Parrondo's paradox and superactivation of classical and quantum capacity of communication channels with memory

Sergii Strelchuk[*]

Department of Applied Mathematics and Theoretical Physics, University of Cambridge, Cambridge CB3 0WA, United Kingdom
(Received 9 May 2013; published 11 September 2013)

There exist memoryless zero-capacity quantum channels that when used jointly result in a channel with positive capacity. This phenomenon is called superactivation. Making use of Parrondo's paradox, we exhibit examples of the superactivation-like effect for the capacity of classical communication channels as well as quantum and private capacity of quantum channels with memory. There are several ingredients necessary for superactivation of quantum capacity to occur in the memoryless case. The first one is the requirement for the quantum channels which are suitable for superactivation to come from two distinct families: binding entanglement channels and erasure channels. The second one is the ability to utilize inputs which are entangled across the uses of the channels. Our construction uses a single family of erasure channels with classical memory to achieve the same superactivation-like effect for quantum capacity without any of the ingredients above.

I. INTRODUCTION

When can we achieve reliable communication over an imperfect channel? Classical information theory, developed by Shannon, gives a satisfactory answer to this question when the information that we want to transmit and the underlying channel are both classical. It characterizes the performance of every classical channel with a single figure of merit: classical capacity. However, if we turn to sending quantum information, it is inadequate because the capacity of quantum channels also depends on what other resources are available. This gives rise to the numerous types of capacities of the quantum channels. In particular, quantum capacity of the quantum channel characterizes its ability to transmit entanglement reliably. Private capacity indicates how suitable the channel is for the task of secret key sharing. The contrast between classical and quantum channels is especially stark when we turn to zero-capacity channels. In classical information theory they are precisely the set of useless channels: the receiver cannot decode the message reliably from the sender because of the adverse effects of noise.

In 2008 Smith and Yard [1] showed that this is not the case for quantum memoryless channels. They provided an intriguing example of two zero-capacity memoryless quantum channels, which, when combined, result in reliable transmission of quantum information reversing the deleterious effects of noise. This phenomenon is called *superactivation*. It is of great interest because it may provide clues about novel methods of transmitting the frail quantum data in the presence of the strongly deleterious noise in the channel.

One important aspect of the channels whose capacity is currently known to be superactivated is the requirement that they must come from two distinct classes: the first one comes from the set of binding entanglement channels [2], and the second one comes either from the set of zero-capacity erasure channels with the rational erasure probability or from depolarizing channels in the regime where they become

antidegradable [3]. The capacity of the channel from each of those classes cannot be superactivated by another channel within the same class [4]. Apart from these two families of channels, it is not known if there exist qualitatively different zero-capacity channels that give rise to the superactivation of the quantum capacity. This is in part because proving that a given channel has zero capacity is hard.

The simplest example of the channel from the latter class is the 50% erasure channel: half of the time it forwards the input state to the receiver, and half of the time it erases the input state, sending Bob the erasure flag. The quantum and private capacity of erasure channels is well studied in the regime without feedback. However, despite its simplicity, the quantum capacity under backward classical communication assistance is unknown [5,6].

Another crucial ingredient for superactivation which was implicitly present in all of the protocols is the use of inputs which are entangled across multiple uses of the channels [1]. This manifestly quantum feature was known to increase the quantum and classical capacity of quantum channels. It is also the reason why there is no computationally tractable way of finding the capacity of a general quantum channel. The extent to which inputs entangled across the uses of the channel may help is unknown.

Unlike the situation with the quantum capacity, showing that the private capacity can be superactivated turned out to be a much harder question that remained without an answer for channels both with and without memory, despite many efforts [7–9]. Private capacity was recently found to be nonadditive: the sum of the private capacities of individual channels turned out to be smaller than the private capacity of the joint channel composed of them [7,8]. Superactivation is the strongest form of nonadditivity: it occurs when all the individual channel capacities are zero, but the joint capacity is strictly positive.

Curiously, for all previously known constructions which exhibit nonadditivity of the private capacity it is also necessary to make use of the inputs, which are entangled across the use of the channel.

For many tasks of information processing considering merely memoryless channels, quantum or classical, is not

----
[*]ss870@cam.ac.uk

sufficient. For most practical purposes one would like to know how well the channel performs given finitely many uses. In this regime the assumption that noise is uncorrelated between channel uses is no longer justified: the environment may have memory. This led to explosive research activity in classical and quantum information theory that quantifies the capacity of channels with memory.

In this paper, we provide a general construction for super-activation of the channels with memory. This construction is in the spirit of Parrondo's paradox [10], originally discovered in the context of Brownian ratchets. It occurs when a player has access to two specific coin-flipping betting games, each of which is a loss for him. However, contrary to the intuitive expectation, playing them in an alternating fashion results in a winning game.

We show that the presence of classical memory leads to superactivation of the capacity of classical and quantum channels. This indicates that the performance of both classical and quantum channels with memory can no longer be characterized with a single figure of merit, its capacity. For quantum channels, we demonstrate superactivation without the above requirements which were necessary for the memoryless channels. We introduce two simple zero-capacity quantum channels which make use only of erasure channels: the first one being a memoryless erasure channel and the second one being the combination of erasure channels with classical memory. Their convex combination results in another channel with shared classical memory (i.e., in which both channels in the mixture have access to a shared memory register) which has positive quantum capacity.

Moreover, this effect also holds for the private capacity of the channels with memory. This is especially interesting because due to its simple construction it may provide clues to constructing the long-sought example of superactivation of the private capacity in the memoryless setting, which despite many efforts still remains elusive. The main obstacle for this lies in the difficulty of proving that a given channel has no private capacity.

For classical channels, we slightly modify the construction used to demonstrate superactivation of the quantum and private capacity of the quantum channel.

Quantum channels with classical memory can be much more powerful when it comes to information transmission. There have been demonstrations of how memory may benefit information transmission [11]. In practice, as our construction will show, supplementing quantum channels with classical memory offers superior error-correcting capabilities at a very small price. However, in theory, such channels are hard to analyze: estimating their properties and computing the capacity turns out to be a much harder task than for their memoryless counterparts. In the quantum case this is especially difficult because for quantum channels with memory there exists no notion of state-channel duality. Consequently, one cannot introduce a Choi-Jamiołkowski state to determine the properties of memoryless channels in the same way as in [2], so new techniques are required.

This paper is structured as follows: in Sec. II we introduce the general construction, which gives rise to the Parrondo effect in the context of communication channels with memory. In Sec. III we find channels which demonstrate the superactivation effect for the quantum and private capacity of quantum channels with memory. Last, in Sec. IV, we provide the evidence for the superactivation-like effect for classical channels with memory, followed by a discussion in Sec. V.

II. PARRONDO'S PARADOX AND GENERAL CONSTRUCTION

It has long been known that noise negatively impacts the performance of classical and quantum systems. But it does not always have a deleterious effect [12]. Sometimes, it is possible to benefit from having noise in the setup [13,14]. There are several features of the system which make it amenable to the noise-enhanced performance: the system must have a certain degree of nonlinearity and asymmetry. The description of noise-enhanced effects in classical or quantum systems is generally very unintuitive and complex, with only a few examples in the quantum case [15]. In 1996, Parrondo devised his eponymous paradox: an example which in its unrivalled simplicity demonstrates how noise can increase the odds of winning in a series of the coin-flipping games [10,16]. This paradox has been explored in connection with a variety of areas and systems [17–19]. To introduce the paradox, we closely follow Ref. [20] and describe it in terms of Parrondo's games. The setup consists of two coin-flipping games. We say that the player gets one pound (i.e., a winning event occurs) if the coin lands heads up (H), and the player loses one pound when the coin lands tails (T). Game A involves a simple coin with the following distribution:

$$\Pr(H) = p_1 - \epsilon, \quad \Pr(T) = 1 - p_1 + \epsilon, \tag{1}$$

where $\epsilon > 0$ and $p_i = 0.5$. The second game, game B, makes use of two biased coins each with $\Pr(H) = p_i$, $i = 2,3$. Game B keeps track of the current amount of money the player has or owns M and is defined as follows:

$$\begin{aligned} \Pr(H | M \bmod n = 0) &= p_2 - \epsilon, \\ \Pr(T | M \bmod n = 0) &= 1 - p_2 + \epsilon, \\ \Pr(H | M \bmod n \neq 0) &= p_3 - \epsilon, \\ \Pr(T | M \bmod n \neq 0) &= 1 - p_3 + \epsilon, \end{aligned} \tag{2}$$

with probabilities $p_2 = 0.1$ and $p_3 = 0.75$ and $n = 3$.

The analysis of both of the games can be performed using elementary probability and properties of Markov chains. Game A is clearly a losing game, as it is biased against the player. Game B can be conveniently represented as a Markov chain having three states, s_0, s_1, and s_2, each corresponding to the possible values of the remainder $M \bmod 3$ with the respective transition probabilities. This game also turns out to be losing in the stationary distribution of the chain.

Now, consider a new game, game C, which consists of alternating between the two losing games A and B. At every round of the game, the player chooses whether to play game A or B at random and depending on the outcome of the game, whether he wins or loses, updates his payoff. One observation which is relevant for the subsequent discussion is that both of the games update the payoff. One may think of it as a shared memory register, which records the status of the player's earnings. Unexpectedly, game C turns out to be a winning game.

This paradox has inspired a construction of the simplest possible quantum channels, which gives rise to superactivation phenomenon below.

We now consider three channels, $\mathcal{A}_{M_A}$, $\mathcal{B}_{M_B}$, and $\mathcal{C}_{M_C}$, where the subscript $M_i \in \mathbb{Z}, i = A, B, C$ denotes the memory register of each channel. At this point, these channels may be classical or quantum, and they encapsulate the channels which give rise to Parrondo's paradox. They will differ depending on the type of capacity for which we demonstrate superactivation and also on the type of channel, quantum or classical, we are investigating. The channel $\mathcal{C}_{M_C}^{\lambda}$ consists of the convex combination of the former two channels with the shared memory:

$$\mathcal{C}_{M_C}^{\lambda} = \lambda \mathcal{A}_{M_C} + (1 - \lambda)\mathcal{B}_{M_C}. \qquad (3)$$

In a nutshell, the general form of Parrondo's paradox for communication channels with memory works as follows. Upon sufficiently large number N of individual uses of each of the channels $\mathcal{A}_{M_A}$ and $\mathcal{B}_{M_B}$, the state of the memory registers M_A^N and M_B^N, respectively, has the property

$$\mathbb{E}\big[M_A^N\big] < M_A^0, \quad \mathbb{E}\big[M_B^N\big] < M_B^0, \qquad (4)$$

where M_i^0 for $i = A, B$ denote the initial state of the memory. However, for $\mathcal{C}_{M_C}^{\lambda}$ one has

$$\mathbb{E}\big[M_C^N\big] > M_C^0. \qquad (5)$$

We will further omit the subscript of the memory register M when it does not cause confusion regarding the channel it belongs to. It is rather counterintuitive to expect that the expected value of the memory of the joint channel can increase if the memory of an individual channel decreases with the number of its applications. Previously, this effect has been demonstrated to occur in a variety of physical systems [10], but it has never been considered in the context of information theory.

The gist of the Parrondo's paradox for communication channels consists of constructing channels which take advantage of this unexpected reversal in the growth of the value of the memory register.

III. SUPERACTIVATION OF THE QUANTUM AND PRIVATE CAPACITY

We now turn to the explicit construction of quantum channels with memory which demonstrate the superactivation-like effect with independent inputs across the channel while utilizing only erasure channels. Each of the channels is an erasure channel $\mathcal{N}_p$ defined as

$$\mathcal{N}_p(\rho) = (1 - p)\rho + p|e\rangle\langle e|. \qquad (6)$$

This channel is arguably one of the simplest nontrivial quantum channels. It is known that for $p \in [\frac{1}{2}, 1]$ the channel has zero quantum and private capacity [5], and it has strictly positive capacity for all values of $p \in [0, \frac{1}{2})$.

We consider two channels with memory $\mathcal{A}_M$ and $\mathcal{B}_M$, where M denotes a classical memory subsystem. We show below that each of the channels has zero capacity, but their convex combination results in a channel with positive capacity.

The first channel, $\mathcal{A}_M$, operates as follows:

(1) $\mathcal{N}_{p_a}$ with $p_a = 0.5$ acts on the input state ρ.

(2) If erasure takes place, the value of M becomes $\widetilde{M} = M - 1$. Otherwise, $\widetilde{M} = M + 1$.

The second one, $\mathcal{B}_M = \mathcal{T}_M^{M_0} \circ \mathcal{P}_M$, is the composition of two channels $\mathcal{P}_M$ and $\mathcal{T}_M^{M_0}$ acting sequentially. The channel $\mathcal{P}_M$ acts on the input state ρ with one of the erasure channels $\mathcal{N}_{p_b}$ or $\mathcal{N}_{p_c}$ depending on the state of the classical memory M:

(1) Depending on the value of M, $\mathcal{P}_M$ acts as follows:

$$\mathcal{P}_M(\rho) = \begin{cases} \mathcal{N}_{p_b}(\rho), & \text{if} \quad M \bmod 3 = 0, \\ \mathcal{N}_{p_c}(\rho), & \text{otherwise.} \end{cases} \qquad (7)$$

(2) If the application of $\mathcal{N}_{p_b}$ or $\mathcal{N}_{p_c}$ results in erasure, then the value of M becomes $\widetilde{M} = M - 1$. Otherwise, $\widetilde{M} = M + 1$.

The channel $\mathcal{T}_M^{M_0}$ acts as follows:

$$\mathcal{T}_M^{M_0}(\rho) = \begin{cases} \mathcal{N}_0(\rho), & \text{if} \quad M > M_0, \\ \mathcal{N}_1(\rho), & \text{if} \quad M \leqslant M_0, \end{cases} \qquad (8)$$

where $\mathcal{N}_0(\rho) = \rho$, $\mathcal{N}_1(\rho) = |e\rangle\langle e|$. Then, $\mathcal{B}_M = \mathcal{P}_M \circ \mathcal{T}_M^{M_0}$ and acts on the input state ρ as follows:

(1) $\mathcal{P}_M$ acts on the input state ρ.

(2) Its output is forwarded to $\mathcal{T}_M^{M_0}$.

(3) $\mathcal{T}_M^{M_0}$ reads out the value of M, and if $M \geqslant M_0$, then $\mathcal{N}_0$ acts on the state. Otherwise, the $\mathcal{N}_1$ acts.

(4) The output of $\mathcal{T}_M^{M_0}$ is forwarded to the receiver.

We note that $\mathcal{B}_M$ is a Markov-type channel in the sense that its action depends only on what happened in the previous step. It is also known to be a *forgetful channel*. Forgetful channels received a significant amount of attention in recent years due to their tractable properties and simple yet powerful structure: an arbitrary quantum channel with memory can be approximated by a sequence of forgetful channels [21]. Being one of the simplest models of the channels with memory, they admit explicit bounds for their classical and quantum capacity [22,23]. In addition to being forgetful it is easy to see that they are also *indecomposable* channels [22]: their operation does not depend on the input state, but only on the internal state of the memory.

The following lemma shows that for certain values of p_b, p_c channel $\mathcal{B}_M$ has zero quantum capacity. The proof of the lemma uses the notion of degradable channels with memory defined below. The notion of degradable channel, first introduced by [24], was only considered in the context of memoryless channels, and what follows is the natural generalization of this notion for the class of quantum channels with memory. For such channels, one cannot talk about the Choi-Jamiołkowski state of the channel, and therefore, the generalization of degradability (and antidegradability) relies on the properties of the output of the channel with memory after the finite number of its uses.

Definition 1. Channel $\mathcal{N}_M$ with memory register M is degradable if

$$\forall n \,\exists\, \text{CPTP map } \mathcal{D}^{(n)} : \mathcal{N}_M^{c,(n)} = \mathcal{D}^{(n)} \circ \mathcal{N}_M^{(n)}, \qquad (9)$$

where $\mathcal{N}_M^{(n)}$ and $\mathcal{N}_M^{c,(n)}$ represent n uses of $\mathcal{N}_M$ and its complementary channel, respectively. Similarly, one defines antidegradable channel with memory as follows.

Definition 2. Channel $\mathcal{N}_M$ with memory M is antidegradable if

$$\forall n \,\exists\, \text{CPTP map } \mathcal{D}^{(n)} : \mathcal{N}_M^{(n)} = \mathcal{D}^{(n)} \circ \mathcal{N}_M^{c,(n)}. \qquad (10)$$

Analogous to the memoryless case, the antidegradable channel with memory has zero quantum capacity.

Observation 1. Let $\mathcal{N}_M$ be the antidegradable channel with memory and $\mathcal{N}_M^{(n)}$ denote its use n times. Then for all inputs σ the following is true:

(1) $I_c(\sigma, \mathcal{N}_M^{(n)}) \leqslant 0$.
(2) $\mathcal{Q}(\mathcal{N}_M^{(n)}) = 0$.
(3) $\mathcal{Q}(\mathcal{N}_M) = 0$.

The proof of the above observation is identical to the proof of Proposition 1 in [25].

Now we are ready to prove the statement of the lemma.

Lemma 1. Consider channel $\mathcal{P}_M$, which uses erasure channels $\mathcal{N}_{p_b}, \mathcal{N}_{p_c}$ with erasure probabilities $p_b = 0.9 + \epsilon$, $p_c = 0.25 + \epsilon$, $\epsilon = 0.01$. The classical memory register M is initialized with 0. Then channel $\mathcal{B}_M$, whose action is defined in steps 1–4, above has

$$\mathcal{Q}(\mathcal{B}_M) = 0. \qquad (11)$$

Proof. As we will show below the *effect* of $\mathcal{B}_M$ on the input will be that of an erasure channel with extremely high erasure probability, and for each block of output states of finite length which arrive for Bob, Eve can degrade her output to match his. This amounts to establishing that channel $\mathcal{B}_M$ is antidegradable and thus has zero capacity.

It will suffice to establish the statement of Lemma 1 for the tensor power inputs $\rho_{AA'}^{\otimes n}$ to the channel because, as will become evident later, neither entangled nor classically correlated inputs across the uses of the channel will provide any advantage.

The state of the memory M changes only upon the action of the Markov channel $\mathcal{P}_M$, and despite the unbounded range of integer values that M can take, its state may be concisely represented by three principal states $\{|i\rangle\langle i|\}_{i=0}^{2}$, which correspond to those values of M for which $M \bmod 3 = i$. Because our channel is Markov, after a large number n of uses of $\mathcal{P}_M$ the state of the memory will approach its stationary state $M_{st} = \pi_0 |0\rangle\langle 0| + \pi_1 |1\rangle\langle 1| + \pi_2 |2\rangle\langle 2|$, where $\{\pi_i\}_{i=0}^{2}$ are the probabilities of being in state $|i\rangle\langle i|$. From the Markov property of the channel it also follows that

$$\|M - M_{st}\|_1 \leqslant \delta(n), \qquad (12)$$

where $\delta(n)$ vanishes exponentially quickly in the limit $n \to \infty$.

By choosing a large enough M_0 we ensure that the state of the underlying Markov chain of $\mathcal{P}_M$ is arbitrary close to the stationary state before any state will get transmitted to Bob via $\mathcal{T}_M^{M_0}$. Therefore, we will limit the analysis of $\mathcal{B}_M$ to this case.

We find $\{\pi_i\}_{i=0}^{2}$ from the system of equations for the stationary distribution of the Markov chain:

$$(1 - p_c)\pi_2 + p_c\pi_1 = \pi_0, \quad (1 - p_b)\pi_0 + p_c\pi_2 = \pi_1,$$
$$(1 - p_c)\pi_1 + p_b\pi_0 = \pi_2, \quad \pi_0 + \pi_1 + \pi_2 = 1. \qquad (13)$$

This gives $\pi_0 \approx 0.3844$. The success probability of Alice transmitting state ρ through $\mathcal{P}_M$ when the memory is in the

stationary state is

$$\Pr(x = \rho) = (1 - p_b)\pi_0 + (1 - p_c)(\pi_1 + \pi_2)$$
$$= (1 - p_b)\pi_0 + (1 - p_c)(1 - \pi_0)$$
$$\approx 0.49914,$$

where x denotes the state that was output by $\mathcal{P}_M$. Now, consider channel $\mathcal{B}_M$ in its entirety. In order for any state from $\mathcal{P}_M$ to be communicated to Bob, the value of M must exceed M_0. To understand the behavior of M we consider a random walk on $\mathbb{Z}$ starting at zero and moving left (decreasing the value of M by 1) with probability p and right (increasing the value of M by 1) with probability $q = 1 - p$, where

$$p = \Pr(x = \rho) = 0.49914, \quad q = \Pr(x = |e\rangle\langle e|) = 0.50086.$$

The value of M after n steps is described as the random walk $S_n = \sum_{i=1}^{n} X_i$ of n random variables $\{X_i\}_{i=1}^{n}$, each of which takes values $+1$ with probability p and -1 with probability q. Note that $\mathbb{E}[S_n] = n\mathbb{E}[X_1] = -2\alpha n$, where $\alpha = 0.009$, and it linearly decreases with each subsequent transmission.

We are now ready to show that the effect of $\mathcal{B}_M$ on the input state is that of an erasure channel $\mathcal{N}_s$, with the probability of erasure $s \gg \frac{1}{2}$. Consider the probability of Bob getting a block $L_k^r = l_1 \cdots l_k$ of fixed length k through $\mathcal{B}_M$ on steps $n, n+1, \ldots, n+k-1$ of transmission. Each l_i is either the erasure flag or a faithfully transmitted state. More formally, let $l_{i_1} = \cdots = l_{i_r} = \rho$, with $i_1 \leqslant \cdots \leqslant i_r, r \geqslant 1$, and $l_j = |e\rangle\langle e|$ for $j \in \{1, \ldots, k\} \setminus \{i_1, \ldots, i_r\}, r \leqslant k$. Because $\mathcal{B}_M$ is the composition of channels $\mathcal{T}_M^{M_0}$ and $\mathcal{P}_M$, the probability of transmitting a sequence of states L_k via $\mathcal{B}_M$ factorizes as follows:

$$\Pr(L_k \text{ is sent to Bob}) = \Pr(\mathcal{P}_M \text{ outputs } L_k)$$
$$\times \Pr(S_{i_1} \geqslant M_0, \ldots, S_{i_r} \geqslant M_0).$$

We can upper bound the right-hand side above as

$$\Pr(\text{Bob receives } L_k) \leqslant \Pr(S_{i_1} \geqslant M_0, \ldots, S_{i_r} \geqslant M_0)$$
$$\leqslant \Pr(S_n \geqslant M_0).$$

To estimate $\Pr(S_n \geqslant M_0)$ we make use of Hoeffding's inequality [26,27]:

$$\Pr(S_n \geqslant M_0) = \Pr(S_n - \mathbb{E}[S_n] \geqslant M_0 - \mathbb{E}[S_n])$$
$$= \Pr(S_n + 2\alpha n \geqslant M_0 + 2\alpha n)$$
$$\leqslant \exp\left(-\frac{2(M_0 + 2\alpha n)^2}{4n}\right). \qquad (14)$$

Therefore,

$$\Pr(\text{Bob receives } L_k) \leqslant \exp(-2\alpha^2 n). \qquad (15)$$

To ensure that entangled inputs do not help, it is enough to show that the effect of $\mathcal{B}_M$ on the input is that of a memoryless erasure channel. Instead of the fixed block L_k, we consider the general form of the block transmitted to Bob of length k, $S_k^B = s_1 \cdots s_k$ and show that for each i the port s_i was obtained by some erasure channel with a probability of erasure greater than 0.5. The ith position in S_k^B contains the mixture $s_i^B = r\rho + (1 - r)|e\rangle\langle e|$, where from the above analysis we get $r = \exp[-(n - k + i)C]$ for some constant $C > 0$, which depends only on p_b and p_c. Picking large enough M_0, we can

guarantee $r \ll 0.5$ for all k, and the action of $\mathcal{B}_M$ on the input state is that of $\mathcal{N}_{1-r}$ with $r \to 0$ when $n \to \infty$.

One can understand this as follows: with each step, the expectation of the random walk $\mathbb{E}[S_n]$ travels towards $-\infty$ with a speed which is linear in n, i.e., the difference $|M_0 - S_n|$ grows linearly. The law of large numbers states that after n steps $S_n \in (-\sqrt{n} + \mathbb{E}[S_n]; \mathbb{E}[S_n] + \sqrt{n})$ with high probability, but this is not enough because we need to guarantee that deviating from the average linearly is highly unlikely. A large deviation bound confirms this intuition: the probability of being polynomially far from this region after n steps vanishes exponentially.

Having established that regardless of input the total effect of $\mathcal{B}_M$ on each of the ports of S_k^B is tantamount to that of a particular erasure channel with probability of erasure $(1 - r) \to 1$, the optimal coding for $\mathcal{B}_M$ must be the same as the optimal coding for the erasure channel $\mathcal{N}_s$. We know that its capacity is achieved on tensor product inputs [5].

Last, what remains to be proved is that $\mathcal{B}_M$ has zero capacity. It suffices to show that the complementary channel $\mathcal{B}_M^c$ is degradable, which implies that $\mathcal{B}_M$ is antidegradable and thus has zero quantum capacity. This amounts to showing that for each output block S_k^B of length k in Bob's possession there exists degrading map $\mathcal{D}_k$ which can degrade Eve's output S_k^E to match Bob's. Again, we will consider the stationary case (number of uses of the channel $n \gg 1$) because in the nonstationary case the probabilities of getting the state or the erasure flag will be weighted by $\delta(n)$ from (12), which can be made arbitrarily small. Also, we consider the nontrivial regime when $M > M_0$ and the channel $\mathcal{T}_M^{M_0}$ becomes $\mathcal{N}_0$. The ith position of S_k^E contains the mixture $s_i^E = (1 - r)\rho + r|e\rangle\langle e|$. Showing that Eve can degrade every position of S_k^E to match S_k^B is sufficient to establish degradability of the whole block. Let $\mathcal{D}_{k,i}$ be the degrading map that acts on the ith position of the block. Then, it can be seen as performing the coin flip, with probability t returning s_i^E and with probability $1 - t$ returning $|e\rangle\langle e|$, i.e., performing the map $(1 - r)\rho + r|e\rangle\langle e| \to t[(1 - r)\rho + r|e\rangle\langle e|] + (1 - t)|e\rangle\langle e| = t(1 - r)\rho + [1 - t(1 - r)]|e\rangle\langle e|$. This yields $t = r/(1 - r)$, and the degrading map for S_k^E has the form $\mathcal{D}_k = \mathcal{D}_{k,1} \circ \cdots \circ \mathcal{D}_{k,k}$. ∎

Remarkably, transmitting quantum states using the convex mixture of two zero-capacity channels, $\mathcal{A}_M$ and $\mathcal{B}_M$, in which both of the channels have access to the shared memory register M enables the sender to convey quantum information reliably.

More precisely, consider the channel $\mathcal{C}_M^\lambda = \lambda\mathcal{A}_M + (1 - \lambda)\mathcal{B}_M$, with $\lambda \in (0,1)$ and the classical memory M shared between the channels. Lemma 2 below shows that the quantum capacity of such a convex mixture for $\lambda = 0.5$ is strictly positive: the capacity of $\mathcal{C}_M^\lambda$ experiences a superactivation-like effect. Note that this is qualitatively different from establishing the nonconvexity condition for the quantum capacity which states that the capacity of the channels in the mixture is larger than the mixture of individual capacities of the respective channels. Nonconvexity is merely the necessary condition for superactivation, which presents the strongest violation of additivity, because the channels in the mixture need not have zero capacity.

We now show that for $\lambda = 0.5$ such a convex mixture of zero-capacity erasure channels with memory has positive capacity.

Lemma 2. Consider the channel $\mathcal{C}_M^{0.5} = 0.5\mathcal{A}_M + 0.5\mathcal{B}_M$, where the underlying erasure channels of $\mathcal{B}_M$ (which form $\mathcal{P}_M$) are $\mathcal{N}_{p_b}, \mathcal{N}_{p_c}$ with erasure probabilities $p_b = 0.9 + \epsilon$, $p_c = 0.25 + \epsilon$, $\epsilon = 0.01$ and M_0 is large. The state of the memory register M is initially zero. Then,

$$\mathcal{Q}\left(\mathcal{C}_M^{0.5}\right) > 0. \tag{16}$$

Proof. It is sufficient to restrict ourselves to the tensor product inputs and the stationary state of the memory M_{st} by picking a large enough M_0 as in the proof of Lemma 1. The value of M after n steps of transmission will be determined by the random walk $S_n = \sum_{i=1}^n X_i$, where each of the random variables $\{X_i\}_{i=1}^n$ takes values $+1$ with probability p and -1 with probability $q = 1 - p$, with p and q computed below. Unlike the case with $\mathcal{B}_M$, the memory register of $\mathcal{C}_M^{0.5}$ can change in two distinct ways: when $M \bmod 3 = 0$, then half of the time $\mathcal{P}_M$ has acted on the input with $\mathcal{N}_{p_b}$ transmitting the input perfectly with probabilities $1 - p_b$ and half of the time with $\mathcal{N}_{p_a}$, which sent the state intact with probability $1 - p_a$. Similarly, in other cases, half of the time $\mathcal{N}_{p_c}$ acted on the input state transmitting the input perfectly with probability $1 - p_c$, and half of the time $\mathcal{N}_{p_a}$ sent the state perfectly with probability $1 - p_a$. This results in new probabilities r_b, r_c of successfully sending the state when the memory is divisible by 3 and when it is not, respectively:

$$r_b = \tfrac{1}{2}(1 - p_a) + \tfrac{1}{2}(1 - p_b), \tag{17}$$

$$r_c = \tfrac{1}{2}(1 - p_a) + \tfrac{1}{2}(1 - p_c). \tag{18}$$

Then, solving system (13) with new probabilities of success $(1 - p_b) = r_b$ and $(1 - p_c) = r_c$, we find $\widetilde{\pi}_0 = 0.345$ and

$$p = \Pr(x_k = \rho) = r_b\widetilde{\pi}_0 + r_c(1 - \widetilde{\pi}_0) \approx 0.5078,$$
$$q = 1 - p \approx 0.4922.$$

Then $\mathbb{E}[S_n] = n\mathbb{E}[X_1] = 2\alpha n$, where $\alpha = 0.0078$. Unlike the random walk, which underpins $\mathcal{B}_M$ in Lemma 1, this random walk is biased in the positive direction, and for a sufficiently large number n of uses of $\mathcal{C}_M^{0.5}$ the value of M will be larger than any fixed constant M_0. To formalize this intuition, define $x_i = \rho$, $i = 1, \ldots, n$ to be Alice's input to $\mathcal{C}_M^{0.5}$ and y_i, $i = 1, \ldots, n$ to be Bob's output. The probability of Alice transmitting the input state ρ to Bob successfully on the nth use of the channel is given by

$$
\begin{aligned}
\Pr(y_n &= \rho) \\
&= \Pr(x_n = \rho)[1 - \Pr(S_n \leqslant M_0)] \\
&= p[1 - \Pr(\widetilde{S}_n \geqslant -M_0)] \\
&= p[1 - \Pr(\widetilde{S}_n - \mathbb{E}[\widetilde{S}_n] \geqslant -M_0 - \mathbb{E}[\widetilde{S}_n])] \\
&= p[1 - \Pr(\widetilde{S}_n + 2\beta n \geqslant -M_0 + 2\beta n)] \\
&= p\left[1 - \exp\left(-\frac{(-M_0 + 2\beta n)^2}{4n}\right)\right], \tag{19}
\end{aligned}
$$

where $\widetilde{S}_n = \sum_{i=1}^n \widetilde{X}_i$ is the symmetric reflection of the random walk S_n, with $\widetilde{X}_i$ being $+1$ with probability q and -1 with probability p and $\mathbb{E}\widetilde{S}_n = 2\beta n$, $\beta = -\alpha$. We applied Hoeffding's inequality in Eq. (19).

After some finite number of steps n_0 channel $\mathcal{T}_M^{M_0}$ starts acting as the identity channel, and the probability of sending state ρ to Bob is given by $\Pr(y_n = \rho)$. Note that $\lim_{n\to\infty} \Pr(y_n = \rho) = \lim_{n\to\infty} p[1 - \exp(-\frac{(-M_0 + 2\beta n)^2}{4n})] = p$. For large (finite) n the effect of $\mathcal{C}_M^{0.5}$ is that of an erasure channel $\mathcal{N}_s$ with $s < 0.5$, which has positive capacity. ∎

The simple structure of the underlying channels in the construction of $\mathcal{A}$, $\mathcal{B}_M$, and $\mathcal{C}_M^{0.5}$ makes it possible to make a similar statement about their private capacity. From the proofs of Lemmas 1 and 2 it immediately follows that the private capacity of the two channels and their convex combination can be superactivated.

Corollary 1. For the quantum channels $\mathcal{A}$, $\mathcal{B}_M$, and $\mathcal{C}_M^{0.5}$ we have

$$\mathcal{P}(\mathcal{A}_M) = 0, \quad \mathcal{P}(\mathcal{B}_M) = 0, \quad \mathcal{P}(\mathcal{C}_M^{0.5}) > 0. \quad (20)$$

IV. SUPERACTIVATION OF THE CAPACITY OF A CLASSICAL CHANNEL WITH MEMORY

The purpose of this section is to show that using Parrondo's paradox, one can construct classical channels with memory for which one can demonstrate superactivation of the capacity.

Consider $\mathcal{A}_{M_A}(x)$ to be the natural generalization of the binary symmetric channel with memory and probability p of confusing the input $x \in \{0,1\}$:

$$\begin{aligned}
\mathcal{A}_{M_A}(0) &= 0 \quad \text{with} \quad p = 0.5 \quad \text{and} \quad \widetilde{M}_A = M_A + 1, \\
\mathcal{A}_{M_A}(0) &= 1 \quad \text{with} \quad p = 0.5 \quad \text{and} \quad \widetilde{M}_A = M_A - 1, \\
\mathcal{A}_{M_A}(1) &= 1 \quad \text{with} \quad p = 0.5 \quad \text{and} \quad \widetilde{M}_A = M_A + 1, \\
\mathcal{A}_{M_A}(1) &= 0 \quad \text{with} \quad p = 0.5 \quad \text{and} \quad \widetilde{M}_A = M_A - 1.
\end{aligned} \quad (21)$$

The capacity of $\mathcal{A}_{M_A}$ is

$$\mathcal{C}(\mathcal{A}_{M_A}) = 1 - H(0.5) = 0, \quad (22)$$

where $H(p)$ is the binary entropy function.

The second channel, $\mathcal{B}_{M_B}$, has an identical construction to the one described in Sec. III with the only difference being that instead of quantum states the channel takes input $x \in \{0,1\}$. Following the proof of Lemma 1 up to the step where we prove degradability amounts to establishing that the probability of successful decoding for the rate $R > 0$ decays exponentially. This implies that $\mathcal{B}_{M_B}$ has vanishing capacity and $\lim_{n\to\infty} \mathcal{C}(\mathcal{B}_{M_B}^{(n)}) = 0$, where the superscript indicates the number of uses of $\mathcal{B}_{M_B}$. Now, consider channel $\mathcal{C}_{M_C}^{0.5}$, composed of channels $\mathcal{A}_{M_C}$ and $\mathcal{B}_{M_C}$ as in Eq. (3). After sufficiently many uses of the channel, the memory of the channel $\mathcal{C}_{M_C}^{\lambda}$ becomes $M_C \gg M_0$, and we are left with the mixture of

$$\widetilde{\mathcal{C}}_{M_C}^{0.5} = 0.5\mathcal{A}_{M_C} + 0.5\mathcal{P}_{M_C}. \quad (23)$$

Following the calculation of Lemma 2, it is straightforward to see that $\mathcal{C}(\mathcal{C}_{M_C}^{0.5}) > 0$.

V. DISCUSSION

In our paper we have shown how Parrondo's paradox naturally leads to constructions which demonstrate the superactivation of the capacity for classical and quantum channels with memory. We constructed two quantum zero-capacity channels which consist of the combination of the simplest erasure channels. Adding classical memory to the construction made it possible to superactivate the capacity of the mixture of those channels. This is especially surprising in light of the fact that the quantum capacity of erasure channels cannot be superactivated by one or more memoryless erasure channels with zero capacity. Adding classical memory lifts this necessary requirement, which manifests itself for the known examples of superactivation in the memoryless case.

We have also shown the analogous construction for classical communication channels with memory, demonstrating superactivation of their classical capacity. The effect which was only known to exist for the quantum capacity of the memoryless quantum channel now becomes possible for other types of capacities of quantum as well as classical communication channels once we allow classical memory.

There were two distinct kinds of erasure channels $\mathcal{N}_p$ used in the construction of the quantum channels with memory: degradable ($p < 0.5$) and antidegradable ($p \geqslant 0.5$). Despite similar descriptions, their properties are significantly different. The former class of channels does not have a tractable analytical characterization. Channels of this kind do not even form a closed set. On the other hand, the class of all antidegradable channels, being closed and convex, admits tractable characterization. For many practical applications one may send quantum information through the channel which may be either degradable or antidegradable without full knowledge of which of the two classes it belongs to. The natural example is sending quantum information through the erasure channel when there is some uncertainty about its erasure probability or when erasure probability varies. The open question is to find the capacity when the transmission is performed using channels which belong to one of the classes above.

In the context of memoryless quantum channels, it is known that the quantum capacity can be superactivated. It is plausible that the private capacity can be superactivated as well, but this question is still open despite attempts to resolve it [7,8]. Finding a suitable way of approximating the channels with memory in our construction with their memoryless counterparts will lead to the resolution of this question.

The nontrivial channel in our construction, $\mathcal{B}_{M_B}$, consists of the composition of two channels, $\mathcal{P}_M$ and $\mathcal{T}_M^{M_0}$. However, the use of $\mathcal{T}_M^{M_0}$ is rather artificial. It was necessitated by the fact that it was not possible to show that channel $\mathcal{P}_M$ alone has zero capacity. It makes the construction insensitive to the nature of the underlying channels, quantum or classical, used for superactivation. Therefore, an open question is whether there exists a way to achieve the genuine quantum superactivation effect for quantum channels with memory, which is impossible for the classical channels with memory.

[1] G. Smith and J. Yard, Science **321**, 1812 (2008).

[2] P. Horodecki, M. Horodecki, and R. Horodecki, J. Mod. Opt. **47**, 347 (2000).

[3] F. G. S. L. Brandao, J. Oppenheim, and S. Strelchuk, Phys. Rev. Lett. **108**, 040501 (2012).

[4] S. Strelchuk and J. Oppenheim, Phys. Rev. A **86**, 022328 (2012).

[5] C. H. Bennett, D. P. DiVincenzo, and J. A. Smolin, Phys. Rev. Lett. **78**, 3217 (1997).

[6] D. Leung, J. Lim, and P. Shor, Phys. Rev. Lett. **103**, 240505 (2009).

[7] K. Li, A. Winter, X. Zou, and G. Guo, Phys. Rev. Lett. **103**, 120501 (2009).

[8] G. Smith and J. A. Smolin, Phys. Rev. Lett. **103**, 120503 (2009).

[9] F. G. S. L. Brandao, J. Eisert, M. Horodecki, and D. Yang, Phys. Rev. Lett. **106**, 230502 (2011).

[10] G. P. Harmer and D. Abbott, Nature (London) **402**, 864 (1999).

[11] C. Macchiavello and G. M. Palma, Phys. Rev. A **65**, 050301 (2002).

[12] P. V. E. McClintock, Nature (London) **401**, 23 (1999).

[13] R. Benzi, A. Sutera, and A. Vulpiani, J. Phys. A **14**, L453 (1981).

[14] M. Bier, Contemp. Phys. **38**, 371 (1997).

[15] F. Caruso, S. F. Huelga, and M. B. Plenio, Phys. Rev. Lett. **105**, 190501 (2010).

[16] J. M. R. Parrondo, G. P. Harmer, and D. Abbott, Phys. Rev. Lett. **85**, 5226 (2000).

[17] G. P. Harmer and D. Abbott, Stat. Sci. **14**, 206 (1999).

[18] T. Philips and A. Feldman, Social Science Research Network, SSRN Scholarly Paper No. 581521, 2004 (unpublished).

[19] L. Schimansky-Geier, *Noise in Complex Systems and Stochastic Dynamics* (SPIE, Bellingham, WA, 2003).

[20] G. P. Harmer, D. Abbott, and P. G. Taylor, Proc. R. Soc. London, Ser. A **456**, 247 (2000).

[21] D. Kretschmann and R. F. Werner, Phys. Rev. A **72**, 062323 (2005).

[22] G. Bowen, I. Devetak, and S. Mancini, Phys. Rev. A **71**, 034310 (2005).

[23] N. Datta and T. C. Dorlas, J. Phys. A **40**, 8147 (2007).

[24] I. Devetak and P. W. Shor, Commun. Math. Phys. **256**, 287 (2005).

[25] A. S. Holevo, Probl. Inf. Transm. (Engl. Transl.) **44**, 171 (2008).

[26] W. Hoeffding, J. Am. Stat. Assoc. **58**, 13 (1963).

[27] W. K. Wootters and W. H. Zurek, Nature (London) **299**, 802 (1982).