two rank-1 projective measurements $E^{(0)}$ and $E^{(1)}$ on a d-dimensional system X with outcomes in $\{1, \ldots, d\}$ we may construct a quantum-classical channel $\mathcal{F}_{Y \leftarrow \mathsf{X}}$ whose input system is X and whose output is a pair $Y = (M, G)$ where M is a bit chosen uniformly at random, and G is the result of performing the measurement $E^{(M)}$ on X. So, M tells us which basis was measured and G tells us the result of that measurement. Without loss of generality we can take $E^{(0)}$ to be the computational basis measurement.

Since the POVM corresponding to this classical-quantum channel has rank-one elements we already know that

$$R_{\leftarrow}(\mathcal{F}) = C_E(\mathcal{F}) = \log d. \tag{17}$$

In Figure 4 we illustrate a protocol which distributes $1 + \log d$ bits of perfectly correlated randomness with one use of $\mathcal{F}$ and a single bit of communication from Bob to Alice, thus attaining a net rate of $\log d$ bits per channel use, perfectly.

On the other hand, if we choose $E^{(1)}$ so that the two measurement bases are mutual unbiased, it is not hard to establish that $C_{\leftarrow}(\mathcal{F}) = C(\mathcal{F}) = \chi(\mathcal{F}) \leq \frac{1}{2} \log d$: The first two equalities are because the channel is entanglement breaking. It remains to upper bound the Holevo information $\chi(\mathcal{F})$. Suppose that the input to the channel is drawn from an ensemble $\{(p(w), \psi^{(w)}) : w = 1, \ldots k\}$ with ensemble average $\rho = \sum_{w=1}^{k} p(w)\psi^{(w)}$. Maximising

$$H(M,G)_\rho - \sum_w p(w) H(M,G)_{\psi^{(w)}} \tag{18}$$

over all ensembles, we obtain the Holevo information $\chi(\mathcal{F})$, and since the channel is entanglement breaking, we know that $C_{\leftarrow}(\mathcal{F}) = C(\mathcal{F}) = \chi(\mathcal{F})$. Clearly

$$H(M,G)_\rho \leq 1 + \log d \tag{19}$$

while, for any state ψ,

$$\begin{aligned} H(M,G)_\psi =& H(M) + H(G|M=0)_\psi \Pr(M=0) + H(G|M=1)_\psi \Pr(M=1) \\ =& 1 + \frac{1}{2}\left(H(G|M=0)_\psi + H(G|M=1)_\psi\right). \end{aligned}$$

If the measurements correspond to mutually unbiased bases then, according to Maassen and Uffink's entropic uncertainty relation [14], we have

$$H(M,G)_\psi \geq 1 + \frac{1}{2}\log d, \tag{20}$$

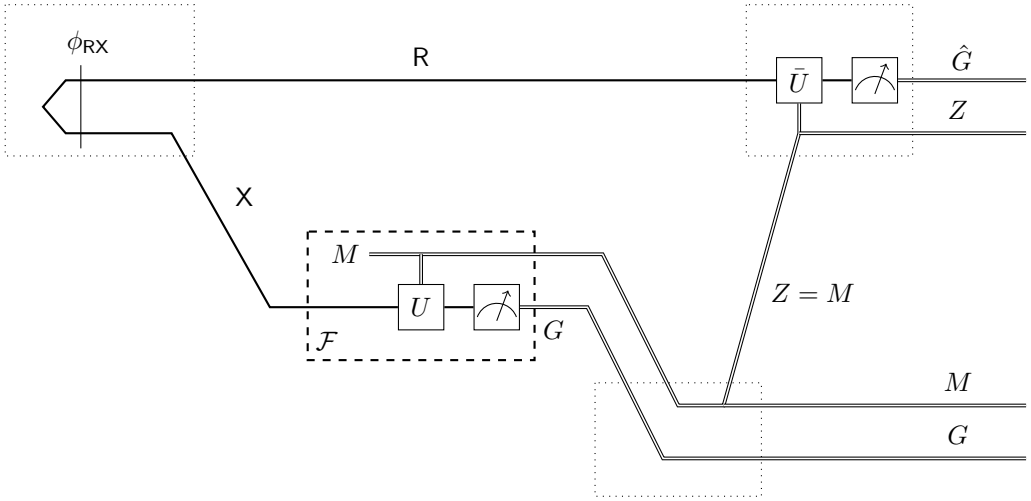

Figure 4 Sharing $1 + \log d$ bits of perfect randomness with one use of the channel $\mathcal{F}$ (the contents of the dashed rectangle) and one bit of back communication: Alice locally prepares a maximally entangled state ϕ_{RX} and inputs X to the channel. We can view the channel as performing a unitary controlled by the bit M and then performing a computational basis measurement to yield G. Alice sets $Z = M$ and sends Z to Bob, who performs $\bar{U}$ (the complex conjugate of U) iff $Z = 1$ and then performs a computational basis measurement on R to yield a value $\hat{G}$. By the $U \otimes \bar{U}$ invariance of ϕ, $\hat{G} = G$ with probability one, so if Alice sets $J = (\hat{G}, Z)$ and Bob sets $K = (G, M)$ then $\Pr(K = J) = 1$, and K is uniformly distributed. Local operations are surrounded by dotted lines.

and substituting the bounds (19) and (20) into (18),

$$C_{\leftarrow}(\mathcal{E}) = C(\mathcal{E}) = \chi(\mathcal{E}) \leq \frac{1}{2} \log d.$$

This upper bound is indeed tight for both, $C_{\leftarrow}(\mathcal{E})$ and $C(\mathcal{E})$, as the channel $\mathcal{E}$ can be transformed with some post-processing on Bob's side into an erasure channel (if $M = 1$ erase register G) of error probability $1/2$. Therefore (feedback-assisted) error-correcting codes for the erasure channel can be used to saturate the bound $C_(\mathcal{E}) = C_{\leftarrow}(\mathcal{E}) = 1/2 \log d$.

5 Conclusion

Despite being, a priori, different things, we have seen that the capacity for a classical channel to distribute shared randomness and to send information are the same, with arbitrary classical assistance. For quantum channels, we have shown that the entanglement-assisted capacity $C_E(\mathcal{E})$ is a general upper bound for $R_{\leftrightarrow}(\mathcal{E})$, and shown that this bound can be achieved using only back-communication for quantum-classical channels. Using this result we have established that strict separations $C_{\leftarrow}(\mathcal{E}) < R_{\leftarrow}(\mathcal{E})$ are possible for quantum-classical channels. We give an explicit example for which $R_{\leftarrow}(\mathcal{E}) = \log d$ while $C_{\leftarrow}(\mathcal{E}) = \frac{1}{2} \log d$.

Our result shows that contrary to what is predicted by classical information theory, where the optimal way of distributing randomness is to generate it locally and distribute it through the channel, quantum mechanics allows for the activation of randomness initially locked inside the channel, which boost the amount of shared randomness generated in the process.

References

1 T. M. Cover and J. A. Thomas, *Elements of Information Theory*, (Wiley, New York, 1991).
2 R. Ahlswede and I. Csiszár, IEEE Trans. Info. Theory **39,** pp. 1121–1132 (1993).
3 G. Bowen and R. Nagarajan, IEEE Trans. Inform. Theory 51, 320 (2005).
4 A. Holevo, Problems of Information Transmission, vol. 48 (2012), pp. 1–10.
5 I. Devetak and A. Winter, Phys. Rev. A 68, 042301 (2003)
6 I. Devetak and A. Winter, IEEE Trans. Info. Theory **50,** 3183 (2004).
7 A. Holevo, IEEE Transactions on Information Theory 44(1):269–273 (1996).
8 B. Schumacher, M. Westmoreland, Phys. Rev. A 56:131–138 (1997).
9 G. Smith and J. A. Smolin, Phys. Rev. Lett. 103, 120503 (2009).
10 Additivity of the classical capacity of entanglement-breaking quantum channels. Journal of Mathematical Physics, 43(9):4334–4340, (2002).
11 C. H. Bennett, I. Devetak, P. W. Shor, and J. A. Smolin, Phys. Rev. Lett. 96, 150502 (2006).
12 C. H. Bennett, P. W. Shor, J. A. Smolin, A. V. Thapliyal, IEEE Trans. Info. Theory **48,** 2637 (2002).
13 L. P. Hughston, R. Jozsa, and W. K. Wootters, Phys. Lett. A **183,** 14 (1993).
14 H. Maassen, and J. Uffink, Phys. Rev. Lett. 60, 1103 (1988)