Noisy Interactive Quantum Communication

Gilles Brassard ∗† Alain Tapp ∗‡ Dave Touchette ∗§

May 11, 2014

Abstract

We study the problem of simulating protocols in a quantum communication setting over noisy channels. This problem falls at the intersection of quantum information theory and quantum communication complexity, and is of particular importance for real-world applications of interactive quantum protocols, which can be proved to have exponentially lower communication costs than their classical counterparts for some problems. To the best of our knowledge, these are the first results regarding the quantum version of this problem, first studied by Schulman in a classical setting (FOCS ’92, STOC ’93). We simulate a length $N$ quantum communication protocol by a length $O(N)$ protocol with arbitrarily small error. Our simulation strategy has a far higher communication rate than the naive one that encodes each particular round of communication to achieve comparable success. In particular, such a strategy would have a communication rate going to 0 in the worst interaction case as the length of the protocols increases, in contrast to our strategy, which has a communication rate proportional to the capacity of the channel used. Under adversarial noise, our strategy can withstand, for arbitrarily small $\varepsilon > 0$, error rates of $\frac{1}{2} - \varepsilon$ when parties preshare perfect entanglement, and this even if they are only allowed noisy classical communication. We show that this is optimal. This is in contrast to the case of the naive strategy, which would not work for any constant fraction of errors in this model. When the parties do not preshare entanglement, we show how to tolerate adversarial error rates close to the maximum tolerable for one-way quantum data transmission. In a random noise setting with a quantum channel of capacity $Q > 0$, the communication rate is proportional to $Q$. We also give simulation protocols with linear communication and entanglement consumption when parties pre-share noisy entanglement and communicate over noisy classical channels. Our results are stated for a general quantum communication protocol in which Alice and Bob collaborate, and are easily seen to hold in a quantum communication complexity setting for Yao’s and Cleve-Buhrman’s models or in other relevant distributed quantum computation models.

∗Laboratoire d’informatique théorique et quantique, Département d’informatique et de recherche opérationnelle, Université de Montréal
†brassard@iro.umontreal.ca
‡alain.tapp@gmail.com
§touchetd@iro.umontreal.ca
1 Introduction

Quantum information theory is well-developed for information transmission over noisy quantum channels, dating back to the work of Holevo in the 70’s [25, 26], either for the transmission of classical information [27, 39], quantum information [32, 41, 19], and even if we allow for pre-shared entanglement between sender and receiver [6, 7]. It describes the ultimate limits for (unidirectional) data transmission over noisy quantum channels without concern for explicit, efficient construction of codes. Closely related is the area of quantum coding theory, which takes a more practical approach toward the construction of quantum error correcting codes (QECC) [40, 42] by providing explicit and efficient constructions [15, 42, 23] of codes, and by providing bounds on their existence [14, 20, 35].

Quantum communication complexity has also been studied in depth since Yao’s seminal paper introduced the field in 1993 [46]. It is an idealized setting in which local computation is free and communication is noiseless but expensive, and two parties want to compute a classical function of their joint input while minimizing the number of qubits they have to exchange to do so. Exponential separations have been shown for some promise problems between their classical and quantum communication complexity [13], even if we allow for some bounded error [43]. Moreover, for both classical and quantum communication complexity, interaction has been proved to be a powerful resource: exponential separations in the communication complexity of some functions have also been established between protocols restricted to some bounded number of messages $k$, and protocols with $k + 1$ messages [34, 24]. Cleve and Buhrman [16] defined an alternative model for communication complexity in a quantum setting, in which players are allowed to pre-share an arbitrary entangled state but transmit classical rather than quantum bits. This model is at least as powerful as Yao’s (up to a factor of 2), since entanglement along with 2 classical bits can be used to teleport an arbitrary quantum state. It is still an open question whether the two models are essentially equivalent, since no good bound on the amount of entanglement required in the Cleve-Buhrman model is known.

However, quantum communication, even more so than classical communication, is prone to transmission errors in the real world. Moreover, with the ubiquity of distributed computing nowadays, it has become increasingly important to develop an information and coding theory for interactive protocols. To the best of our knowledge, this problem has never been studied before in a quantum setting. In the realm of classical communication, Schulman initiated the field with his pioneering works [36, 37, 38] showing that it is possible to simulate any protocol defined over a noiseless channel over a noisy channel with exponentially small probability of error while only dilating the protocol by a constant factor. This multiplicative dilation factor, in the case of a binary symmetric channel, is proportional to the inverse of the capacity, as in the data transmission case. However, it does not go to 1 asymptotically in this case. For the case of adversarial errors, Schulman also shows how to withstand up to a corruption rate of $\frac{1}{240}$. Recent work by Braverman and Rao [12] shows how to withstand error rates of $\frac{1}{4} - \varepsilon$ in the case of an adversarial channel, and they also show this is optimal in their model of noisy communication. Even more recently, Franklin, Gelles, Ostrovsky and Schulman [21] were able to show that in an alternative model in which Alice and Bob
are allowed to share a secret key unknown to the adversary Eve, they can withstand error rates up to \( \frac{1}{2} - \varepsilon \), which is also shown to be optimal in their model. All of the above works use the notion of tree codes, introduced by Schulman. These tree codes are shown to exist for various parameters, but no efficient construction is known. A relaxation of the tree code condition still strong enough for most applications in interactive coding was stated by Gelles, Moitra and Sahai [22], and they were able to provide an efficient randomized construction for these so-called potent tree codes. Using these in a random error model leads to efficient decoding on the average, hence to efficient simulation protocols (of course, given black-box access to the original protocol, which might be inefficient in itself) but in a worst-case adversarial scenario, the decoding might still take exponential time with these potent tree codes.

It was only recently that an alternative coding strategy developed by Brakerski and Kalai [9] was able to take care efficiently of the adversarial error case by cleverly splitting the whole communication into blocks of logarithmic length in which tree encoding is used but also some history information is sent in between the blocks that enables efficient decoding. This construction was further improved by Brakerski and Naor in [10]. A survey article by Braverman [11] provides a good overview of results and open questions in the area of classical interactive communication circa 2011, though some of the important questions raised there have been addressed since. In particular, the question of interactive capacity of binary symmetric channels that was recently investigated by Kol and Raz [29] for which they find that indeed, the communication capacity of the binary symmetric channel with capacity close to unity behaves differently in the asymptotic limit of long interactive protocols than in the data transmission case.

The approach taken in all of the above is inherently classical and does not generalize well to the quantum setting. In particular, the fact that classical information can be copied and resent multiple times is implicitly used, and therefore the fact that the information in the communication register can be destroyed by noise is without consequence. By contrast, in the quantum case, if the information in some communication register is destroyed, it could not have been copied before, and in particular it cannot be resent. A naive strategy, which applies in the quantum as well as in the classical case, would be to encode each round separately. But, in a random error model, a constant dilation of each round would not be sufficient in the worst case of one-qubit transmission to reach good fidelity, and a superconstant dilation leads to a communication rate of zero asymptotically. Moreover, in the case of adversarial errors, no constant rate of error can be withstood with such a strategy: the adversary can then always disrupt a whole block (unless the number of round is constant). Using properties of classical information, it was possible to design clever simulation protocols that were able to withstand constant error rates at constant communication rates, and succeed in simulating protocols designed for noiseless classical channels over noisy channels by reproducing the whole transcript of the noiseless protocol. However, it is not clear that it is possible, given an arbitrary quantum protocol designed for a noiseless bidirectional quantum channel, to simulate it over noisy quantum channels with constant error rate at a constant communication rate. Even in the case of protocols in the Cleve-Buhrman model, in which the communication is classical, it is not clear that we can achieve results similar
to what is done for classical communication protocols when we replace noiseless classical channels by noisy ones. Indeed, if a quantum measurement is performed on the entangled state shared by the two parties and it is later realized that the choice of the measurement was based on wrong information, such a measurement will in general be irreversible, and the naive approach to adapt the previous simulation to the Cleve-Buhrman model does not work.

2 Overview of Results

We show that despite the above setbacks, it is indeed possible to simulate arbitrary quantum protocols over noisy quantum channels with good communication rates. We consider two models for interaction over noisy channels: one analogous to Yao’s model, in which all communication is over noisy quantum channels but the parties do not pre-share entanglement, and one analogous to Cleve-Buhrman’s model in which all communication is over noisy classical channels but parties are allowed to pre-share noiseless entanglement. We call these models the quantum and shared entanglement models, respectively. We show that with only a constant dilation factor, it is possible to withstand error rates of $\frac{1}{2} - \varepsilon$ in the shared entanglement model, for arbitrarily small $\varepsilon > 0$, thus matching the highest tolerable error rate in the analogous shared secret key model for classical interactive communication [21].

In the quantum model, we can withstand error rates close to the best achievable for quantum data transmission. The result for the shared entanglement model is optimal when the noiseless protocol requires bidirectional communication and protocols over the noisy channel are restricted to be non-adaptive, i.e. the order in which the parties talk depends only on the protocol and maybe on the input, but not on the previous actions of the adversary. This restriction is natural for protocols over noisy channels: the view of each party is different and depends on previous errors, so this restriction enforces that they do not simultaneously try to speak over the channel. Moreover, in our simulation protocol for this model, all communication is classical, which is in general a much less expensive resource than quantum communication. The approach we take is to teleport the quantum communication register back and forth. When the register is in some party’s possession, he tries to evolve the simulation by applying one of his unitaries in the noiseless protocol, or one of its inverses if he realizes at some point he applied it wrongly before. The important point is that all operations on the quantum registers are reversible, being a sequence of noiseless protocol unitaries and random Pauli operators. Of particular importance to our work is the notion of tree codes, introduced by Schulman for the purpose of simulating classical protocols over noisy classical channels, which we use to transmit our classical information. We can adapt the techniques we develop in the shared entanglement model for the quantum communication model in which parties do not pre-share entanglement, but have access to a noisy quantum channel: we first distribute a linear amount of entanglement using standard quantum information and coding theory techniques. We can tolerate an adversarial error rate of up to $\frac{1}{6}$ in that case. We can also adapt our techniques for an adversarial error model to the case of a random error model. Then, dilation factors proportional to $\frac{1}{Q}$ for a depolarizing channel of
quantum capacity $Q$ in the quantum model, and proportional to $\frac{1}{C}$ for a binary symmetric channel of capacity $C$ in the shared entanglement model, are sufficient. We also show that the result in the shared entanglement model is asymptotically optimal: there exist a family of binary functions for which a dilation factor proportional to $\frac{1}{C}$ is necessary. We further extend the study in the shared entanglement model to consider noisy entanglement in the form of noisy EPR pairs in the so-called Werner states. For any non-separable Werner state, we give simulation protocols with linear noisy classical communication and noisy EPR pair consumption.

If we compare our approach for the quantum case to those for the classical case, as described in a recent paper on efficient interactive coding [10], the high level logic of all proposed solutions until now for classical protocol simulation can be described as follow: try to evolve the protocol, and if it is later realized there has been some error, try to have the parties go back to where they last agreed (in a protocol tree representation, this would be their least common ancestor). However, in our case, the parties roughly try to follow the same idea, but are not able to do this passively as in the classical case for two reasons. First, the underlying classical communication is not fixed at the beginning of the protocol but depends on the random outcomes of the teleportation measurement, so even when they would try to synchronize based on previous errors, they would have to actively teleport during resynchronization, leading to potentially more errors on the joint quantum register. Second, there is no underlying transcript (or protocol tree) that the parties try to synchronize on, except that they want to evolve their sequence of unitaries, and so they have to actively rewind previous unitaries and wrong teleportation decoding instead of just going back to a point in the protocol where they agreed on the previous transcript.

Since the parties need to backtrack the simulation of the noiseless protocol, we remark that it is surprising that we can tolerate error rates as high as $\frac{1}{2} - \varepsilon$. Indeed, all recent classical schemes tolerating high error rates had the property that the parties were always going forward with the communication by using the tree structure of classical protocols, in comparison to Schulman’s original tree code based scheme in which there is some form of backtracking and for which he can tolerate a much lower adversarial error rate of $\frac{1}{240}$. To obtain such a result, we follow [21] and use the fact that a blueberry code effectively turns most adversarial errors into erasure, so that concatenating such a code on top of a tree code yields a tree code with an erasure symbol. Since actual errors are twice as harmful as erasures for the tree code condition, which is stated in terms of Hamming distance, it was shown in [21] that if the error rate is below $\frac{1}{2} - \varepsilon$, then the number of rounds in which both parties correctly decode a long prefix is large enough to imply success of the simulation. This condition is not sufficient for our purpose: the number of rounds in which both parties decode correctly even the whole string could be high, but if these rounds alternate with rounds in which at least one of the parties makes a decoding error, then the protocol could stall, and simulation would fail. This is due to the fact that in our case, the parties could agree on some transcript at the end of the simulation, but this transcript could be completely useless, consisting mostly of random teleportation measurement outcome. To circumvent this possibility, Lemma [5] develops a new bound on tree codes with an erasure symbol, which
might be of independent interest for classical interactive coding, and is sufficient to obtain the same maximum tolerable error rate of $\frac{1}{2} - \varepsilon$ for our quantum protocol as was obtained in [21] with blueberry codes over the protocol of [12]. Another important ingredient in our simulation protocols is the representation for noisy quantum protocols that we develop, which is quite powerful and will be used in forthcoming papers to adapt classical results on computationally efficient interactive computation over adversarial channels [9] and on the interactive capacity of the binary symmetric channel [29] to the quantum regime. Note that due to the use of tree codes in the present paper, the protocols presented are not computationally efficient. However, as was just stated, it is possible to extend classical results on efficient interactive coding to our case, and will do so in upcoming work.

The paper is structured as follows: in section 3 we set up notation and state definitions we use, in particular those relevant for the different models of communication. In section 4 we state and prove a simpler version of our main result for the adversarial case in the shared entanglement model, and then show how to modify it to obtain optimal results in section 5. Finally, section 6 shows how to adapt the results of the previous sections to obtain various other interesting results, in particular for the quantum model and in the case of a random error model. We conclude with a discussion of our results, and further directions of research are also explored.

3 Preliminaries

3.1 Quantum Mechanics

Let us set some notation. We first briefly review the quantum formalism, mainly to set notation; for a more thorough treatment, we refer the interested reader to good introductions in a quantum information theory context [33, 44, 45]. For every quantum system $A$, we associate a finite dimensional Hilbert space, which by abuse of notation we also denote by $A$. The state of quantum system $A$ is represented by a density operator $\rho_A$, a positive semi-definite operator over the Hilbert space $A$ with unit trace. We denote by $\mathcal{D}(A)$ the set of all density operators representing states of system $A$. Composite quantum systems are represented by the (Kronecker) tensor product space of the underlying spaces, i.e. for systems $A$ and $B$, the allowed states of the composite system $A \otimes B$ are (represented by) the density operators in $\mathcal{D}(A \otimes B)$. We sometimes use the shorthand $AB$ for $A \otimes B$. The evolution of a quantum system $A$ is represented by a completely positive, trace preserving linear map (CPTP maps) $\mathcal{N}^A$ such that if the state of the system was $\rho \in \mathcal{D}(A)$ before evolution through $\mathcal{N}$, the state of the system is $\mathcal{N}(\rho) \in \mathcal{D}(A)$ after. We refer to such maps as quantum channels, and to the set of all channels acting on $A$ as $\mathcal{L}(A)$. An important quantum channel that we consider is the quantum depolarizing channel $T_\varepsilon$ with depolarizing parameter $0 \leq \varepsilon \leq 1$: it takes as input a qubit $\rho$, and outputs a qubit $T_\varepsilon(\rho) = (1-\varepsilon)\rho + \varepsilon \frac{I}{2}$, i.e. with probability $1 - \varepsilon$ it outputs $\rho$ and with complementary probability $\varepsilon$ it outputs a completely mixed state. We also consider quantum channels with different input and output systems; the set of all quantum channels from a system $A$ to a system $B$ is denoted $\mathcal{L}(A, B)$. 

6
Another important operation on a composite system $A \otimes B$ is the partial trace $\text{Tr}_B(\rho^{AB})$ which effectively gets rid of the part of the quantum state $\rho^{AB}$ in the $B$ subsystem and keeps the corresponding marginal state of the $A$ subsystem. Fixing a basis $\{|i\rangle\}$ for $B$, the action of the partial trace can be evaluated as $\text{Tr}_B(\rho^{AB}) = \sum_i \langle i | \rho | i \rangle$, and this is a valid quantum channel in $\mathcal{L}(A \otimes B, A)$.

An important special case for quantum systems are pure states, whose density operators have a special form: rank-one projectors $|\psi\rangle\langle \psi|$. In such a case, a more convenient notation is provided by the pure state formalism: a state is represented by the unit vector $|\psi\rangle$ (up to an irrelevant complex phase) the density operator projects upon. We denote by $\mathcal{H}(A)$ the set of all such unit vectors (up to equivalence of global phase). Pure state evolution is represented by a unitary operator $U^A$ acting on $|\psi\rangle^A$, denoted $U|\psi\rangle^A$. Evolution of the $B$ register of a state $|\psi\rangle^{AB}$ under the action of a unitary $U^B$ is represented by $(I^A \otimes U^B)|\psi\rangle^{AB}$, for $I^A$ representing the identity operator acting on the $A$ system, and is denoted by the shorthand $U^B|\psi\rangle^{AB}$ for convenience. We might drop the superscripts when the systems are clear from context. The evolution under consecutive action of unitaries $U_j$’s is denoted by:

$$\prod_{j=1}^\ell U_j |\psi\rangle = U_\ell \cdots U_1 |\psi\rangle. \quad (3.1)$$

We represent a classical random variable $X$ with probability density function $p_X$ by a density operator $\sigma^X$ that is diagonal in a fixed basis $\{|x\rangle\}_{x \in \mathcal{X}}$: $\sigma^X = \sum_{x \in \mathcal{X}} p_X(x) |x\rangle\langle x|^X$. For a quantum system $A$ classically correlated with a random variable $X$, we represent the corresponding classical-quantum state by the density operator $\rho^{XA} = \sum_{x \in \mathcal{X}} p_X(x) |x\rangle\langle x|^X \otimes \rho^x_A$, in which $\rho^x_A$ is the state of system $A$ conditioned on the random variable $X$ taking value $x \in \mathcal{X}$. The extraction of classical information from a quantum system is represented by quantum instruments: classical-quantum CPTP maps that take classical-quantum states on a composite system $X \otimes A$ to classical-quantum states. Viewing classical random variables as a special case of a quantum system, quantum instruments can be viewed as a special case of quantum channels.

Our simulation protocols make heavy use of the teleportation protocol between Alice and Bob [3], which uses the following resource state shared by Alice and Bob, called an EPR pair: $|\Phi^+\rangle^{TA} = \frac{1}{\sqrt{2}}(|00\rangle + |11\rangle)$, with the qubit in the $T_A$ register held by Alice, and the qubit in the $T_B$ register held by Bob. The teleportation protocol then uses one of these resource states to teleport one qubit either from Alice to Bob, or from Bob to Alice. If Alice wants to teleport a qubit $|\psi\rangle$ in the register $C$ to Bob, with whom she shares an EPR pair, she applies a joint Bell measurement, which can perfectly distinguish the Bell states $\{|\Phi_{xz}\rangle = \frac{1}{\sqrt{2}}(|0x\rangle + (-1)^z |1x\rangle)\}_{x,z \in \{0,1\}}$, to the registers $CT_A$ she holds, and obtains uniformly random measurement outcomes $xz \in \{0,1\}^2$. After this measurement, the state in the $T_B$ register is $X^z Z^z |\psi\rangle$, for $X$ and $Z$ the Pauli operators corresponding to bit flip and phase flip in the computational ($Z$) basis, respectively. If Alice transmits the two bits $xz$ to Bob, he can then decode the state $|\psi\rangle$ on the $T_B$ register by applying $(X^z Z^z)^{-1} = Z^z X^z$. Teleportation from Bob to Alice is performed similarly (EPR pairs are symmetric).
Another technique we use is that of making classical operations coherent: measurements and classically-controlled operations are replaced by corresponding unitaries (and ancilla register preparation). The coherent version of a measurement is called a pseudo-measurement: instead of measuring a qubit to obtain a binary classical outcome in \( \{0,1\} \) with probability \( p_0 \) and \( p_1 \), respectively, a classical value that can be distributed classically among two parties (or more), a pseudo-measure is applied to the quantum state, leaving it in a pure quantum state in which the two qubits will then act the same as if they had been measured, provided they do not further interact. The technique to do so uses a controlled-\( X \) gate, i.e. a gate mapping \( |x\rangle^S |b\rangle^T \) to \( |x\rangle^S |b \oplus x\rangle^T \) for some source qubit in register \( S \) to be kept at the previous measurement point, and some target qubit in register \( T \) to be distributed. Then a pseudo-measure is done by preparing a fresh ancilla qubit in state \( |0\rangle \) in some register \( T \), and if we relabel the register of the previously measured qubit by register \( S \), we simply use the controlled-\( X \) gate as described above. Then, if the \( S \) and \( T \) registers are left as is, a subsequent measurement on any one of these registers will still output 0 or 1 with probability \( p_0 \) and \( p_1 \), respectively, and if both registers are measured, both outcomes are perfectly correlated. The usefulness of this operation is that, contrary to actual measurements, they are not irreversible, and if it is later realized that a qubit should not have been measured, the pseudo-measure can be undone.

To measure the success of the simulation, we use the trace distance \( \|\rho - \sigma\|_1^A \) between two arbitrary states \( \rho^A \) and \( \sigma^A \), in which \( \|O\|_1^A = \text{Tr}(O^JO)\frac{1}{2} \) is the trace norm for operators on system \( A \). We might drop the \( A \) subscript if the system is clear from context. The trace distance has the operational interpretation to be (four times) the best possible bias to distinguish between the two states \( \rho^A \) and \( \sigma^A \), given a single unknown copy of one of these two states. To distinguish between quantum channels, we first consider the induced norm for quantum channels \( \mathcal{N} \in \mathcal{L}(A,B) \): \( \|\mathcal{N}\| = \max \{\|\mathcal{N}(\sigma)\|_1^B : \sigma \in \mathcal{D}(A)\} \). Correlations with another quantum system can help distinguish between quantum channels, so the appropriate norm to use to account for this fact is the diamond norm \([1]\): \( \|\mathcal{N}\|_{\diamond} = \|\mathcal{N} \otimes I^R\| \) for some reference system \( R \) of the same dimension as the input system \( A \). Then, for two quantum channels \( \mathcal{N}, \mathcal{M} \in \mathcal{L}(A,B) \), \( \|\mathcal{N} - \mathcal{M}\|_{\diamond} \) has the operational interpretation to be (four times) the best possible bias possible to distinguish between the two channels, given a single unknown use of one of these two channels.

### 3.2 Quantum Communication Complexity

In Yao’s model for quantum communication complexity \([16]\), Alice is given a classical input \( x \in X \), Bob a classical input \( y \in Y \), and they want to compute a classical function \( f : X \times Y \rightarrow Z \) (often \( X = Y = \{0,1\}^n, Z = \{0,1\} \)) of their joint input by communicating as few quantum bits as possible, but without regard to the local computation cost. Often, we are only interested in \( x \in X, y \in Y \) satisfying some promise \( P : X \times Y \rightarrow \{0,1\} \). A global quantum system is split into three subsystems: the \( A \) register is the register held by Alice, the \( B \) register is the one held by Bob, and the \( C \) register is the communication register, initially held by Alice, and exchanged back-and-forth by Alice and Bob in each round. Our formal description of the protocols in this model is based upon the one given in \([30]\).
An length $N$ protocol is defined by a sequence of unitaries $U_1, \cdots, U_{N+1}$ in which for $i$ odd, $U_i$ acts on the $AC$ register, and for $i$ even, $U_i$ acts on the $BC$ register. Initially, all the qubits in the $A,B,C$ registers are set to the all $|0\rangle$ state, except for $n$ qubits in the $A$ register initially set to $x \in X$, and $n$ in the $B$ register set to $y \in Y$. The number of qubits $m_A, m_B \in \mathbb{N}$ in the $A$ and $B$ registers is arbitrary (of course, $m_A, m_B \geq n$) and not taken into account in the cost of the protocol, and neither is the complexity of the $U_i$'s, since local computation is considered free. However, the number of qubits $c$ in the $C$ register is important and is taken into account in the communication cost, which is $N \cdot c$. The outcome of the protocol is obtained by measuring the first qubit of the $C$ register after application of $U_{N+1}$. The protocol succeeds if the outcome of the measurement is $f(x,y)$ with good probability for any $x,y$ (satisfying the promise).

Another model for quantum communication complexity is the one introduced by Cleve and Buhrman [16]. In their model, communication is classical, but parties are allowed to pre-share an arbitrary entangled quantum state at the outset of the protocol. We can view protocols in this model as a modification on those of Yao's model in which the initial state $|\psi\rangle$ on the $ABC$ register is arbitrary except for $n$ qubits in each of the $A,B$ registers initialized to $x,y$ respectively. Also, each qubit in the $C$ register is measured in the computational basis, and it is the outcome of these measurements that is communicated to the other party. Note that by using pseudo-measurements instead of actual measurements in each round, the parties can use quantum communication instead of classical communication. Then the two models become almost identical, except for the initial state, which is arbitrary in the Cleve-Buhrman model, and fixed to the all 0 state in Yao's model (not including each party’s classical input). Since our simulation protocols consider general unitary local processing but do not assume any particular form for the initial state, they work on this slight adaptation of the Cleve-Buhrman model as well as for Yao’s model of quantum communication complexity.

3.3 Quantum Communication Model

3.3.1 Noiseless Communication Model

In the noiseless quantum communication model that we want to simulate, there are four quantum registers: the $A$ register held by Alice, the $B$ register held by Bob, the $C$ register, which is the communication register exchanged back-and-forth between Alice and Bob and initially held by Alice, and finally the $E$ register, which purifies the initial (and then also the final) state of the $ABC$ registers and might be held by Eve, a potential adversary. The initial state $|\psi_{\text{init}}\rangle_{ABCE} \in \mathcal{P}$ is chosen arbitrarily from the set of allowed inputs $\mathcal{P}$, and is fixed at the outset of the protocol, but possibly unknown (totally or partially) to Alice and Bob. Note that to allow for composition of quantum protocols in an arbitrary environment, we consider arbitrary quantum states as input, maybe entangled with some reference system $E$. A protocol $\Pi$ is then defined by the sequence of unitaries $U_1^{AC}, U_2^{BC}, \cdots, U_{N+1}$, with $U_{2i+1}$ known at least to Alice (or given to her in a black box), and $U_{2i+2}$ known at least to Bob (or given to him in a black box). Without loss of generality, we assume $N$ is even: this affects the total cost of communication by at most one communication of the
C register. On a particular input state $|\psi_{\text{init}}\rangle$, the protocol generates the output state $|\psi_{\text{final}}\rangle^{ABCE} = U_{N+1} \cdots U_1 |\psi_{\text{init}}\rangle^{ABCE}$, for which at the end of the protocol the A and C registers are held by Alice, the B register is held by Bob, and the E register is held by Eve. We sometimes also write $\Pi(|\psi_{\text{init}}\rangle)$ for $\text{Tr}_E(|\psi_{\text{final}}\rangle^{ABCE})$, and by abuse of notation also represent the induced quantum channel from $ABCE$ to $ABC$ simply by $\Pi$. Since we consider local computation to be free, the sizes of A and B can be arbitrarily large, but still of finite size, say $m_A$ and $m_B$ qubits, respectively. Also, we consider the case of a single-qubit C register, which is the worst case for interaction. This can be done without affecting the cost of communication by more than a factor of two (if a party has to speak when it is not his turn, he sends a $|0\rangle$ qubit), but maybe at the expense of much more interaction.

Note however that it is straightforward to apply our results to registers C of arbitrary size. Also note that both Yao’s and Cleve-Buhrman’s models of quantum communication complexity can be recast in this framework by making all operations coherent: put the initial classical registers into quantum registers, replace classically controlled operations by quantumly controlled operations, also replace measurements by pseudo-measurements, and then replace any classical communication by quantum communication. In particular, this gets rid of the problem of the non-reversibility of measurements, which is especially present in the Cleve-Buhrman model.

We need to embed length $N$ protocols into others of larger length $N' > N$. To perform such noiseless protocol embedding, we define some dummy registers $\tilde{A}, \tilde{B}, \tilde{C}$ isomorphic to $A, B, C$, respectively. $\tilde{A}$ and $\tilde{C}$ are part of Alice’s scratch register and $\tilde{B}$ is part of Bob’s scratch register. Then, for any isomorphic quantum registers $\tilde{D}, \tilde{D}$, let $\text{SWAP}_{\tilde{D} \leftrightarrow \tilde{D}}$ denote the quantum unitary that swaps the $\tilde{D}, \tilde{D}$ registers. In a noiseless protocol embedding, for $i \in \{1, 2, \cdots, N - 1\}$, we leave $U_i$ untouched. We replace $U_N$ by $\text{SWAP}_{\tilde{B} \leftrightarrow \tilde{B}} \circ U_N$ and $U_{N+1}$ by $\text{SWAP}_{\tilde{A} \leftrightarrow \tilde{C}} \circ U_{N+1}$. Finally, for $i \in \{N + 2, N + 3, \cdots, N' + 1\}$, we define $U_i = I$, the identity operator.

We refer later to the unidirectional model; in this noiseless model, we allow for large local registers $A', B'$ and for a large communication register $C'$ that is used only once, either from Alice to Bob or from Bob to Alice, depending on the protocol. These registers can be further decomposed such that when used for simulation, the A register of the protocol to be simulated is a subsystem of $A'$, B is one of $B'$, and C of $A'$. We also allow for classical registers $X, Y$ held by Alice and Bob, respectively. For concreteness we consider here the case of communication from Alice to Bob; the other case is symmetric. A simulation protocol $U$ in the unidirectional model is defined by two quantum instruments $M_1^{X'A'C'}, M_2^{Y'B'C'}$, and the output of the protocol on input $|\psi\rangle \in \mathcal{H}(A \otimes B \otimes C \otimes E)$ is the $ABC$ subsystem of $M_2 M_1(|\psi\rangle)$, and is denoted $U(|\psi\rangle)$. By abuse of notation, the induced quantum channel from $ABCE$ to $ABC$ is also denoted $U$.

### 3.3.2 Noisy Communication Model

There are many possible models for noisy communication. We consider two in particular: one analogous to Yao’s model with no shared entanglement but noisy quantum communication, which we call the quantum model, and one analogous to Cleve-Buhrman’s model.
with noiseless pre-shared entanglement but noisy classical communication, which we call the shared entanglement model. A further variation on the shared entanglement model in which the entanglement is also noisy is considered in section 5.4. For simplicity, we formally define in this section what we sometimes refer to as alternating communication models, in which Alice and Bob alternatively transmit each other the communication register, and this is the model most of our protocols are defined in. However, somewhat more general models to which our definitions easily adapt are referred to as oblivious communication models, following [12], in which Alice and Bob do not necessarily transmit their messages in alternation, but nevertheless in a fixed order known to all (Alice, Bob and Eve) depending only on the round, and not on the particular input or the actions of Eve.

For the quantum model, Alice possesses a local classical-quantum register $X \otimes A'$ in which $X$ is the classical register and the quantum register $A'$ contains five subsystems of interest: to act a noiseless protocol $\Pi$ as a black-box, the $A$ and $C_A$ parts correspond to the registers of the noiseless communication protocol, while $\tilde{A}$ and $\tilde{C}_A$ are the corresponding registers defined by the noiseless protocol embedding, and $A''$ is some scratch register used for her local quantum computation in the simulation. Similarly, Bob possesses a local classical-quantum register $Y \otimes B'$ in which $Y$ is the classical register and the quantum register $B'$ contains four subsystems of interest: to act $\Pi$ as a black-box, the $B$ and $C_B$ parts correspond to the registers of the noiseless communication protocol, while $\tilde{B}$ is the corresponding register defined by the noiseless protocol embedding, and $B''$ is some scratch register used for his local quantum computation in the simulation. Eve possesses a local classical-quantum register $Z \otimes E'$ in which $Z$ is the classical register and the quantum register $E'$ contains two subsystems of interest: the $E$ part corresponds to the reference register of the noiseless communication protocol and $E''$ is some scratch register used for her local quantum computation in the simulation. A quantum communication register $C'$, of some fixed size $q$ independent of the length $N$ of the protocol to be simulated, is exchanged back-and-forth between Alice and Bob by passing through Eve’s hand; it is held by Alice at both the beginning and the end of the simulation protocol. A simulation protocol $Q$ in the quantum model is defined by a sequence of quantum instruments $\mathcal{M}_1^{A'BC'}$, $\mathcal{M}_2^{B'C'}$, $\cdots$, $\mathcal{M}_N^{A'C'}$ such that, on input a state $|\psi'_\text{init}\rangle_{A'B'C'}E' = |\psi_{\text{init}}\rangle^{ABC_AE} \otimes |0\rangle$, given black-box access to a noiseless protocol $\Pi$, and against an adversary $A^Q$ (which only has to make the simulation fail on some particular protocol, and on some particular input, to characterize the simulation protocol as bad against her) defined by a sequence of quantum instruments $\mathcal{N}_1^{ZEC'}$, $\cdots$, $\mathcal{N}_N^{ZEC'}$, the protocol outputs the $\tilde{A}\tilde{B}\tilde{C}$ subsystems of

$$\rho_{\text{final}} = \mathcal{M}_N^{\Pi} N_N \mathcal{M}_N^{\Pi} \cdots M_1^{\Pi} N_1 \mathcal{M}_1^{\Pi} (|\psi_{\text{init}}\rangle \langle \psi_{\text{init}}|).$$

We denote this output by $Q^{\Pi}(A^Q(|\psi_{\text{init}}\rangle))$, and the induced quantum channel from $ABC_E$ to $\tilde{A}\tilde{B}\tilde{C} \cong ABC$ by $Q^{\Pi}(A^Q)$. The success of the simulation is measured by how close the simulation output state is to the final state of the noiseless protocol on the $ABC$ registers, and is captured by the following definition:

**Definition 1** A simulation protocol $Q$ in the quantum model of length $N'$ succeeds with error $\varepsilon$ at simulating all length $N$ noiseless protocols against all adversaries in some class $A^{N'}$ if,
for all noiseless protocols \( \Pi \) of length \( N \), for all adversaries \( A^Q \in \mathcal{A}^N \), \( \| \Pi - Q^\Pi(A^Q) \| \leq \varepsilon \). The communication rate \( R_Q \) of \( Q \) is \( R_Q = \frac{N}{N^l \log q} \) for \( q \geq 2 \) the alphabet size of the communication register \( C' \).

In a random error model (analogous to that studied in quantum information theory, à la Shannon), Eve is a non-malicious passive environment, and \( \mathcal{N}_i = \mathcal{N}^Q \) for some fixed quantum channel \( \mathcal{N}^Q \), and the class \( \mathcal{A}^N \) contains a single element \((\mathcal{N}^{C'} \otimes I \otimes E') \otimes \mathcal{N}^N \). For simplicity, we then say that the simulation succeeds over \( \mathcal{N}^Q \). In an adversarial error model (analogous to that studied in quantum coding theory, à la Hamming), Eve is a malicious adversary who wants to make the protocol fail, and we are interested in particular classes of adversaries which we denote \( \mathcal{A}_0^Q \) for some parameter \( 0 \leq \delta < 1 \). The class \( \mathcal{A}_0^Q \) contains all adversaries with a bound \( \delta \) on the fraction of communications of the \( C' \) register they corrupt, in the following sense: for any reference register \( R \) and classical register \( X \), for any state \( \rho \in \mathcal{D}(Z^\otimes \mathcal{N}^N \otimes C'^\otimes \mathcal{N}^N \otimes R \otimes X) \) and for all possible classical states \( z \in Z \) of the classical register \( Z^\otimes \mathcal{N}^N \), if the action of some adversary in \( \mathcal{A}_0^Q \) on \( \rho \) is

\[
(\mathcal{N}_1^{E'C'} \otimes \cdots \otimes \mathcal{N}_N^{E'C'})(\rho) = \sum_{x \in X, z \in Z} p_{XZ}(x, z)|x \rangle \langle x|^X \otimes |z \rangle \langle z|^Z_{\otimes \mathcal{N}^N} \otimes (\mathcal{N}_1(z)^{E'C'} \otimes \cdots \otimes \mathcal{N}_N(z)^{E'C'})(\rho(x, z)), \quad (3.3)
\]

for some channels \( \mathcal{N}_i(z) \), quantum states \( \rho(x, z) \), and a probability density function \( p_{XZ} \), then we must have that the size of \{ \( i : \mathcal{N}_i(z)^{E'C'} \neq \mathcal{N}_i(z)^{E'} \otimes IC' \) \} is bounded by \( \delta N \).

Note that this allows for adaptive, probabilistic, entangled strategies for Eve, but such that conditioned on any sequence of measurement outcomes \( z \) (recorded in the \( Z \) registers), at most a \( \delta \) fraction of the the actions of Eve act non-trivially on the \( C' \) register, and so we say that the fraction of error is bounded by \( \delta \) for all adversary in \( \mathcal{A}_0^Q \).

For the shared entanglement model, Alice, Bob and Eve possess local classical-quantum registers split analogously to those in the quantum model. In addition to the entanglement inherent in \( |\psi_{\text{init}} \rangle^{ABCE} \), Alice and Bob also share entanglement to be consumed during the simulation in the form of a large state \( |\phi \rangle^{TA} \) with the registers \( T_A, T_B \) held by Alice and Bob, respectively. In general, the entanglement registers have a product decomposition \( T_A = T_A^1 \otimes \cdots \otimes T_A^{N} \), \( T_B = T_B^1 \otimes \cdots \otimes T_B^{N} \). A classical communication register \( C'' \), of some fixed size \( q \) independent of the length \( N \) of the protocol to be simulated, is exchanged back-and-forth between Alice and Bob by passing through Eve’s hand; it is held by Alice at both the beginning and the end of the simulation protocol. A simulation protocol \( \mathcal{S} \) in the shared entanglement model is defined by a sequence of quantum instruments \( \mathcal{M}_{1}^{XATAC''}, \mathcal{M}_{2}^{YBTBC''}, \ldots, \mathcal{M}_{N+1}^{XATAC''} \) such that, on input a state \( |\psi_{\text{init}} \rangle^{A'B'C'D'} = |\psi_{\text{init}} \rangle^{ABCA'E} \otimes |0 \rangle \), given black-box access to a noiseless protocol \( \Pi \), and against an adversary \( A^S_0 \) defined by a sequence of quantum instruments \( N_1^{ZEC''}, \ldots, N_N^{ZEC''} \), the protocol outputs the \( \tilde{A}BC \) subsystems of

\[
\rho_{\text{final}} = \mathcal{M}_{N+1}^{\Pi} N_N \mathcal{M}_N^{\Pi} \cdots \mathcal{M}_2^{\Pi} N_1 \mathcal{M}_1^{\Pi} (|\psi'_{\text{init}} \rangle \langle \psi'_{\text{init}} |). \quad (3.4)
\]

We denote this output by \( S^\Pi(A^S(\langle \psi_{\text{init}} \rangle)) \), and the induced quantum channel from \( A^SCE \) to \( \tilde{A}BC \cong ABC \) by \( S^\Pi(A^S) \). The success of the simulation is measured by how close the
simulation output state is to the final state of the noiseless protocol on the $ABC$ registers, and is captured by the following definition:

**Definition 2** A simulation protocol $S$ in the shared entanglement model of length $N'$ succeeds with error $\varepsilon$ at simulating all length $N$ noiseless protocols against all adversaries in some class $A^{N'}$ if, for all noiseless protocols $\Pi$ of length $N$, for all adversaries $A^S \in A^{N'}$, $\|\Pi - S^\Pi(A^S)\| \leq \varepsilon$. The communication rate $R_C$ of $S$ is $R_C = \frac{N}{N' \log q}$ for $q \geq 2$ the alphabet size of the classical communication register $C''$, and the entanglement consumption rate $R_E$ is $R_E = \frac{\log \left( \min (\dim T_A, \dim T_B) \right)}{N' \log q}$ for $T_A, T_B$ the entanglement registers used for the simulation by Alice and Bob, respectively.

In a random error model, Eve is a non-malicious passive environment, and $N_i = N^S$ for some fixed classical channel $N^S$, and the class $A^{N'}$ contains a single element $(N^C'' \otimes I^Z \otimes E^r)^{\otimes N''}$. For simplicity, we then say that the simulation succeeds over $N^S$. In an adversarial error model, Eve is a malicious adversary who wants to make the protocol fail, and we are interested in particular classes of adversaries which we denote $A^{S}_\delta$ for some parameter $0 \leq \delta < 1$. The class $A^{S}_\delta$ contains all adversaries with a bound $\delta$ on the fraction of communications of the $C''$ classical register they corrupt, in the following sense: for any reference register $R$ and classical register $X$, for any state $\rho \in D(Z^{\otimes N'} \otimes E^r \otimes N^r \otimes C'' \otimes N' \otimes R \otimes X)$ and for all possible classical states $z \in Z$ of the classical register $Z^{\otimes N'}$, if the action of some adversary in $A^{S}_\delta$ on $\rho$ is

$$\sum_{x \in X, z \in Z} p_{xz}(x, z) [x]_x [z]_z \otimes (N_i(z)^{E''} \otimes \cdots \otimes N_i(z)^{E''})(\rho(x, z)),$$

for some channels $N_i(z)$, quantum states $\rho(x, z)$, and a probability density function $p_{xz}$, we must have that the size of $\{ i : N_i(z)^{E''} \neq N_i'(z)^{E'} \otimes \Delta^{C''} \}$ is bounded by $\delta N'$, for $\Delta$ the noiseless classical channel (in the communication basis) on $C''$. Note that this allows for adaptive, probabilistic strategies for Eve, but such that conditioned on any sequence of measurement outcome $z$ (recorded in the $Z$ registers), at most a $\delta$ fraction of the the actions of Eve act non-trivially on the $C''$ register, even though she can copy all classical transmission in the $Z$ registers, and so we say that the fraction of error is bounded by $\delta$ for all adversary in $A^{Q}_\delta$.

Note that the adversaries in the quantum and in the shared entanglement models are fundamentally different: in the shared entanglement model, Eve can copy all classical messages and gather the corresponding information to establish her strategy, but she cannot modify Alice or Bob’s quantum information, except for what is possible by corrupting their classical communication and by using the information in the quantum register $E$ purifying the input state. By contrast, in the quantum model, she cannot always “read” the quantum messages, but she can apply entangled, fully quantum corruptions to the quantum register when she chooses to.
3.4 Classical Communication Protocols

Our simulation protocols contain an important classical component. In our setting, we are interested in protocols in which each party sends a message from some message set \([d] = \{1, 2, \ldots, d-1, d\}\) of size \(d\) in alternation, for some fixed number of rounds \(N'\) (actually, \(N'/2\) in our protocols). A round consists of Alice sending a message to Bob and then Bob sending a message back. Parties only have access to some noisy channels, so they need to encode these messages in some way. The codes used to do so in an interactive setting are described in the next subsection. For the moment, let us focus on the actual messages the parties wish to transmit.

In round \(i\), Alice transmits a message \(a_i \in [d]\) to Bob, and then Bob sends back a message \(b_i \in [d]\). These messages depend on the previous messages \(a_1, a_2, \ldots, a_{i-1} \in [d]\) and \(b_1, b_2, \ldots, b_{i-1} \in [d]\) Alice and Bob have sent in the previous rounds, respectively. Following [38], we refer to these sequences of messages (at the end of round \(i\)) as Alice’s state \(s_A = a_1 \cdots a_i \in [d]^i\) and Bob’s state \(s_B = b_1 \cdots b_i \in [d]^i\), respectively. Note that these states are updated in each round, and that each state, at the end of round \(i\), can be represented as a node at depth \(i\) in some \(d\)-ary tree of depth \(N'\). This tree is called a state tree. The whole (noiseless) communication can be extracted from the information in these two states.

Since the communication is noisy, in some rounds the parties make errors when trying to guess the other party’s state. When comparing the actual state \(s = s_1 \cdots s_i \in [d]^i\) of a party in round \(i\) with the other party’s best guess \(s^i = s_1^i \cdots s_i^i \in [d]^i\) about that state based on the communication he received up to that point, the least common ancestor of \(s\) and \(s^i\) is the node at depth \(i - \ell\) such that \(s_1 \cdots s_{i-\ell} = s_1^i \cdots s_{i-\ell}^i\) but \(s_{i-\ell+1} \neq s_{i-\ell+1}^i\). We call \(\ell\) the magnitude of the error of such a guess \(s^i\), and in general for two states \(s, s^i \in [d]^i\) satisfying the above (with least common ancestor at depth \(i - \ell\)) we write \(L(s, s^i) = \ell\). Note that we can compute \(\ell\) as \(i - \max \{t : (\forall j \leq t)[s_j = s^i_j]\}\).

3.5 Online Classical Codes

3.5.1 Tree Codes [38]

Standard error correcting codes are designed for data transmission and therefore are not particularly well suited for interactive communication over noisy channels. In his breakthrough papers on interactive communication [37, 38], Schulman defined tree codes, which are particular codes designed for such interactive communication. Indeed, these tree codes can perform encoding and decoding by rounds (following [21], we refer to such codes as online codes), such that for each round, a message from the message set \([d]\) is transmitted, but even if there is some decoding error in this round, for each additional round we perform (without transmission error), the more likely it is that this previous decoding error is correctly decoded. We describe this property in more details after formally defining these tree codes. We use the following for our definition. Given a set \(A\) and its \(k\)-fold cartesian product \(A^k = A \times \cdots \times A\) (\(k\)-times), we denote, for any \(n \in \mathbb{N}\), \(A^\leq n = \cup_{k=1}^{n} A^k\). Also, given a transmission alphabet \(\Sigma\) and two words \(\overline{e} = e_1 \cdots e_t \in \Sigma^t\) and \(\overline{e'} = e'_1 \cdots e'_t \in \Sigma^t\) over this
alphabet, we denote by $\Delta(\bar{e}, \bar{e}')$ (the Hamming distance) the number of different symbols, i.e. $\Delta(\bar{e}, \bar{e}') = |\{i : e_i \neq e_i'\}|$.

**Definition 3** (Tree codes [38]) Given a message set $[d]$ of size $d > 1$, a number of rounds of communication $N' \in \mathbb{N}$, a distance parameter $0 < \alpha < 1$ and a transmission alphabet $\Sigma$ of size $|\Sigma| > d$, a $d$-ary tree code of depth $N'$ and distance parameter $\alpha$ over alphabet $\Sigma$ is defined by its encoding function $E : [d]^{\leq N'} \rightarrow \Sigma$. It must also satisfy the following distance property, called the tree code property, in which we define $\bar{e} = e_1 \cdots e_t = \bar{E}(a), \bar{e}' = e_1' \cdots e_t' = \bar{E}(a')$:

$$(\forall t \leq N') (\forall a, a' \in [d]^t)[L(a, a') = \ell \rightarrow \Delta(\bar{e}, \bar{e}') \geq \alpha \cdot \ell],$$

in which, given an encoding function $E$, we also define its extension $\bar{E} : [d]^{\leq N'} \rightarrow \Sigma^{N'}$ satisfying

$$(\forall t \leq N') (\forall a = a_1 \cdots a_t \in [d]^t)[\bar{E}(a) = E(a_1)E(a_1a_2) \cdots E(a_1 \cdots a_{t-1})E(a_1 \cdots a_t) \in \Sigma^t].$$

The decoding function $D : \Sigma^{\leq N'} \rightarrow [d]^{\leq N'}$ satisfies

$$(\forall t \leq N') (\forall \bar{e}' \in \Sigma^t)[D(\bar{e}') \in \{a : a \in [d]^t \text{ minimizes } \Delta(\bar{E}(a), \bar{e}')\}].$$

Note that the decoding function is not uniquely defined for a given tree code: we could avoid ambiguity by outputting a special failure symbol for $D(\bar{e}')$ whenever $|\{a : a \in [d]^t \text{ minimizes } \Delta(\bar{E}(a), \bar{e}')\}| > 1$. Also note that we can view tree codes in the following alternate way, connecting them with the state tree representation defined above. Starting with a state tree, we can label the arcs out of each node by a symbol from $\Sigma$ corresponding to the encoding of that path in the tree code. The $\bar{E}$ encoding function represents the concatenation of the symbols on the path from root to node $a$, and the distance property is related to the distance of $a, a'$ to their least common ancestor in the protocol tree, and to the number of errors during these corresponding $L(a, a')$ last transmissions. The following was proved in [38] about the existence of tree codes:

**Lemma 1** Given a message set $[d]$ of size $d > 1$, a number of rounds of communication $N' \in \mathbb{N}$ and a distance parameter $0 < \alpha < 1$, taking transmission alphabet $\Sigma$ with $|\Sigma| = 2\lfloor 2 \cdot 2^{H(\alpha)} \cdot d \rfloor - 1$ suffices to label the arcs of some tree code, i.e. there exist an encoding function $E$ satisfying the tree code property, and the required alphabet size is independent of $N'$, the number of rounds of communication. Here, $H(\alpha) = -\alpha \cdot \log \alpha - (1 - \alpha) \cdot \log (1 - \alpha)$ is the binary entropy function.

In fact, the result of Schulman is even stronger: there exists an unbounded depth tree code with $\Sigma$ of the size discussed above. This stronger result could be useful in the case in which the number of rounds $N'$ is not bounded at the beginning of the protocol, and has been used to authenticate streams of classical data in [21].

The distance property of tree codes assures us of the following: if in round $t$ the decoding is good for the first $t - \ell$ messages sent ($\ell \geq 0$), but wrong for the message sent in round
the last \( t - \ell + 1 \) (and possibly also for some other messages), then the reencoding of the sequence of decoded messages must be distinct from the transmitted one in at least \( \alpha \cdot \ell \) positions in the last \( \ell \) rounds. Then, bad decoding implies that there must have been at least \( \frac{1}{2} \cdot \alpha \cdot \ell \) transmission errors during those rounds, independently of what was sent in the first \( t - \ell \) rounds. More precisely, given a transmitted message \( \bar{a} \in [d]^t \), encoded as \( \bar{e} = E(\bar{a}) \in \Sigma^t \), received as \( \bar{e}' \in \Sigma^t \), and decoded as \( \bar{a}' = D(\bar{e}') \in [d]^t \), with \( \bar{e}' = E(\bar{a}') \), if we have \( a_1 \cdots a_{t-\ell} = a'_1 \cdots a'_{t-\ell} \) but \( a_{t-\ell+1} \neq a'_{t-\ell+1} \), i.e. \( L(a, a') = \ell \), then \( \Delta(\bar{e}, \bar{e}') \geq \alpha \cdot \ell \) and \( \Delta(e_{t-\ell+1} \cdots e_t, e''_{t-\ell+1} e''_t) \geq \frac{1}{2} \cdot \alpha \cdot \ell \) (Note \( e_1 \cdots e_{t-\ell} = e'_1 \cdots e'_{t-\ell} \)). This property is the one so useful for interactive communication: even if bad decoding of a message is performed in some round, with enough correct transmissions from further rounds, we can later correct that previous error. This property is essential to our analysis of the simulation protocol.

3.5.2 Blueberry Codes [21]

Another kind of online codes we need to withstand the highest possible error rates are randomized error detection codes called blueberry codes in [21]. To use these, Alice and Bob encode and decode messages with a shared secret key in a way that weakly authenticates and encrypts each message, and in this way the adversary Eve cannot apply a corruption of her choosing. Such codes unknown to the adversary were termed private codes in [31]. At best, with some small (but constant) probability she is able to corrupt a message in such a way that Alice and Bob do not detect it and this results in an effective decoding error, but most of the time a corruption of Eve results in an effective erasure decoding. Since the tree code property, and hence also its decoding, is defined in terms of Hamming distance, actual errors are twice as harmful as erasures in the tree decoding (we can view the erasure flag \( \perp \) as a special symbol in \( \Sigma \) never used in the encoding, but which helps in decoding). Moreover, in rounds in which actual decoding errors arise, parties are not immediately aware of it and on the basis of this wrong information might perform wrong operations on the quantum registers that need to later be corrected, while a party immediately realizes when an erasure happens and this prevents him from performing such wrong operations. Hence, concatenating a blueberry encoding on the tree encoding enables significant improvement in the allowed error rates.

These blueberry codes were defined in [21] for the purpose of authenticating streams of classical messages and for the simulation of interactive classical protocols. The authors gave the following definition for them and proved the following properties:

**Definition 4** (Blueberry codes) For \( i \geq 1 \) let \( B_i : \Gamma \to \Gamma \) be a random and independent permutation. The blueberry code maps a string \( e \in \Sigma^t \subset \Gamma^t \) of arbitrary length \( t \) to \( B(e) = B_1(e_1)B_2(e_2) \cdots B_t(e_t) \). We denote such a code as \( B : \Sigma^* \to \Gamma^* \), and define the erasure parameter of this code as \( \beta = 1 - \frac{|\Sigma| - 1}{|\Gamma| - 1} \), and its complement \( \varepsilon_{\beta} = 1 - \beta = \frac{|\Sigma| - 1}{|\Gamma| - 1} \).

**Definition 5** Assume that at some time \( i \), \( d_i = B_i(e_i) \) is transmitted and \( d'_i \neq d_i \) is received. If \( B^{-1}_i(d'_i) \notin \Sigma \), we mark the transmission as an erasure, and the decoding algorithm outputs \( \perp \). Otherwise, this event is called an error.
Corollary 1 Let \( e \in \Sigma^t \) and assume \( B(e) \) is communicated over a noisy channel. Every symbol corrupted by the channel causes either an error with probability \( \varepsilon_\beta \), or an erasure with probability \( \beta \).

Lemma 2 Assume a blueberry code \( B : \Sigma^* \rightarrow \Gamma^* \) is used to transmit a string \( e \in \Sigma^t \) over a noisy channel. For any constant \( 0 \leq c \leq 1 \), if the channel’s corruption rate is \( c \), then with probability \( 1 - 2^{-\Omega(t)} \) at least a \( c(1 - 2\varepsilon_\beta) \)-fraction of the transmissions are marked as erasures.

Corollary 2 If out of \( t \) received transmissions, \( ct \) were marked as erasures by a blueberry code \( B : \Sigma^* \rightarrow \Gamma^* \), then except with probability \( 2^{-\Omega(t)} \) over the shared randomness, the adversarial corruption rate is at most \( c/(1 - 2\varepsilon_\beta) \).

4 Basic Simulation Protocol

We start by describing a basic simulation protocol, which attains the first goal of simulating quantum protocols with asymptotically positive communication and error rates, and constant entanglement consumption rate. This provides an interactive analogue of a family of good quantum codes.

4.1 Result

We focus on the shared entanglement model since techniques to distribute entanglement in both random and adversarial error models are well-studied, so we can combine our findings with these entanglement distribution techniques to translate these results in the quantum model. Also, we focus on an adversarial model of error, and then we can adapt these results to a random error model. Such extensions of this result to other models of communication are studied in section 6. For the basic simulation protocol described in this section, entanglement is only used to teleport the quantum information back-and-forth between the two parties. In section 5 we show how to tolerate maximum error rates by also using entanglement to generate a shared secret key unknown to the adversary, thus enabling the two honest parties to detect most adversarial errors as effective erasures.

Theorem 1 Given an adversarial channel in the shared entanglement model with low enough error rate, we can simulate perfectly any noiseless protocol of length \( N \) over this channel using a number of transmission linear in \( N \), and consuming a linear number of EPR pairs. More precisely, there exists constant error rate \( \delta > 0 \), communication rate \( R_C > 0 \), transmission alphabet size \( q \in \mathbb{N} \), and entanglement consumption rate \( R_E \in \mathbb{R}^+ \) such that for all noiseless protocol lengths \( N \in 2\mathbb{N} \), there exist a universal simulator \( S \) in the shared entanglement model of length \( N' \) with communication rate at least \( R_C \), transmission alphabet size \( q \), entanglement consumption rate \( R_E \), which succeeds with zero error at simulating all noiseless protocols of length \( N \) against all adversaries in \( \mathcal{A}_\delta \).
4.2 Intuition for the simulation protocol

Before describing in detail the basic simulation protocol, let us first give some intuition on how it succeeds in simulating a noiseless quantum protocol over a noisy channel. The strategy to avoid losing the quantum information in the communication register over the noisy channel is to teleport the $C$ register of the noiseless protocol back and forth into Alice’s $C_A$ register and Bob’s $C_B$ register, creating a virtual $C$ register which is either in Alice’s or in Bob’s hand. They use the shared entanglement in $T_A T_B$ to do so, as well as the provided noisy classical channels to transmit their teleportation measurement outcome. Whenever Alice possesses the virtual $C$ register she can try to evolve the simulation of the noiseless protocol by applying one of her noiseless protocol unitaries on the virtual $AC$ register, and similarly for Bob on the virtual $BC$ register. If it is later realized that there has been some error in the teleportation decoding, they might have to apply inverses of these operations, but overall, everything acting on the virtual $ABC$ quantum register can be described as an intertwined sequence of Pauli operators acting on the $C$ register and noiseless protocol unitaries (and their inverses) acting on the $AC$ and the $BC$ registers. There are two important things to notice here. First, the sequence of operations acting on the joint register is a sequence of reversible unitaries acting either only on the $C$ register (for the Pauli operators appearing during teleportation) or on pairs $AC$ or $BC$ of registers (for noiseless protocol unitaries and their inverses). Hence, if the parties can keep track of the sequence of operations on the joint register, at least one of the parties can reverse any of the operations when he is in possession of the virtual $C$ register. Second, both parties know the order in which these operators have been applied while only one knows exactly which one was applied: for Pauli operators, both parties know $\pm X^x Z^z$ is applied at some point, but only one knows for sure the value of $xz \in \{0, 1\}^2$, and similarly both know $U_j^M$ (with $U_j^{+1} = U_j, U_j^{-1} = U_j^\dagger, U_j^0 = I$) is applied at some point, but only one knows for sure the values of $j \in \{1, \cdots, N^j + 1\}$ and $M \in \{-1, 0, +1\}$. This is the classical information they try to transmit each other so that both can know exactly the sequence of operations that have acted on the joint register up to some point. The tree codes of Schulman are particularly well suited for noisy communication in this interactive scenario.

More concretely, in each round the parties first need to decode the teleportation before trying to evolve the simulation of the quantum protocol and finally teleporting back the communication register to the other party. We want the parties to be able to know exactly where they are in the simulation of the protocol when they are able to correctly decode the classical messages sent by the other party up to that point. To enable a party to learn exactly what actions were taken by the other party in each previous round, the message set in each round is $\{0, 1\}^2 \times \{-1, 0, +1\} \times \{0, 1\}^2$, and messages are encoded with a tree code before being sent. The first pair of bits corresponds to the teleportation decoding operation done at the beginning of a party’s turn. Then the trit is associated to the evolution in the noiseless protocol: $+1$ stands for going forward with the protocol, a unitary of the noiseless protocol was applied to the joint state of the party’s local register and the communication register; $-1$ stands for going back with the protocol, the inverse of an unitary of the noiseless protocol applied by that party to the joint state was performed; $0$ stands for holding the protocol idle,
no action is done by that party to evolve the protocol in that round. Note that the index \( j \) of the unitary \( U^M_j \) a party applies can be computed solely from the sequence of trits sent by that party, and such an explicit calculation is defined in the simulation description. Finally, the last pair of bits corresponds to the outcome of the measurement in the teleportation of the communication register, to enable the other party to correctly decode the teleportation.

For each party, we call his state at some point the sequence of these triplets of messages he transmitted up to that point (see section 3.4). If a party succeeds in correctly decoding the state of the other party, he then possesses all the information about what operations were applied on the joint quantum register, and can choose his next move accordingly. Note that the information about which Pauli operator was used to decode the teleportation might appear to be redundant, but it is not when there are decoding errors. In such a case, the wrong Pauli operators might be applied to do the teleportation decoding. Even though the party who applied the wrong Pauli operator will realize his mistake later when the tree code enables him to finally decode this message correctly, the other party still need to be informed that the decoding of the teleportation in that particular round was different from what it should have been based on the initial teleportation measurement outcome. Sending the information about which Pauli operator was used to do the teleportation decoding implicitly provides that information, and even enables the other party to correct this wrong teleportation decoding by himself if needs be. We indeed use this property in the simulation.

4.3 Description of the simulator

All communication is done with a tree encoding over some alphabet \( \Sigma \). To later simplify the analysis, we fix the distance parameter to \( \alpha = \frac{39}{40} \). The message set consist of \( \{0, 1\}^2 \times \{-1, 0, +1\} \times \{0, 1\}^2 \cong [4] \times [3] \times [4] \cong [48] \), so we take arity \( d = 48 \). Also, taking \( N' = 4(1 + \frac{1}{N})N \) is sufficient. By Lemma 1, we know that there exist a \( q \in \mathbb{N} \) independent of \( N' \) such that an alphabet \( \Sigma \) of size \( q \) suffices to label the arcs of a tree code of any depth \( N' \in \mathbb{N} \). Both parties have already agreed before the protocol begins on such a tree code of depth \( N' \) with corresponding encoding and decoding functions \( E \) and \( D \) (both parties use a different instance of the same tree code to transmit their messages to the other party). Also, we want to tolerate error rate \( \delta = \frac{1}{80} \).

The convention we use for the variables of the protocol is the following: on Alice’s side, in round \( i \), \( x_{iAD}z_{iAD} \in \{0, 1\}^2 \) correspond to the bits she uses for the teleportation decoding on the \( X \) and \( Z \) Pauli operators, respectively, \( x_{iAM}z_{iAM} \in \{0, 1\}^2 \) correspond to the bits of the teleportation measurement on the corresponding Pauli operators, \( j_{iA} \in \mathbb{Z} \) and \( M_{iA} \in \{-1, 0, +1\} \) correspond respectively to the index of the unitary she is using in round \( i \) and to whether she is using \( U_{j_{iA}} \), its inverse \( U^{-1}_{j_{iA}} = U^\dagger_{j_{iA}} \), or simply acting the identity channel \( U^0_{j_{iA}} = I \) on the quantum register, and the counter \( C_{iA} \) keeps track of the sum of all previous messages \( M_{lA}, l \leq i \). Similarly, on Bob’s side, all the same variables are used, with A’s changed to B’s. When discussing variables obtained from decoding in round \( i \), a superscript \( i \) is added to account for the fact that this decoding might be wrong and could be corrected in later rounds, and similarly when discussing other variables which are round dependent.

19
The actions Alice and Bob take in round $i$ is based on the following two representations for the form of the state $|\psi_i\rangle$ of the joint register at the beginning of round $i$ ($|\psi_1\rangle = |\psi_{\text{init}}\rangle$) which can be classically computed from the information in their two state trees. The first one can be directly computed as

$$|\psi_i\rangle^{ABCE} = \prod_{\ell=1}^{i-1} (X^{i+1}Z^{i+1}U^{M_{iB}} Z^{i+1} X^{i+1} X^{i+1} Z^{i+1} U^{M_{iA}} Z^{i+1} X^{i+1}) |\psi_{\text{init}}\rangle^{ABCE}. \tag{4.1}$$

Here, from the state $s_A$ of Alice’s state tree, we can directly obtain from the $\ell$th message sent by Alice, for $\ell = 1 \cdots i - 1$, the two bits $x_{iA}z_{iA}$ used to decode the teleportation, the trit $M_{iA}$ corresponding to the evolution of the protocol performed in round $\ell$, and then the two bits $x_{iAM}z_{iAM}$ corresponding to the outcome of the teleportation measurement. We then use counters $C_{iA}$’s that maintain the sums of the $M_{iA}$’s to compute the indices $j_{iA}$’s of the noiseless protocol unitaries used by Alice in round $\ell$: $C_{0A} = 0, C_{iA} = C_{(\ell-1)A} + M_{iA}, j_{iA} = 2C_{(\ell-1)A} + M_{iA}$. Note that $j_{iA}$ depends only on the sequence of messages $M_1, M_2, \cdots, M_{(i-1)A}, M_i$. Similarly, the state $s_B$ of Bob’s state tree is used to obtain $x_{iBD}z_{iBD}, x_{iBM} z_{iBM}$, as well as $M_{iB}$, and to compute $C_{0B} = 0, C_{iB} = C_{(\ell-1)B} + M_{iB}, j_{iB} = 2C_{(\ell-1)B} + M_{iB} + 1$. We define $U_j^M = I$ whenever $j \leq 0$ or $M = 0$. Note that if $M \neq 0$ then $j_{iA}$ is odd and $U_j^{M_{iA}}$ acts on Alice’s side, $j_{iB}$ is even and then $U_j^{M_{iB}}$ acts on Bob’s side. Also note that $j \leq N^i + 1$ so the $U_j$’s are well defined by the noiseless protocol embedding described in section 3.3.1.

From this first representation of the form of the state $|\psi_i\rangle$, we can classically compute a second one by recursively cleaning up the first representation. The cleanup is performed by collapsing together as many of the operators as possible (Pauli operators together, $U_i$’s with $U_{i-1}$’s) to obtain something in the form:

$$|\psi_i\rangle^{ABCE} = \hat{\sigma}_i \cdot \hat{U}_{iA} \cdot \hat{\sigma}_{i-1} \cdot \hat{U}_{i-1} \cdots \hat{U}_1 \cdot \hat{\sigma}_1 \cdot U_{r_i} \cdot U_{r_{i-1}} \cdots U_2 \cdot U_1 |\psi_{\text{init}}\rangle^{ABCE}. \tag{4.2}$$

with $\hat{\sigma}_i = \pm X^{x_i} Z^{z_i}, \hat{\sigma}_i = X^{x_i} Z^{z_i}$ for $x_i, z_i \in \{0, 1\}$, and $\hat{U}_i = U_i^{\pm 1}$ for some $r_i - 2t_i \leq \ell' \leq r_i + 2t_i$. The rules to be used recursively to perform the cleanup are the following: in the case that $\hat{\sigma}_i = I$, we require, if $\ell > 1$, that $\hat{U}_i \neq (\hat{U}_{i-1})^{-1}$, and if $\ell = 1$, that $\hat{U}_i \neq U_{r_{i+1}}$. This last rule is what determines the cut between $U_{r_i}$ and $\hat{U}_i \hat{\sigma}_i$. The parameter $r_i$ determines the number of noiseless protocol unitaries the parties have been able to successfully apply on the joint register before errors start to arise on it, and the parameter $t_i$ determines the number of errors the parties have to correct before being able to resume the simulation. Note that this is well-defined: there is a unique representation in the form (4.2) corresponding to any in the form (4.1).

To decide which action to take in round $i$, Alice starts by decoding the possibly corrupted messages $f^i_1, \cdots, f^i_{i-1} \in \Sigma$ received from Bob up to this point to obtain her best guess $s_B^i = D(f^i_1, \cdots, f^i_{i-1})$ for the state $s_B$ of his state tree. Along with the state $s_A$ of her state tree, she uses it to compute her best guess of the form (4.2) of the joint state. If her decoding
of Bob’s state is good, then she has all the information she needs to compute the form of the joint state $|\psi_i\rangle$. She can then choose the right actions to take to evolve the simulation. She takes the following actions based on the assumption that her decoding is good. If it is not, errors might accumulate on the joint register $ABC$, which she will later have to correct.

Alice’s next move depends on whether (she thinks) $t_i = 0$ or not. If $t_i = 0$, then she wishes to evolve the protocol one round further, if it is her turn to do so. That is, if $r_i$ is even, then she sets $M_{iA} = +1$ to apply $U_{r_i+1}^{AC}$, but if $r_i$ is odd, Bob should be the next to apply a unitary of the protocol, so she sets $M_{iA} = 0$. If $t_i \neq 0$, then she wishes to correct the last error not yet corrected, if she is the one who applied it. That is, if $\hat{U}_{t_i} = U_{\ell'}^M$ for $\ell'$ odd, then she sets $M_{iA} = -M' \in \{\pm 1\}$ (note that in this case it holds that she sets $\hat{j}_i = \ell'$), else she sets $M_{iA} = 0$ and she hopes Bob will next correct $\hat{U}_{t_i}$. In all cases, with $\hat{\sigma}_i^C = \pm X^{\hat{x}_i} Z^{\hat{z}_i}$, she sets $s_{iAD} = \hat{x}_i$, $z_{iAD} = \hat{z}_i$ and computes $C_{iA} = C_{(i-1)A} + M_{iA}$, $\hat{j}_i = 2C_{(i-1)A} + M_{iA}$. Note that she does not care about the irrelevant global phase factor $\pm 1$ appearing in $\hat{\sigma}_i$ during the cleanup from the form (4.1) to the form (4.2) because of the fact that the $X$ and $Z$ Pauli operators anticommute.

After this classical preprocessing, she can now perform her quantum operations on the $AC$ registers: she first decode the teleportation operation (and possibly some other Pauli errors remaining on the $C$ register) by applying $Z^{s_{iAD}} X^{x_{iAD}}$ on the $T_A^{2(i-1)}$ register (note that in round 1, Alice already possesses the $C$ register so this part is trivial: let $T_A^0 = C_A$ and set $x_{1AD} \cdot z_{1AD} = 00$) before swapping registers $T_A^{2(i-1)}$ and $C_A$, effectively putting the virtual $C$ register into $C_A$. She then performs $U_{j_i A}^{M_{iA}}$ on the virtual $AC$ register to try to evolve the protocol (or correct a previous error), before teleporting back the virtual $C$ register to Bob using the half of entangled state in the $T_{B}^{2i-1}$ register, obtaining measurement outcome $x_{iAM} z_{iAM} \in \{0,1\}^2$. She updates her state $s_A$ by following the edge $a_i = (x_{iAD} z_{iAD}, M_{iA}, x_{iAM} z_{iAM})$ in the state tree, and transmits message $e_i = E(a_1 \cdots a_i)$ over the noisy classical channel, with $E$ the encoding function of the tree code.

Upon reception of the message $e'_i$, a possibly corrupted version of $e_i$, Bob obtains his best guess $s'_A$ for Alice’s state $s_A$ by computing, with previous messages $e'_1 \cdots e'_{i-1}$, $s'_A = D(e'_1 \cdots e'_{i}).$ He uses it along with his own state $s_B$ to compute his best guess of the representation of
\[
(X^{x_{iAM}} Z^{z_{iAM}} U_{j_i A}^{M_{iA}} Z^{z_{iAD}} X^{x_{iAD}}) |\psi_i\rangle
\] (4.3) analogous to the form in (4.1), then clean this up to obtain a representation analogous to (4.2), and based on this latest representation chooses in the same way as Alice his $x_{iBD} z_{iBD}, M_{iB}$, and then uses $M_{iB}$ to compute $C_{iB}, j_{iB}$. After this classical preprocessing, he can then perform his quantum operations: he first decodes the teleportation operation by applying $Z^{s_{iBD}} X^{x_{iBD}}$ on the $T_{B}^{2i-1}$ register and by swapping it with $C_B$, creating a virtual $C$ register, then performs $U_{j_i B}^{M_{iB}}$ on the virtual $BC$ register to try to evolve the protocol, before teleporting back the virtual $C$ register to Alice using the half of entangled state in the $T_{B}^{2i}$ register, and obtains measurement outcome $x_{iBM} z_{iBM}$. He updates his state $s_B$ by following the edge $b_i = (x_{iBD} z_{iBD}, M_{iB}, x_{iBM} z_{iBM})$, and transmits message $f_i = E(b_1 \cdots b_i)$ over the channel. The round completes when Alice receives message $f'_i$, a possibly corrupted version
of \( f_i \). After the \( \frac{N'}{2} \) rounds, Alice and Bob take the particular registers \( \tilde{A}, \tilde{B} \) and \( \tilde{C} \) specified by the noiseless protocol embedding (see section 3.3.1), and use them as their respective outcome for the protocol. If the simulation is successful (and it is if the error rate is below \( \frac{1}{80} \)), the output quantum state corresponds to the \( ABC \) subsystem of \( |\psi_{\text{final}}\rangle^{ABCE} \) specified by the original noiseless protocol.

We summarize the protocol as follow: Alice and Bob repeat the following for \( i = 1 \cdots \frac{N'}{2} \):

1. Alice computes \( s_B^i = D(f_1^i \cdots f_{i-1}^i) \), and extracts \( b_\ell^i = (x_{iBD}^i z_{iBD}^i, M_{iB}^i, x_{iBM}^i z_{iBM}^i) \), \( \ell = 1 \cdots i - 1 \), her best guess for Bob’s messages, and the corresponding \( C_{iB}^\ell, j_{iB}^\ell \).

2. Also using \( s_A \), she computes her best guess for the form (4.2) of the state \( |\psi_i\rangle \) of the joint register, and the corresponding \( x_{iAD}^i z_{iAD}^i, M_{iA}^i, C_{iA}^i, j_{iA}^i \).

3. She decodes the teleportation by applying \( Z_{iAD} X_{iAD} \) to register \( T_{2(i-1)}^A \) and swaps this with the \( C_A \) register.

4. She tries to evolve the simulation by applying \( U_{j_{iA}}^{M_{iA}} \) to the \( AC_A \) register.

5. She teleports back the \( C_A \) register to Bob using entanglement in register \( T_{2i-1}^A \) and gets outcomes \( x_{iAM} z_{iAM} \).

6. Alice updates her state \( s_A \) by following edge \( a_i = (x_{iAD} z_{iAD}, M_{iA}, x_{iAM} z_{iAM}) \) and transmits message \( e_i = E(a_1 \cdots a_i) \) using the channel to Bob, who receives \( e_i' \), a possibly corrupted version of \( e_i \).

7. Bob computes \( s_A^i = D(e_1^i \cdots e_i^i) \) and also using \( s_B \), performs actions on his side analogous to Alice’s, first swapping register \( T_{2i-1}^B \) with \( C_B \), then using the \( T_{2i}^B \) register to teleport back the \( C_B \) register to Alice, transmits \( f_i \), and round \( i \) completes upon reception by Alice of \( f_i' \), a possibly corrupted version of \( f_i \).

After these \( \frac{N'}{2} \) rounds, they both extract their protocol outcome from the \( \tilde{A} \tilde{B} \tilde{C} \) registers specified by the noiseless protocol embedding.

### 4.4 Analysis

The analysis is done conditioned on some overall classical state (and in particular, some respective views of Alice and Bob of the transcript) at each round. In particular, if the adversary has an adaptive, probabilistic strategy, we condition on some strategy based on the outcome of her previous measurements. We come back later to this issue.

We define two kinds of rounds: good rounds in which both parties decode correctly the other party’s state, and bad rounds in which at least one party makes a decoding error. We define a quantity \( P(i) \in \mathbb{Z} \) which increases at least by some (strictly positive) amount in good rounds, and decreases by at most some other (bounded) amount in bad rounds, and such that \( P\left(\frac{N'}{2} + 1\right) \geq N + 1 \) implies success of the simulation. Hence, it is sufficient to
bound the ratio of good to bad rounds as a function of the error rate to prove the success of the simulation.

Let us now define $P(i)$ more formally. To do so, we use the representation (1.2) for the form of the quantum state of the joint registers at the beginning of round $i$ (or equivalently, at the end of round $i-1$). Remembering $r_i$ determines the number of noiseless protocol unitaries the parties have been able to successfully apply on the joint register before errors start to arise on it, and $t_i$ determines the number of errors the parties have to correct before being able to resume the simulation, we define

$$P(i) = r_i - 2t_i,$$  \hspace{1cm} (4.4)

in which the factor of 2 in front of $t_i$ is due to the fact that in the worst case all remaining $\hat{U}^i_j$'s are applied by the same party who applied $U_{r_i-1}$ and $\hat{U}^i_{t_i} = U_{r_i-1-2(t_i-1)}^{-1}$. We now prove the following technical lemma which bounds $P(i)$ as a function of the number of good and bad rounds.

**Lemma 3** At the end of round $i$, define

$$N^i_g = |\{j : j \leq i, \text{round } j \text{ was good}\}|,$$

$$N^i_b = |\{j : j \leq i, \text{round } j \text{ was bad}\}|.$$

Then $P(i + 1) \geq N^i_g - 4N^i_b$.

**Proof.** Let us adopt the following notation: for $m \in [\frac{N_i}{2}]$, $V_1 = U_1, V_2 = U_3, \ldots V_{m+1} = U_{2m+1}$, i.e. the $V_m$'s are the $U_i$'s acting on Alice’s side, and $W_1 = U_2, W_2 = U_4, \ldots W_m = U_{2m}$, i.e. the $W_m$'s are the $U_i$’s acting on Bob’s side. We can then observe the following three facts whenever $t_i \geq 1$ at the end of round $i$, given our way to compute $\ell_{iA}$ and $\ell_iB$ defined in the protocol description above. Their proofs follow by noting that a statement analogous to the statement still holds at each step of the recursive cleanup until arriving at form (1.2).

First, looking at the $\hat{U}$ acting on Alice’s side (if such a $\hat{U}$ exists), the first one, say $\hat{U}^i_{\ell_0}$ for some $1 \leq \ell_0 \leq t_i$, satisfies the following: $\hat{U}^i_{\ell_0} \in \{V_{\ell+1}, V_{\ell}^{-1}\}$ for $V_{\ell} = U_{r_{\ell}}$ or $U_{r_{\ell}-1}$ (whichever acts on Alice’s side). A similar statement holds for Bob with the $W_i$’s.

Second, for any two successive $\hat{U}$’s acting on Alice’s side (if two such $\hat{U}$’s exist), say $\hat{U}^i_{\ell_1}$ and $\hat{U}^i_{\ell_2}$ for some $\ell_1 < \ell_2$, if $\hat{U}^i_{\ell_1} = V_{\ell_1}^{M_1}$ for some $1 \leq \ell_1' \leq N'$ and $M_1 \in \{\pm 1\}$, then $\hat{U}^i_{\ell_2} = V_{\ell_2+1}^{M_2}$, for $M' = \frac{M_1+M_2}{2}$. A similar statement holds for Bob.

Third, the choice of $M_{iA}$ and $j_{iA}$ are good, i.e. in a good round in which Alice tries to correct the last $\hat{U}$ acting on her register, say $\hat{U}_{\ell_3}^i = V_{\ell_3}^{M_3}$, then $U_{j_{iA}} U_{M_3}^{(-M_3)}$ indeed. Note that the choice of $j_{iA}$ is also good when $t_i = 0$, and similar statements hold for Bob.

Using these facts, we can prove Lemma 3 by induction. For the base case, $|\psi_1\rangle = |\psi_{\text{init}}\rangle$, so $P(1) = 0$ and the statement holds trivially. To give us a flavor, let us look at $P(2)$ at the end of round 1. In round 1, Alice applies $U_1$ then teleport. Then on Bob’s side if there is no decoding error he applies $U_2$ and teleport back, leaving $\hat{d}U_2 U_1 |\psi_{\text{init}}\rangle$ on the
register, so \( P(2) = 2 \geq 1 = N_g^1 \) in this case \( (N_g^1 = 0) \), else there is a decoding error and at worst he badly decodes the teleportation and still applies \( U_2 \), leaving \( \hat{\sigma}U_2\hat{U}_1|\psi_{\text{init}}\rangle \) on the register and \( P(2) = 1 - 2 = -1 \geq -4 = -4N_b^1 \) in this case \( (N_g^1 = 0) \). For the induction step, given the state \( |\psi_{i}\rangle \) at the end of round \( i - 1 \), if the \( i \)th round is a good round \( (N_g^i = N_g^{i-1} + 1, N_b^i = N_b^{i-1}) \), then at least one of Alice or Bob can act on the joint register, and so, by the way they choose their actions and by the above argument, if \( t_i \geq 1 \), then \( t_{i+1} \leq t_i - 1 \), and \( r_{i+1} \geq r_i \), else \( t_{i+1} = t_i = 0 \) and \( r_{i+1} \geq r_i + 1 \), so in all cases

\[
P(i + 1) = r_{i+1} - 2t_{i+1} \\
\geq r_i - 2t_i + 1 \\
= P(i) + 1 \\
\geq N_g^{i-1} - 4N_b^{i-1} + 1 \\
= N_g^i - 4N_b^i.
\]

If it is a bad round \( (N_g^i = N_g^{i-1}, N_b^i = N_b^{i-1} + 1) \), then the worst that can happen is if both parties apply a wrong unitary and \( t_{i+1} = t_i + 2, r_{i+1} = r_i \) \((t_{i+1} = t_i + 1, r_{i+1} = r_i - 1 \) or \( t_{i+1} = t_i, r_{i+1} = r_i - 2 \) and others are also possible, but not as bad for \( P(i + 1) \)) so

\[
P(i + 1) = r_{i+1} - 2t_{i+1} \\
\geq r_i - 2t_i - 4 \\
= P(i) - 4 \\
\geq N_g^{i-1} - 4N_b^{i-1} - 4 \\
= N_g^i - 4N_b^i.
\]

In all cases, \( P(i + 1) \geq N_g^i - 4N_b^i \) which proves our claim. ■

**Corollary 3** If \( P(\frac{N_i^r}{2} + 1) \geq N + 1 \), then the simulation succeeds.

**Proof.** For notational convenience, in this proof let \( r = r_{\frac{N_i^r}{2} + 1}, t = t_{\frac{N_i^r}{2} + 1} \), and also let the superscript \( \frac{N_i^r}{2} + 1 \) be implicit in all the \( \hat{U} \)'s. The proof of Lemma 3 also establishes that for two successive indices acting on either Alice's or Bob's side, these do not differ by more than 2. Also, \( P(\frac{N_i^r}{2} + 1) = r - 2t \geq N + 1 \) implies \( r \geq N + 1 + 2t \) for \( i \geq 0 \), and in particular we have \( r \geq N + 1 \). Then we know that after \( U_r \), the noiseless protocol embedding has already put the \( ABC \) registers of the noiseless protocol into safe local registers \( \hat{A}, \hat{B}, \hat{C} \), which are never accessed by \( U_{N+2} \cdots U_{N'+1}, \) and neither by \( X^C, Z^C \). Is left to verify that all \( \hat{U} \)'s have indices strictly higher than \( N + 1 \). But the indices of the \( \hat{U} \)'s of Alice decrease by at most two at once, and similarly for Bob, so clearly the worst case is if all \( \hat{U} \)'s are for the same party, and are inverses of the noiseless protocol unitaries. Without loss of generality, we consider only this case. If it is the same party who applied \( U_r \) who applies all the \( \hat{U} \)'s, then \( \hat{U}_1 = U_{r-1}, \hat{U}_2 = U_{r-1}^{-1} \cdots \hat{U}_i = U_{r-2(t-1)}^{-1} \) and \( r - 2(t - 1) > r - 2t = P(\frac{N_i^r}{2} + 1) \geq N + 1 \), so we are good. Similarly if it is the same party who applied \( U_{r-1} \) who applies the \( \hat{U} \)'s, then
\[ \tilde{U}_1 = U_{r-1}, \tilde{U}_2 = U_{r-3}, \ldots \tilde{U}_t = U_{r-2t+1} \text{ and } r - 2t + 1 > r - 2t = P(\frac{N'}{2} + 1) \geq N + 1. \]

In all cases, we are good, and the safe registers \(\tilde{ABC}\) to be outputted by the parties hold the \(ABC\) subsystem of \(|\psi_{\text{final}}\rangle\) at the end of round \(\frac{N'}{2}\) whenever \(P(\frac{N'}{2} + 1) \geq N + 1. \) ■

We now want to show that if the number of errors as a fraction of \(N'\) (the total number of classical symbols transmitted over the adversarial channel) is bounded by a particular constant \(\delta > 0\), we are then sure that the simulation succeeds. We do this in two steps: we first give a bound on the fraction of bad rounds as a function of the error rate, and then use it to show that below a certain error rate, the simulation succeeds.

The bound on the fraction of bad rounds as a function of the error rate we use follows as a corollary from the more general result in Lemma 5, which we prove in the next section when studying the way to tolerate the highest possible error rates. The result we use here is the following: if the error rate is bounded by \(\delta\) (so there are at most \(\delta N'\) errors) and the tree code distance of both Alice and Bob’s tree code is at least \(\alpha = 1 - \varepsilon\alpha\), then the number of bad round \(N_b\) is bounded by \(N_b \leq 2\delta N' + \varepsilon\alpha N' = (2\delta + \varepsilon\alpha)N'.\) Note that since we use a standard tree code without an erasure symbol, we could also obtain results with a weaker dependence on \(\varepsilon\alpha\) to improve on the (binary) communication rates.

We are now ready to prove that the simulation succeeds with the parameters of our protocol. We have \(\varepsilon\alpha = \frac{1}{40}, \delta = \frac{1}{80}, l_r = \frac{N'}{N} = 4(1 + \frac{1}{N}),\) so

\[
P(\frac{N'}{2} + 1) \geq N_g - 4N_b
\]

\[
= \frac{N'}{2} - 5N_b
\]

\[
\geq \frac{N'}{2} - 5(2\delta + \varepsilon\alpha)N'
\]

\[
= N'(\frac{1}{2} - \frac{10}{80} - \frac{5}{40})
\]

\[
= \frac{1}{4}N'
\]

\[
= N + 1,
\]

in which the first inequality is from Lemma 5, the first equality is by definition of \(N_g, N_b,\) i.e. \(\frac{N'}{2} = N_g + N_b,\) and the second inequality is from our bound on \(N_b\) due to Lemma 5. The fact that the simulation succeeds is then immediate from Corollary 3.

Note that the simulation protocol does not depend on the particular protocol to be simulated, but only on its length \(N\) and the noise parameter of the adversarial channel we want to tolerate. Also note that even if the adversary is adaptive and probabilistic (with adaptive, random choices depending on her measurement outcomes, as allowed by the model), the simulation succeeds no matter her choice of action, as long as the corruption rate is bounded by \(\delta,\) since then in each branch of the probabilistic computation our analysis holds. We use the definition of the class \(A^S_\delta\) to prove that indeed, the simulation succeeds with zero error.

For \(|\psi\rangle \in \mathcal{H}(A \otimes B \otimes C \otimes E \otimes R),\) with \(R\) a purifying system of the same size as
A ⊗ B ⊗ C ⊗ E, with have that
\((Π ⊗ I^R)(|ψ⟩) = \text{Tr}_E(U_N \cdots U_1|ψ⟩⟨ψ|U_1^T \cdots U_N^T),\)
while for \(A^s ∈ A^s_\delta,\)
\((S^H(A^s) ⊗ I^R)(|ψ⟩) = \text{Tr}_{(ABCR)}(M^{II}_{N'+1}N'_N M^{II}_N \cdots M^{II}_2 N_1 M^{II}_1 (|ψ⟩⟨ψ|)),\)
in which the \(-{(ABCR)}\) subscript argument for the partial trace means that we trace everything except the \(ABCR\) registers. But then we can rewrite
\((S^H(A^s) ⊗ I^R)(|ψ⟩) = \sum_{xTyTz} p_{xTyTz}(x_T, y_T, z)|x_T⟩⟨x_T|^{xT} ⊗ |y_T⟩⟨y_T|^{yT} ⊗ |z⟩⟨z|^{zT} ⊗ \rho(x_T, y_T, z)\)
for \(X_T, Y_T\) the registers containing the views \(x_T, y_T\) of the transcript as seen by Alice and Bob, respectively, for some quantum states \(\rho(x_T, y_T, z)\), and a probability density function \(p_{xTyTz}\). But, by definition of the class \(A^s_\delta\), we have that, conditioned on some classical state \(z\) of Eve, \(\rho(x_T, y_T, z)\) has suffered at most \(δN'\) corruptions by Eve, for any possible transcript views \(x_T, y_T\), and so, by the above analysis, its \(ABCR\) subsystems contains \(\text{Tr}_E(U_N \cdots U_1|ψ⟩⟨ψ|U_1^T \cdots U_N^T),\) a perfect version of \((Π ⊗ I^R)(|ψ⟩)\) for any views \(x_T, y_T\) of the transcripts of Alice and Bob, respectively. Hence, tracing over all subsystems but \(ABCR\), we obtain \((Π ⊗ I^R)(|ψ⟩),\) and the simulation protocol succeeds with zero error at simulating any noiseless protocol of length \(N\) against all adversaries in \(A^s_\delta\).

We have thus established the following: with \(q = |Σ|\) chosen according to Lemma \(Π\) (we use a tree code of arity \(d = 48\) and distance parameter \(α = \frac{26}{39}\)), \(R_C = \frac{\frac{1}{2} \log q}{4(1 + \frac{1}{80}) \log q} ≥ \frac{1}{8 \log q}\), \(R_E = \frac{1}{\log q}\) and \(δ = \frac{1}{80}\), we have that for all \(N\), there exists a universal simulation protocol in the shared entanglement model which, given black-box access to any two-party quantum protocol of length \(N\) in the noiseless model, succeeds with zero error at simulating the noiseless protocol on any input (independent of what is in the purifying register held by Eve) while transmitting \(\frac{1}{R_C \log q} N\) symbols in an alphabet \(Σ\) of size \(q\) over any adversarial channel with error rate \(δ\), and consuming \(\frac{R_E}{R_C} N\) EPR pairs, which proves Theorem \(Π\).

## 5 Tolerating Maximal Error Rates

We show how we can modify the basic protocol described in the last section such that an improved analysis can show that it tolerates up to \(\frac{1}{2} - ε\) error rate, for arbitrarily small \(ε > 0\), in the shared entanglement model. This is optimal: we also prove that no interactive protocol can withstand an error rate of \(\frac{1}{2}\) in that model. More formally, we prove the following results.

**Theorem 2** Given any two-party quantum protocol of length \(N\) in the noiseless model, no protocol in the shared entanglement model can tolerate an error rate of \(\frac{1}{2}\) and succeed in simulating the protocol with lower error than the best unidirectional protocol in the worst case. This result holds in oblivious as well as alternating communication models. More precisely,
for all noiseless protocol lengths $N \in \mathbb{N}$, for all communication rates $R_C > 0$, transmission alphabet sizes $q \in \mathbb{N}$, entanglement consumption rates $R_E \geq 0$, for all simulation protocols $S$ in the shared entanglement model with the above parameters, there exists an adversary $A^S \in A^S_{\frac{1}{2}}$ and an unidirectional protocol $U$ such that for all noiseless protocols $\Pi$ of length $N$, $\|S^U(A^S) - \Pi\|_\diamond \geq \|U - \Pi\|_\diamond$.

**Theorem 3** Given an adversarial channel in the shared entanglement model with error rate strictly smaller than $\frac{1}{2}$, we can simulate any noiseless protocol of length $N$ with negligible error over this channel using a number of transmission linear in $N$, and consuming a linear amount of EPR pairs. More precisely, there exists a constant $c > 0$ such that for arbitrary small $\varepsilon > 0$, there exist a communication rate $R_C > 0$, an alphabet size $q \in \mathbb{N}$, and an entanglement consumption rate $R_E \geq 0$ such that for all noiseless protocol lengths $N \in 2\mathbb{N}$, there exists a universal simulator $S$ in the shared entanglement model of length $N'$ with communication rate $R_C$, transmission alphabet size $q$, entanglement consumption rate $R_E$, which succeeds with error $2^{-cN}$ at simulating all noiseless protocols of length $N$ against all adversary in $A^S_{\frac{1}{2}-\varepsilon}$.

**5.1 Proof of Optimality**

To prove Th. 2, the argument of [21] in the classical case applies here as well: we only need to notice that if the error rate is $\frac{1}{2}$ with alternating communication in the shared entanglement model, then an adversary can completely corrupt all of the transmissions of either Alice or Bob, at his choosing, say Bob. In particular, he could replace all of his transmission by a fixed message, and leave Alice’s message unchanged. But then effectively Bob does not transmit any information to Alice, and this protocol can be simulated in the unidirectional model. Indeed, for a fixed register $E$, transmission alphabet $\Sigma$ of size $q$, noiseless protocol length $N$, simulation protocol length $N'$, taking the adversary $A^S_{\frac{1}{2}}$ described above which maps all transmissions from Bob to Alice to a fixed symbol $e_0 \in \Sigma$, for any simulator $S$ of length $N'$ that tries to simulate a noiseless protocol $\Pi$ of length $N$, we can take $M^U_1$ which runs sequentially all operations of Alice in $S$ while replacing all messages of Bob by $e_0$, then the quantum communication from Alice to Bob would be the simulation protocol messages to Bob along with Bob’s share of the entanglement in $T_B$, who would then take $M^U_2$ to be the sequential application of all his operations in $S$. This actually simulate $S$ running against adversary $A^S_{\frac{1}{2}}$ for any noiseless protocol and any input, and the outputs are the same.

Note that the proof also applies in an oblivious model for noisy communication, since in an oblivious model, the order in which the parties speak is fixed by the protocol and does not depend on the input or the actions of the adversary, and then the adversary can choose to disrupt all the messages of the party who communicate at most half the messages. Hence, the proof also extends to the case of oblivious communication, but not necessarily alternating. In such a case, the simulation protocol would also define a function $\text{peak} : [N'] \rightarrow \{A, B\}$ known to all (Alice, Bob and Eve) which tells whose turn it is to speak and is independent of both the input and of the action of Eve.
We can even extend the argument to the case of a speak function which depends on some secret key and is unknown to Eve, so Eve does not always know who is going to speak more often. In that case, Eve can flip a random bit (for example by measuring a $|+\rangle$ state in the computational basis) to decide which party’s communication she is going to corrupt (of course, with the reasonable assumption in the case of classical communication that Eve can see who speaks before she decides whether or not to corrupt a message). In this case, the statement is changed to $\|S^\Pi(A^S) - \Pi\|_\diamond$ is bounded away from zero, as can be seen by considering, for increasing $N$, some family of protocols computing, for example, the bitwise parity function of $\frac{N}{2}$ bits output by both parties or the swap function in which Alice and Bob want to exchange their $A, B$ registers. An extension of the argument of the proof of Theorem 5 shows that the fidelity is also bounded away from 1 for the case of protocols computing the inner product binary function. To reach the $\frac{1}{2}$ bound on the tolerable error rate, the parties would then need an adaptive strategy which depends on the sequence of errors applied by the adversary. However, this is dangerous in a noisy model: depending on the error pattern, the parties might not agree on whose turn it is to speak, and they could run into synchronisation problems.

5.2 Proof of Achievability

5.2.1 Description of the Simulation

The proof of achievability is somewhat more involved. It follows ideas similar to that of the basic simulation, but everything must be carefully analysed and optimized. We start by setting up the new notation enabling us to do so. The simulation protocol is essentially the same as the basic simulation one, the intuition given in section 4.2 still applies here, but different parameters which were fixed in the basic case now depend on the parameter $\varepsilon$. In particular, the distance parameter $\alpha = 1 - \varepsilon_\alpha$ now varies, as well as the length of the protocol $N' = l_\varepsilon N$. Since the parties have access to shared entanglement, they do not need to distribute it at the beginning of the protocol, and they can also use it to generate a secret key unknown to the adversary Eve. The secret key is used to generate a blueberry code with erasure parameter $\varepsilon_\beta = \frac{|\Sigma| - 1}{|\Gamma| - 1}$, for $\Sigma$ the tree code alphabet and $\Gamma$ the blueberry code alphabet. Each of the tree code transmission alphabet symbols are then reencoded with the blueberry code before transmission over the noisy channel, and an error caused by the adversary is detected as an erasure with probability $1 - \varepsilon_\beta$. When an erasure is detected by either party in a round, that party does not try to evolve the protocol in that particular round, so the corresponding trit sent is going to be 0, and also the teleportation decoding bits are 00. Otherwise, the structure of the protocol is mainly unchanged. The summary of the optimized protocol is as follows: Alice and Bob repeat the following for $i = 1 \cdots \frac{N'}{2}$:

1. Alice decodes the blueberry encoding of Bob’s possibly corrupted last transmission: if she detects an erasure, she sets $M_iA = 0, x_{iAD} = z_{iAD} = 0$ and $f_{i-1} = \perp$, and skips to step 4 below. Else, she decodes the transmission as $f_{i-1} \in \Sigma$, a possibly corrupted version of Bob’s last tree encoding $f_{i-1}$, and continue with step 2.
2. Alice computes \( s^i_B = D(f'_1 \cdots f'_{i-1}) \), and extracts \( b^i_\ell = (x^i_B z^i_B, M^i_B, x^i_{BM} z^i_{BM}) \), \( \ell = 1 \cdots i - 1 \), her best guess for Bob’s messages, and the corresponding \( C^i_{IB}, j^i_{IB} \).

3. Also using \( s_A \), she computes her best guess for the form \( |\psi_i\rangle \) of the joint register, and the corresponding \( x^i_{AD} z^i_{AD}, M^i_A, C^i_A, j^i_A \).

4. She decodes the teleportation by applying \( Z^i_{AD} X^i_{AD} \) to register \( T^2_A \) and swaps this with the \( C_A \) register.

5. She tries to evolve the simulation by applying \( U^{M^i_A} \) to the \( AC_A \) register.

6. She teleports back the \( C_A \) register to Bob using entanglement in register \( T^2_{i-1} \) and gets outcomes \( x^i_{AM} z^i_{AM} \).

7. Alice updates her state \( s_A \) by following edge \( a_i = (x^i_{AD} z^i_{AD}, M^i_A, x^i_{AM} z^i_{AM}) \), computes \( e_i = E(a_1 \cdots a_i) \) and transmits the blueberry encoding of \( e_i \) using the channel to Bob.

8. Upon reception of a possibly corrupted version of Alice’s last transmission, Bob decodes the blueberry code layer: he either detects an erasure and sets \( e'_i = \perp \), or else decode the transmission as \( e'_i \in \Sigma \), a possibly corrupted version of \( e_i \).

9. Bob computes \( x^i_{BD} z^i_{BD}, M^j_B \) analogously to Alice, depending on whether or not he detects an erasure. If not, he decodes \( s^i_A = D(e'_1 \cdots e'_i) \) and also uses \( s_B \) to compute these. He then performs actions on his side analogous to Alice’s, first swapping register \( T^2_{i-1} \) with \( C_B \), then using the \( T^2_B \) register to teleport back the \( C_B \) register to Alice, computes \( f_i \) and transmits the blueberry encoding of \( f_i \) to Alice. Round \( i \) completes upon reception by Alice of a possibly corrupted version of this message.

After these \( \frac{N'}{2} \) rounds, they both extract their protocol outcome from the \( \tilde{A} \tilde{B} \tilde{C} \) registers specified by the noiseless protocol embedding.

5.2.2 Analysis

Similar to section 4.4, the analysis is first carried conditional on some respective views of Alice and Bob of the transcript at each round, but now averaging over the shared secret key used for the blueberry code, and also conditional on some classical state \( z \) of the \( Z \) register of Eve, and the conclusion holds in that case. In particular, if the adversary has an adaptive and probabilistic strategy, we condition on some strategy consistent with the transcript already conditioned on. We come back later to this issue.

To analyse this protocol, we once again define a function \( P(i) \) such that we know the protocol succeeds whenever \( P(\frac{N'}{2} + 1) \geq N + 1 \). Here also, if we refer to the form of the state \( |\psi_i\rangle \) on the joint register \( \tilde{A} \tilde{B} \tilde{C} \tilde{E} \) at the beginning of round \( i \) (or at the end of round \( i - 1 \) rewritten as in (4.2), then \( P(i) = r_i - 2t_i \) will do. We now have three kinds of rounds: good rounds in which both parties can decode correctly the other party’s state, bad rounds in which at least one party makes a decoding error, and the new erasures rounds, in which
no party makes an actual decoding error, but at least one of them decodes an erasure from
the blueberry code, and so does not try to do anything on the quantum register before
teleporting back. We have an analogue of the technical Lemma 3 and its corollary. The
proofs are omitted since they are nearly identical to the proofs in the basic simulation case,
the only difference being that a party who detects an erasure does not take any action and
by consequence does not affect \( P(i) \).

**Lemma 4** At the end of round \( i \), define

\[
\begin{align*}
N^i_g &= |\{j : j \leq i, \text{round } j \text{ was good}\}|, \\
N^i_b &= |\{j : j \leq i, \text{round } j \text{ was bad}\}|, \\
N^i_e &= |\{j : j \leq i, \text{round } j \text{ was an erasure round}\}|.
\end{align*}
\]

Then \( P(i + 1) \geq N^i_g - 4N^i_b \).

**Corollary 4** If \( P(\frac{N'}{2} + 1) \geq N + 1 \), then the simulation succeeds.

Hence, we once again want to bound the ratio of bad to good rounds as a function of the
corruption rate to prove the success of the simulation. To do so, we show that depending on
a given tolerable error rate \( \frac{1}{2} - \varepsilon \), we can vary the distance parameter \( \alpha = 1 - \varepsilon_\alpha \) of the tree
codes used by Alice and Bob and also the erasure parameter \( \beta = 1 - \varepsilon_\beta \) of the blueberry
codes they use, and get this ratio as low as desired (except with negligible probability in the
random choice of the shared secret key used for the blueberry code). However, since there
is now a third kind of rounds, we also need to make sure that the ratio of good rounds vs.
erasure rounds does not get arbitrarily low, so that we can show \( P(\frac{N'}{2} + 1) \geq N + 1 \). We focus
on the number \( N_g = N^{N'}_g + 1 \), \( N_b = N^{N'}_b + 1 \) and \( N_e = N^{N'}_e + 1 \) of good, bad and erasure rounds
in the whole simulation, respectively. To bound the fraction of bad rounds as a fraction of
the corruption rate, we need the corollary of the following technical lemma, which derives a
new bound on tree codes with an erasure symbol. This result only talks about the structure
of such codes independently of our application, and so might have applications in a classical
interactive coding setting as well.

**Lemma 5** If there is a bound \( \delta \) on the fraction of the total number of transmission \( N' \) that
are corrupted and not detected as erasure by the blueberry code, then the number \( N_b \) of bad
rounds in the whole simulation is bounded by \( N_b \leq (2\delta + \varepsilon_\alpha)N' \) for \( \alpha = 1 - \varepsilon_\alpha \) the distance
parameter of the tree code used by Alice and Bob.

**Proof.** For any \( 1 \leq i \leq j \leq \frac{N'}{2} \), let \( I^A(i,j), I^A_b(i,j), I^A_g(i,j) \) be the subset of rounds
\( i, i+1, \ldots, j-1, j \) in which the symbol Alice gets from the blueberry decoding is an erasure,
an actual error, or the actual non-corrupted transmission, respectively. Note that these are
disjoint sets satisfying \( I^A_e(i,j) \cup I^A_b(i,j) \cup I^A_g(i,j) = [i,j] = \{i,i+1,\ldots,j-1,j\} \). Similarly, let
\( J^A_e(i,j), J^A_b(i,j), J^A_g(i,j) \) be the subset of \( [i,j] \) in which the sequence of messages Alice gets
from the tree decoding corresponds to a failure (note \( I^A_e(i,j) \subseteq J^A_e(i,j) \)), an actual decoding
error, or a correct decoding, respectively. Again note that \( J^A(i, j) \cup J^B(i, j) \cup J^A(i, j) = [i, j] \), a disjoint union. We can set up similar notation for Bob with \( A \)'s replaced by \( B \)'s, and then we have

\[
N_b = |J^A_b(1, \frac{N'}{2}) \cup J^B_b(1, \frac{N'}{2})|,
\]

\[
|I^A_b(1, \frac{N'}{2})| + |I^B_b(1, \frac{N'}{2})| \leq \delta N',
\]

so the statement we wish to prove is

\[
|J^A_b(1, \frac{N'}{2}) \cup J^B_b(1, \frac{N'}{2})| \leq 2\delta N' + \varepsilon N'.
\]

We prove the following stronger statements:

\[
|J^A_b(1, \frac{N'}{2})| \leq 2|I^A_b(1, \frac{N'}{2})| + \frac{1}{2} \varepsilon N'
\]

and

\[
|J^B_b(1, \frac{N'}{2})| \leq 2|I^B_b(1, \frac{N'}{2})| + \frac{1}{2} \varepsilon N'.
\]

Note that everything is symmetric from Alice’s and Bob’s point of view, so we only prove the statement from Alice’s. To lighten the notation, we drop the \( A \) superscripts. For any subset \( C = \{c_1, \ldots, c_C\} \) of \( \frac{N'}{2} \) and any two strings \( \bar{e} = e_1 \cdot e_t, \bar{e}' = e'_1 \cdot e'_t \in \Sigma' \), define \( \Delta_C(\bar{e}, \bar{e}') = |\{i \in C : i \leq t, e_i \neq e'_i\}| \). Note that with \( C = [\frac{N'}{2}] \setminus C \), \( \Delta(\bar{e}, \bar{e}') = \Delta_C(\bar{e}, \bar{e}') + \Delta_C(\bar{e}, \bar{e}') \), and \( \Delta_C(\bar{e}, \bar{e}') \leq |C| \).

We are now ready to prove the statement. We prove by induction on \( t \) that \( |J_b(1, t)| \leq 2|I_b(1, t)| + \varepsilon t \). The base case is obvious: for \( t = 1 \), if there is no transmission error during the first round, then there is no decoding error, and otherwise \( 1 \leq 2 + \varepsilon \). If in round \( t \) Alice detects an erasure or decodes correctly, then the induction is trivial. Hence, for the induction step, we consider the case of a bad decoding. Let \( \bar{a} \in [d]^t \) be the sequence of transmitted messages, \( \bar{e} = E(\bar{a}) \in \Sigma' \) the corresponding sequence of transmissions, \( \bar{e}' \in \Sigma' \) the sequence of possibly corrupted receptions, \( \bar{a}' = D(\bar{e}') \in [d]^t \) the sequence of decoded messages, and \( \bar{e}'' = E(\bar{a}') \) its reencoding. Then, by the decoding condition, \( \Delta(\bar{e}'', \bar{e}') \leq \Delta(\bar{e}, \bar{e}') \). Let \( \ell = L(\bar{a}, \bar{a}') \) be the distance of \( \bar{a}, \bar{a}' \) to their least common ancestor, then \( \Delta_{[1, t-\ell]}(\bar{e}'', \bar{e}) = 0 \). Note that \( 1 \leq \ell \leq t \). By the induction hypothesis,

\[
|J_b(1, t - \ell)| \leq 2|I_b(1, t - \ell)| + \varepsilon t - \ell,
\]

in which we vacuously set \( J_b(1, 0) = I_b(1, 0) = \emptyset \). We then have by definition

\[
|J_b(1, t)| = |J_b(1, t - \ell)| + |J_b(t - \ell + 1, t)|,
\]

\[
|I_b(1, t)| = |I_b(1, t - \ell)| + |I_b(t - \ell + 1, t)|,
\]

31
and \(|t - \ell + 1, t| = \ell\), so we only have to prove
\[
|J_\ell(t - \ell + 1, t)| \leq 2|I_\ell(t - \ell + 1, t)| + \varepsilon_\alpha \ell
\]
and we are done.

Let \(K_\ell = [t - \ell + 1, t]\) and \(K_s = \{i \in K_\ell : e_i'' = e_i\}\), then \(|K_\ell| = \ell = \Delta_{K_\ell}(\bar{e}'', \bar{e}) + |K_s|\), and
in fact \(\Delta(\bar{e}'', \bar{e}) = \Delta_{K_s}(\bar{e}'', \bar{e})\) since \(\Delta_{[1, t-\ell]}(\bar{e}'', \bar{e}) = 0\). But then by the tree code condition, \(\Delta_{K_s}(\bar{e}'', \bar{e}) \geq \alpha \ell\), and since we have \(\alpha = 1 - \varepsilon_\alpha\), we find \(|K_s| \leq \varepsilon_\alpha \ell\). If we define, for \(v \in \{e, b, g\}\),
\[
J_v = J_v(t - \ell + 1, t),
I_v = I_v(t - \ell + 1, t),
K_d = \{i \in K_\ell \setminus (K_s \cup I_e) : e_i' \neq e_i \text{ and } e_i' \neq e_i''\},
K_a = [t] \setminus ([1, t - \ell] \cup K_s \cup I_e \cup K_d)
= (I_b \cup I_g) \setminus (K_s \cup K_d),
\]
then we can notice that, since \(\Delta_{[1, t-\ell]}(\bar{e}'', \bar{e}) = \Delta_{K_s}(\bar{e}'', \bar{e}) = 0\), \(\Delta_{I_e}(\bar{e}'', \bar{e}') = \Delta_{I_e}(\bar{e}, \bar{e}') = |I_e|\)
and \(\Delta_{K_d}(\bar{e}'', \bar{e}') = \Delta_{K_d}(\bar{e}, \bar{e}') = |K_d|\), the decoding condition \(\Delta(\bar{e}'', \bar{e}') \leq \Delta(\bar{e}, \bar{e}')\) is equivalent to \(\Delta_{K_s}(\bar{e}'', \bar{e}') \leq \Delta_{K_a}(\bar{e}, \bar{e}')\). But for all \(i \in K_a\), \(e_i'' \neq e_i\) and either \(e_i' = e_i''\) or \(e_i' = e_i\), so that exactly one of the two equalities holds. Hence,
\[
\Delta_{K_a}(\bar{e}'', \bar{e}') = |\{i \in (I_b \cup I_g) \setminus (K_s \cup K_d) : e_i' = e_i\}|
= |I_g \setminus (K_s \cup K_d)|
\]
and
\[
\Delta_{K_a}(\bar{e}, \bar{e}') = |I_b \setminus (K_s \cup K_d)|,
\]
so we can restate the equivalent decoding condition as
\[
|I_g \setminus (K_s \cup K_d)| \leq |I_b \setminus (K_s \cup K_d)|.
\]
We have
\[
K_\ell = J_e \cup J_b \cup J_g
= I_e \cup I_b \setminus (K_s \cup K_d) \cup I_g \setminus (K_s \cup K_d) \cup K_s \cup K_d,
\]
so
\[
\ell = |K_\ell|
= |J_e| + |J_b| + |J_g|
\leq |I_e| + |I_b \setminus (K_s \cup K_d)| + |I_g \setminus (K_s \cup K_d)| + |K_s| + |K_d|
\leq |I_e| + 2|I_b \setminus (K_s \cup K_d)| + |K_s| + |K_d|
\leq |I_e| + 2|I_b| + |K_s|
\]
in which we used the fact that \(K_d \subseteq I_b\) in the last inequality. But then \(|I_e| \leq |I_e|, |J_g| \geq 0\) and \(|K_s| \leq \varepsilon_\alpha \ell\), so \(|J_b| \leq 2|I_b| + \varepsilon_\alpha \ell\), as required. \(\blacksquare\)
Corollary 5 If the corruption rate satisfies \(0 \leq c < \frac{1}{2}\), then except with probability smaller than \(2^{-\Omega(N')}\) for \(N'\) the length of the simulation protocol, the total number of bad rounds in the simulation is bounded by \(N_b \leq (2\varepsilon\beta + \varepsilon\alpha)N'\) for \(\alpha = 1 - \varepsilon\alpha\) the distance parameter of the tree code and \(\beta = 1 - \varepsilon\beta\) the erasure parameter of the blueberry code.

Proof. If the transmitted symbol is \(g_i \in \Gamma\) after a blueberry encoding \(B_i\) (actually, \(B_i^A\) or \(B_i^B\)) and, conditional on the classical state of Eve and based on some measurement outcomes \(z_i\), she chooses to corrupt \(g_i\) into a different \(g_i' \in \Gamma\), this action is independent from the randomness used in \(B_i\), and then it holds that \(\Pr[B_i^{-1}(g_i') \in \Sigma | z_1, \cdots, z_i] = \varepsilon\beta\). This is independent of the classical state and any measurement outcome \(z_i\) of Eve. Then with a corruption rate \(c\) bounded by some constants \(\varepsilon\beta \leq c < \frac{1}{2}\), the proof of Lemma 2 tells us that with probability \(1 - 2^{-\Omega(N')}\) at least a \(c(1 - 2\varepsilon\beta)\)-fraction of the transmissions are detected as erasures. But the total number of corruption is \(cN'\), so there are at most \(cN' - (c - 2c\varepsilon\beta)N' = 2c\varepsilon\beta N' < \varepsilon\beta N'\) actual transmission error, except with probability negligible in \(N'\). Taking \(\delta = \varepsilon\beta\) in the statement of Lemma 5 gives the result. If \(0 \leq c \leq \varepsilon\beta\), then the result is immediate from Lemma 5 and the total number of corruption, also with \(\delta = \varepsilon\beta\).

With the above result in hand, we can show that if the corruption rate is below \(\frac{1}{2}\) and we take \(\varepsilon\alpha = \frac{1}{20}\varepsilon, \varepsilon\beta = \frac{1}{40}\varepsilon, l_r = \frac{N'}{N} \geq 2\varepsilon(1 + \frac{1}{N})\), then except with negligible probability, the simulation succeeds:

\[
P\left(\frac{N'}{2} + 1\right) \geq N_g - 4N_b
= \frac{N'}{2} - N_e - 5N_b
\geq \varepsilon N' - 5N_b
\geq \varepsilon N' - 5(2\varepsilon\beta + \varepsilon\alpha)N'
= N'(\varepsilon - \frac{5}{10}\varepsilon - \frac{5}{20}\varepsilon)
= \frac{1}{2}\varepsilon N'
= \frac{1}{2}\varepsilon l_r N
\geq N + 1.
\]

The first inequality is from Lemma 4, the first equality is by definition of \(N_g, N_b, N_e\), i.e. \(\frac{N'}{2} = N_g + N_b + N_e\), the second inequality is from the fact that the number of erasure rounds is bounded by the number of corruption, i.e. \(N_e \leq (\frac{1}{2} - \varepsilon)N'\), and the third inequality is from our bound on \(N_b\) due to Corollary 5 which holds except with negligible probability. The fact that the simulation succeeds is then immediate from Corollary 4.

The above statement holds conditioned on some classical state \(z\) of the \(Z\) register of Eve, and some respective views of Alice and Bob of the transcript at each round. To prove Theorem 3 we have to argue similar to what is done in section 4.4 how to translate these
results to the state output by the protocols, even when we consider inputs entangled with some reference register $R$. We do not repeat this whole analysis here, since it is nearly identical once we make the following note. An arbitrary Eve fitting in the framework of the shared entanglement model could have adaptive, probabilistic behavior based on previous measurement outcomes. But these probabilistic choices must be independent of the secret key generated by Alice and Bob for the blueberry code, so similarly to section 4.4 for each probabilistic choice of Eve, the above result holds, so summing over all such choices, the result stays the same, proving Theorem 3.

6 Results in Other Models

By adapting the results we have obtained in the shared entanglement model for an adversarial error model, we can obtain many other interesting results. We first complete our study of the shared entanglement model with results in a random error setting. We then consider the quantum model and obtain results for both adversarial and random error settings. We also present a result hinting at the fact that the standard forward quantum capacity of the quantum channels used might not be the quantity that is best suited for our interactive communication scenario. We also consider a variation on the shared entanglement model in which, along with the noisy classical communication, the shared entanglement is also noisy.

6.1 Shared Entanglement Model with Random Errors

**Theorem 4** Given a two-party quantum protocol of length $N$ in the noiseless model and any $C > 0$, there exists a simulation protocol in the shared entanglement model that is of length $O(\frac{1}{C}N)$ and succeeds in simulating the original protocol with negligible error over classical binary symmetric channels of capacity $C$. More precisely, there exists constants $c, l_r > 0$ such that given any classical binary symmetric channel $\mathcal{M}_C$ of capacity $C > 0$ and noiseless protocol length $N \in 2N$, there exist a universal simulator $S$ in the shared entanglement model of length $N'$ with communication rate $R_C \geq l_rC$, transmission alphabet of size 2, entanglement consumption rate $R_E \leq 1$, which succeeds with error $2^{-cN}$ at simulating all noiseless protocols of length $N$ over $\mathcal{M}_C$.

**Theorem 5** There exists a sequence of two-party quantum protocols of increasing length $N$ in the noiseless model such that for all $C > 0$, any corresponding sequence of simulation protocol of length $o(\frac{1}{C^2}N)$ in the shared entanglement model fails at outputting the final state with low error on some input over classical binary symmetric channels of capacity $C$. Moreover, the family of quantum protocol can be chosen to be one computing a distributed binary function. More precisely, there exists a sequence $\{\Pi_N\}_{N \in 2N}$ of two-party quantum protocols and constants $d, \varepsilon > 0$ such that for all $N_0 \in \mathbb{N}$, there exist $N \geq N_0$ and $C > 0$ such that for any $R_E \geq 0$ and any simulation protocol $S$ in the shared entanglement model of length $N' = \frac{d}{C^2}N$ with communication rate $R_C = \frac{N}{N'}$ and arbitrary entanglement consumption rate $R_E$, the simulation does not succeed with error $\varepsilon$ over the binary symmetric channels.
6.1.1 Discussion About Optimality

The above results show that, in the regime where we use binary symmetric channels of classical capacity close to 0, we cannot expect to do much better than what we achieve, up to a multiplicative constant in front of the $\frac{1}{d}$ dilation factor. If we want to perform better in that regime, we would have to use the specifics of the operations implemented by the noiseless protocol instead of just using it as a black-box, even if we are restricting to protocols computing binary functions. We could however hope to be able to get much better hidden constants, since ours do not match the case of one-way communication in which the constant can be made arbitrarily close to $\frac{1}{2}$ as the quantum message size increases. Another regime of interest would be for channels of capacity close to 1, in which our techniques dilate the length of the protocols by a large multiplicative constant even when the error rate is low. In the classical case, recent results of Kol and Raz [29] show how to obtain communication rates going to 1 as the capacity goes to 1. Using our representation for quantum protocols, we are able to adapt their techniques with ideas similar to those used here to obtain comparable results in the shared entanglement model (up to a factor of 2 for teleportation), and this result will appear in a forthcoming paper.

6.1.2 Proof of Theorem 4

In [38], it is stated that, given a transmission alphabet $\Sigma$ and a desired bound $\varepsilon$ on the probability of transmission error, there exists a $d > 0$ such that given a binary symmetric channel $\mathcal{M}_C$ of capacity $C$, there is a $p \in \mathbb{N}, p \leq d\frac{1}{C}$, an encoding function $E : \Sigma \rightarrow \{0, 1\}^p$ and a decoding function $D : \{0, 1\}^p \rightarrow \Sigma$ such that $\Pr[D(\mathcal{M}_C(E(e))) \neq e] \leq \varepsilon$. We use this in conjunction with the result of Theorem 1 and the Chernoff bound to obtain the following result: Taking $\varepsilon = \frac{1}{90} < \frac{1}{80}$, $\Sigma$ given by Lemma 1 for a tree code of arity 48 and distance parameter $\alpha = \frac{39}{40}$ and the corresponding $d > 0$, given a binary symmetric channel of capacity $C$ and the corresponding $p \in \mathbb{N}, E$ and $D$, if all the $\Sigma$ transmissions in the basic simulation protocol are done by reencoding over $\{0, 1\}^p$ with $E$ (and decoding with $D$), then except with probability $2^{-\Omega(N''')}$ for $N'' = 4(1 + \frac{1}{N})N$ the length of the basic simulation protocol over alphabet $\Sigma$, $N' = pN''$ the length of the oblivious simulation protocol over the binary symmetric channel, and $N$ the length of the noiseless protocol to be simulated, the error rate for transmission of $\Sigma$ symbols is going to be below $\frac{1}{80}$ and then by Theorem 1 the simulation succeeds.

6.1.3 Proof of Theorem 5

It is known that for a classical discrete memoryless channel such as the binary symmetric channel, entanglement-assistance cannot increase the classical capacity [4], and it is also known that allowing for classical feedback does not allow neither to increase the classical capacity. However, we might hope that allowing for both simultaneously might lead to improvements. This is not the case: classical feedback augmented by shared entanglement can be seen to be equivalent to quantum feedback, and it is also known that for discrete memoryless quantum channels, the classical capacity with unlimited quantum feedback is
equal to that with unlimited entanglement assistance [8]. Hence, in the shared entanglement model, the classical capacity of the binary symmetric channels used is not increased by the entanglement assistance and the other binary symmetric channel’s feedback. It is clear that for some protocols of length $N$ fitting our general framework in the noiseless model, like those accomplishing a quantum swap function or even a classical swap or bitwise XOR functions on inputs of size $\frac{N}{2}$, the parties must effectively exchange their whole inputs to output their final state. Hence, a dilation factor proportional to the inverse of the capacity $\frac{1}{C}$ is necessary since these protocols are equivalent to a communication of $\frac{N}{2}$ bits or qubits in each direction.

What we want to prove is even stronger: there exists a family of distributed binary functions such that this is necessary. We consider the inner product function $IP_n : \{0,1\}^n \times \{0,1\}^n \rightarrow \{0,1\}, IP_n(x, y) = \bigoplus_{i=1}^{n} x_i \land y_i$, which has been proved to have communication complexity of $\Theta(n)$ in both Yao’s and Cleve-Buhrman’s quantum communication complexity model [17].

But we know from [17] that any black-box protocol evaluating coherently the $IP_n$ function with small error can be used to transmit $n$ bits of classical information with small probability of error, and that any non-coherent unitary protocol to compute a classical function can be made coherent by doubling the amount of quantum communication (make a pseudo-copy of the output, and then run the protocol backward to get rid of the junk) while keeping the error parameter small. Hence, these protocols can be used to transmit $n$-bit strings over a channel of classical capacity $C$ with some small probability of failure, and by consequence for small enough error it requires at least $\frac{1}{C}n$ uses of the channel. Since for any small enough error, the communication complexity of $IP_n$ is $\Theta(n)$, $N \in \Theta(n)$ (if the protocol does not waste communication), and $N' \in \Omega(\frac{1}{C}n) = \Omega(\frac{1}{C}N)$ is required for the simulation to succeed. Note that we have made the reasonable assumption that we can run the simulation backward over the noisy channel at the same communication cost (or else that we start with a coherent protocol for the inner product function; the restriction of having the protocol compute the function in a coherent way is natural if we want to compose our quantum simulation protocols, since then they may be called on arbitrary quantum inputs). Details will appear in a future version of this work.

6.2 Quantum Model with Adversarial Errors

**Theorem 6** Given an adversarial channel in the quantum model with error rate strictly smaller than $\frac{1}{3}$, we can simulate any noiseless protocol of length $N$ over this channel using a number of transmission linear in $N$. More precisely, there exists a constant $c > 0$ such that for arbitrary small $\varepsilon > 0$, there exist a communication rate $R_C > 0$ and an alphabet size $q \in \mathbb{N}$ such that for all noiseless protocol lengths $N \in 2\mathbb{N}$, there exists a universal simulator $S$ in the quantum model of length $N'$ with communication rate at least $R_C$, transmission alphabet size $q$, which succeeds with error $2^{-cn}$ at simulating all noiseless protocols of length $N$ against all adversary in $A^{Q}_{\frac{1}{3} - \varepsilon}$.  

36
6.2.1 Discussion About Optimality

If we consider only perfect quantum error correcting codes for quantum data transmission, it is known that we cannot tolerate error rates of more than $\frac{1}{4}$ asymptotically, and so with the approach of first distributing entanglement and then using our $\frac{1}{2} - \varepsilon$ error rate protocol, this leads to overall tolerable error rates for the simulation of less than $\frac{1}{6}$. However, Crépeau, Gottesman and Smith [18] showed how to tolerate up to $\frac{1}{2}$ error rate asymptotically for data transmission if we consider approximate quantum error correcting codes, and using these would lead to $\frac{1}{4} - \varepsilon$ tolerable error rate for a two phase simulation protocol as described above. However, their register size as well as the number of communicated registers are linear in the number of transmitted qubits, so for our purpose using these would lead to communication rates of 0 asymptotically. It would be interesting to see whether we can do something similar with register size independent of the transmission size, but possibly dependent on the fidelity we want to reach and how close to $\frac{1}{2}$ (or some other fraction strictly larger than $\frac{1}{4}$) we want to get. Using these kinds of codes, if we want to do the simulation in two steps, an entanglement distribution part and then an actual simulation part, this is the best we can do. To tolerate higher error rates than what we achieve, we might hope to develop a fully quantum analogue of tree codes that does not require to first distribute entanglement before it can be used while robustly transmitting quantum information. However, to be able to coherently apply the noiseless protocol unitaries in the simulation, the developed quantum codes would require some properties for fault-tolerant computation, a problem not present in the classical case, since we can copy classical information and perform the computation on the copy. Finally, note that the proof of Theorem 2 also establishes that the bound of $\frac{1}{2}$ on the maximum error rate tolerable in an oblivious communication model applies here as well: no simulation protocol in the quantum model can succeed with arbitrarily small error against all adversaries in $A^Q_{\frac{1}{2}}$.

6.2.2 Proof of Theorem 6

The approach we take in the quantum model is to emulate the approach in the shared entanglement model by first using the provided quantum channels to distribute sufficient entanglement, and then by using them effectively as classical channels along with the entanglement to run the simulation protocol of section 5. Let us look at the parameters of the quantum error correcting codes (QECCs) we use to distribute entanglement.

For a given $\varepsilon > 0$, let $s = \frac{(|\Gamma|)}{(|\Gamma| - |\Sigma|)!}$ be the size of the shared secret key used to do the blueberry encoding in each round, so that in each round, two maximally entangled states of size $2s$ (i.e. states of the form $\sum_{j=0}^{2s-1} |j\rangle_{TA} |j\rangle_{TB}$) are used to generate the secret keys required in the protocol of section 5 and to create the EPR pairs required for teleportation. Then, for any given communication register size $q$ and simulation protocol in the shared entanglement model of length $N'$, we need to distribute a maximally entangled state of $N' \log_q (2s)$ registers of that size to perform the whole protocol of section 5.

If we allow, in the entanglement distribution phase, for $l_c N'$ transmissions of registers of size $q$ from Alice to Bob, we want quantum error correcting codes on a register of size $q$, a
transmission rate \( R_Q = \frac{1}{c} \log_q(2s) \), and some corresponding maximum tolerable error rate \( \delta \). We only consider here exact quantum error correcting codes, but the analysis extends to approximate ones for which we might also allow for some deviation from perfect transmission. To choose \( q, l_c, \) and \( \delta \), we first note that in the actual simulation part we need to transmit classical messages chosen from a set of size \(|\Gamma|\) over the same quantum channel used to distribute entanglement, so a first constraint is \( q \geq |\Gamma| \). Then, to make sure that the simulation succeeds in the second part, the total number of corruptions should be bounded by \( \left( \frac{1}{2} - \varepsilon \right) N' = \frac{N'}{2} - \varepsilon N' \). Hence, since an adversary could choose to put all of her allowed corruptions in the first part instead of the second part, the QECC should also be able to recover from the same number of errors, \( \frac{N'}{2} - \varepsilon N' \). If the QECC can tolerate an error rate of \( \delta \), we need the length of the entanglement distribution part to satisfy \( l_c \geq \frac{1-2\varepsilon}{2\delta} \), and then the whole simulation protocol can tolerate \( \frac{N'}{2} - \varepsilon N' \) adversarial errors during a total of \((l_c+1)N'\) communications, i.e. it can tolerate an error rate of \( \frac{1-2\varepsilon}{2(l_c+1)} \) (note that if we restrict ourselves to an alternating instead of oblivious communication model, a factor of 2 appears in front of \( l_c \) due to the fact that the adversary can choose to corrupt the transmissions of one particular party during the entanglement transmission phase, but there is now twice as much communication during that phase).

We now use a high-dimensional quantum Gilbert-Varshamov bound \([2, 20]\) stating that for arbitrarily small \( \varepsilon' > 0 \), there exists strictly positive communication rate \( R_Q > 0 \) and large enough transmission alphabet size such that families of quantum codes of arbitrarily large length exist which can tolerate a fraction \( \frac{1}{4} - \varepsilon' \) of errors and still perfectly correct the quantum state. We use these and the above analysis to tolerate error rate \( \frac{1}{6} - \varepsilon \) for our simulation protocols (this result is obtained in an oblivious model of communication; in an alternating model of communication, we are able to tolerate error rates of \( \frac{1}{10} - \varepsilon \)).

Taking \( l_c = 2\left(\frac{1-2\varepsilon}{1-4\varepsilon}\right) \) for \( 0 < \varepsilon < \frac{1}{4} \) and choosing a \( q \) large enough as a function of \( \varepsilon \) such that \( R_Q \) is low enough for a QECC with the required parameters to exist, Alice uses her first \( l_c N' \) transmissions to distribute perfectly entanglement to Bob with the above QECC. By the above analysis, since the overall error rate is bounded by \( \frac{1}{6} - \varepsilon \), the error rate in the entanglement distribution phase is bounded by \( \frac{1}{4} - \varepsilon \) and the QECC can perfectly recover from this error rate and produce perfect entanglement. They then share enough entanglement to run the simulation of section 5. During the simulation phase, before transmission and after reception of an element of \(|\Gamma|\) through the channel, both the sender and the receiver measure the quantum communication register. These measurements have the effect of transforming all possible quantum actions of Eve into effectively classical actions. Indeed, conditioned on the results of the two measurements, the corresponding branches of the simulation proceed exactly as if the sender and the receiver had transmitted and received classical information over a classical channel, and doing so restricted the action of Eve into an essentially classical one. Moreover, if \( q \) is larger than \( \Gamma \) and Eve maps some of these classical messages outside of the span of \( \Gamma \), Alice and Bob only have to mark these as erasures so Eve does not gain anything by leaving the span of \( \Gamma \).

But then, the corresponding corruption rate of the adversary during the actual simulation phase is lower than \( \frac{1}{2} - \varepsilon \), and so, given any strategy which must be independent of the
generated secret key, the fraction of branches in which the secret key enables Alice and Bob to succeed is overwhelming over the measurement outcomes, and the remainder of the analysis goes as in section 5.2.2 proving Theorem 6.

6.3 Quantum Model with Random Errors

Theorem 7 Given a two-party quantum protocol of length $N$ in the noiseless model and any $Q > 0$, there exists a simulation protocol in the quantum model that is of length $O(\frac{1}{Q} N)$ and succeeds in simulating the original protocol with arbitrarily small error over quantum depolarizing channels of quantum capacity $Q$. More precisely, there exist a constant $l_r > 0$ and a function $f : \mathbb{N} \rightarrow \mathbb{R}^+$ with $\lim_{N \rightarrow \infty} f(N) = 0$ such that given any depolarizing channel $M_Q$ of quantum capacity $Q > 0$ and noiseless protocol length $N \in 2\mathbb{N}$, there exist a universal simulator $P$ in the quantum model of length $N'$ with communication rate $R_Q \geq l_r Q$, transmission alphabet size 2, which succeeds with error $f(N)$ at simulating all noiseless protocols of length $N$ over $M_Q$.

Theorem 8 There exist a sequence of two-party quantum protocols of increasing length $N$ in the noiseless model such that for all $Q_B > 0$, any corresponding sequence of simulation protocol of length $o(\frac{1}{Q_B} N)$ in the quantum model fails at outputting the final state with low error on some input over quantum depolarizing channels of quantum capacity with classical feedback $Q_B$. Moreover, the family of quantum protocol can be chosen to be one computing a distributed binary function. More precisely, there exists a sequence $\{\Pi_N\}_{N \in 2\mathbb{N}}$ of two-party quantum protocols and constants $d, \varepsilon > 0$ such that for all $N_0 \in \mathbb{N}$, there exists $N \geq N_0$ and $Q_B > 0$ such that for any simulation protocol $P$ in the quantum model of length $N'$ with communication rate $R_Q = \frac{N}{N'}$, the simulation does not succeed with error $\varepsilon$ over the quantum depolarizing channels.

Theorem 9 Given a two-party quantum protocol of length $N$ in the noiseless model, there exists a quantum depolarizing channel of unassisted forward quantum capacity $Q = 0$ and a simulation protocol in the quantum model with asymptotically positive rate of communication which succeeds in simulating the original protocol with arbitrarily small error over that quantum channel. More precisely, there exists constants $c, R_Q > 0$ such that given a particular depolarizing quantum channel $M_0^Q$ of forward quantum capacity $Q = 0$ and any noiseless protocol length $N \in 2\mathbb{N}$, there exist a universal simulator $P$ in the quantum model of length $N'$ with communication rate at least $R_Q$, transmission alphabet size 2, which succeeds with error $2^{-cN}$ at simulating all noiseless protocols of length $N$ over $M_0^Q$.

6.3.1 Discussion About Optimality

It is known that for some range of the depolarizing parameter, the quantum capacity with classical feedback $Q_B$ of the depolarizing channel is strictly larger than its unassisted forward quantum capacity $Q \geq 0$. In particular, there exists values for which $Q = 0$ but $Q_B > 0$. A careful analysis of the related 2-way entanglement distillation protocols (in particular their...
communication cost and their amount of interaction) reveals that there is some range of the depolarizing parameter for which we can achieve successful simulation even though $Q = 0$, by using the depolarizing channels in each direction to transmit the classical information. Note that $Q_B > 0$ if and only if the depolarizing parameter $\varepsilon' < \frac{2}{3}$, and so $Q_B > 0$ if and only if the quantum capacity assisted by two-way classical communication $Q_2 > 0$. In the case where we are given a depolarizing channel with $Q_B > 0$, we can modify the method used in the proof of Theorem 9 by iteratively using the recurrence method a constant number of times (constant in $N$, not in the depolarizing parameter!) on the noisy distributed EPR pairs until the depolarizing channels induced through teleportation over the noisy distilled EPR pairs has $Q > 0$, and then distribute entanglement over the induced channels using standard QECCs. We achieve asymptotically positive rates of communication for our simulation protocols. It is an interesting open question whether we can close the gap between our lower and upper bounds and always achieve successful simulation at a rate $O\left(\frac{1}{Q_B N}\right)$. The separation result regarding the forward, unassisted quantum capacity of the depolarizing channel requires some technical work, but the case of the erasure channel already makes it clear that in general for discrete memoryless quantum channels, the unassisted forward quantum capacity is not the most suitable quantity to consider in the setting of interactive quantum communication.

### 6.3.2 Proof of Theorem 7

For the random error case in the quantum model, we use techniques similar to the adversarial error case. Indeed, we split the protocol into two phases: an entanglement distribution part and an actual simulation phase.

To avoid technicalities, it is sufficient to adapt the result from section 4 for a basic simulation protocol of length $N''$ over some large alphabet $\Sigma$. We then only need to distribute $N''$ EPR pairs. For any depolarizing channel of quantum capacity $Q > 0$, we use standard quantum Shannon theory type coding to distribute entanglement at a rate of $\frac{d}{Q}$ for some $d > 0$ with low error. Then, for the actual simulation part, we use both the fact that the classical capacity $C$ is at least as large as the quantum capacity $Q$ for any quantum channel, and that a classical capacity achieving strategy for the depolarizing channel is just to simulate a binary symmetric channel (BSC) of capacity $C$ for each transmission by measuring the output in the computational basis, and then do block coding over the corresponding BSC (details are provided in [45]). We can then translate the arguments of the proof of Theorem 4 to design our classical strategy which succeeds with overwhelming probability (assuming perfect entanglement for now), and the output is arbitrarily close to the noiseless protocol one. Combining the bound on this error with the one from the entanglement distribution part, the simulation can be made to succeed with error less than $f(N)$ over the depolarizing channel of quantum capacity $Q$, for some function $f : \mathbb{N} \to \mathbb{R}^+$ which asymptotically goes to zero. Details will appear in a future version of this work.
6.3.3 Proof of Theorem 8

The idea for this proof is to use the symmetry of the depolarizing channel for entanglement distribution to actually simulate one direction of the use of the quantum depolarizing channel with classical feedback used for teleportation. Then apply a coherent version of the idea to use the inner product protocol to communicate, as in the proof of Theorem 5, to use the depolarizing channel to distribute quantum entanglement, and then further use the depolarizing channel (again with the inner product protocol used this time to communicate classical information) to teleport.

Similar to what was argued in the proof of Theorem 5 for classical communication, it is clear that for some protocols of length $N$ fitting our general framework in the noiseless model can be used to communicate up to $\frac{N}{2}$ qubits in each direction. Hence, since our simulation protocols of length $N'$ can be simulated by $N'$ uses of a depolarizing channel from Alice to Bob supplemented by classical feedback from Bob to Alice, we cannot have a rate of communication better than $\frac{N}{2Q_B}$ for small enough error. To prove that a protocol to compute a binary function is sufficient, we once again consider the inner product function $IP_n$. Note that what we achieved in the proof of Theorem 5 using the protocol for $IP_n$ is actually stronger than $\Theta(N)$ bits of classical communication: we had a coherent bit channel for $\Theta(N)$ cobits (coherent bits), which can be used to distribute $\Theta(N)$ ebits (EPR pairs). But then we can perform teleportation of $\Theta(N)$ qubits from Alice to Bob by once again using the $IP_n$ protocol, but this time to transmit the classical teleportation measurement information.

We have thus used the length $N'$ simulation protocol at most 4 times (depending whether or not the noiseless one was computing the $IP_n$ function coherently to begin with) over the depolarizing channel from Alice to Bob, with (free, perfect) classical feedback from Bob to Alice, and succeeded at transmitting $\Theta(N)$ qubits, so we must have that $N' \in \Omega(\frac{1}{Q_B}N)$ for the simulation to succeed with small error. Note that we once again make the reasonable assumption that in the case in which the initial protocol is not coherent, we can run the simulation protocol backward over the noisy channel at the same communication cost. Details will appear in a future version of this work.

6.3.4 Proof of Theorem 9

The case of the depolarizing channel requires some technical work, so for simplicity we first consider the case of the quantum erasure channel. For the quantum erasure channel, we use the fact that, for erasure probability $\frac{1}{2} \leq p < 1$, the (forward, unassisted) quantum capacity is 0 while the classical capacity is $1 - p$ and the entanglement generation capacity with classical feedback is at least $1 - p$. Moreover, the feedback required to achieve this bound is only one message of length linear in the size of the quantum communication. The strategy we use is the following: for a basic simulation protocol of length $N''$ over $\Sigma$, Alice distribute $N''$ EPR pairs to Bob by sending $\frac{4N''}{1-p}$ halves of EPR pairs over the quantum erasure channel. Then, except with negligible probability, at least $N''$ of them are received correctly, and Bob knows which these are. The feedback consist of informing Alice which
$N''$ pairs to use in the protocol, so that they both agree. This can be done over the quantum erasure channel (again except with negligible probability) with a classical message of length linear in $N''$.

Then, given a message set $\Sigma$ we can use the quantum erasure channel a constant number of times to decrease the probability of error in a classical transmission of any symbol $e \in \Sigma$ below $\frac{1}{60}$. Except with negligible probability, the fraction of $N''$ transmissions of symbols of $\Sigma$ transmitted in this way is below $\frac{1}{80}$. We can then use ideas similar to those in the proof of Theorem 6 to argue that the output is arbitrarily close to the noiseless protocol one. Details will appear in a future version of this work.

Now for the depolarizing channel, the idea is mostly the same, but we have to work harder to obtain (almost) noiseless entanglement. The unassisted forward capacity of the depolarizing channel is shown in [5] to be equivalent to one-way entanglement distillation yield. To separate one-way and two-way entanglement distillation, they use a combination of the recurrence method of [4], which is an explicitly two-way entanglement distillation protocol which can purify highly noisy entanglement, but does not have a positive yield in the limit of high fidelity distillation, along with their hashing method, a one-way protocol with positive yield in the perfect fidelity limit, but which does not work on highly noisy entanglement. However, we cannot hope to use this strategy to distill near perfect EPR pairs in our scenario since the hashing method as they describe it requires too much communication (however, we could probably use some derandomization argument to avoid communicating the random strings). To reduce the communication cost, we instead use a hybrid approach of entanglement distillation followed by quantum error correction.

Starting with a depolarizing channel with depolarizing parameter as high as possible, but still low enough to have $Q = 0$, we use it to distribute imperfect EPR pairs. This yields (rotated) Werner states with the highest possible fidelity to perfect EPR pairs, but such that one-way entanglement distillation protocols cannot have a positive yield of EPR pairs while two-way entanglement distillation protocols can (see section 6.4 for a definition of Werner states). We then do one round of the recurrence method for entanglement distillation to obtain a lesser number of Werner states of higher fidelity to perfect EPR pairs, and so we could now use one-way distillation protocols on these to obtain a positive yield of near perfect EPR pairs. Note that the amount of classical communication required up to this point is one message from Alice to Bob of linear length informing him of her measurement outcomes, and then one classical message of linear length from Bob to Alice informing her which states to keep as well as which rotation to apply to these (to go back to symmetric Werner form; log 12 bits of information per pair is sufficient for this purpose [5]). But we can now use these EPR pairs along with teleportation to effectively obtain a depolarizing channel of quantum capacity $Q > 0$, and so we use standard Quantum Shannon theory type coding over this quantum channel to distribute $N''$ near perfect EPR pairs. This new step has only required a linear amount of classical communication, and so after the initial very noisy entanglement distribution step, we only have three classical messages to send over the depolarizing channel of classical capacity $C > 0$, and so we can generate near perfect entanglement using the depolarizing channel a linear amount of times, and then go on to do
the actual simulation phase as above. Note that we are not yet insured of an exponential
decay of the error at this point, only that the error tends to zero in the limit of large $N$. To get exponential decay, adapt the above protocol such that before using teleportation and
QECC to distribute good entanglement, we perform a few more rounds of the recurrence
method until the Werner states reach fidelity parameter above 0.82. Except with negligible
probability, starting with some linear amount of noisy EPR pairs, after a constant number
of rounds of the recurrence method, we are left with sufficiently many less noisy EPR pairs
for our next step. At this point, it is known that there exist stabilizer codes achieving
the hashing bound (which has strictly positive yield for this noise parameter) and which
have negligible error. Using the fact that some classical capacity achieving strategy for the
depolarizing channel also has negligible error, we get the stated exponential decay in the
error. Details will appear in a future version of this work.

6.4 Noisy Entanglement

The last model we consider is a further variation on the shared entanglement model, in
which, along with the noisy classical links between the honest parties, the entanglement
these parties share is also noisy. Details will appear in a future version of this work.

There are many possible models for noisy entanglement; we consider a simple one in this
section, in which parties share noisy EPR pairs instead of perfect pairs. Following [4], we consider the so-called (rotated) Werner states

$$W_F = F|\Phi_{00}\rangle\langle\Phi_{00}| + \frac{1-F}{3}(|\Phi_{01}\rangle\langle\Phi_{01}| +
|\Phi_{10}\rangle\langle\Phi_{10}| + |\Phi_{11}\rangle\langle\Phi_{11}|),$$

which are mixtures of the four Bell states parametrized by $0 \leq F \leq 1$. Note that these are the result of passing one qubit of an EPR pair through a $T_{\varepsilon'}$ depolarizing
channel, for $F = 1 - \frac{3\varepsilon'}{4}$. The purification of these noisy EPR pairs is given to Eve. We use
the result of [4] to show that for any $F > \frac{1}{2}$, simulation protocols with asymptotically (in
$N \to \infty$, not in $F \to \frac{1}{2}$) positive communication rates and which can tolerate a positive error
rate can succeed with asymptotically zero error. This is optimal: at $F = \frac{1}{2}$, Werner states
are separable, so there is no way to use them in conjunction with classical communication
to simulate quantum communication.

6.4.1 Adversarial Errors on the Classical Communication

We first consider the case of adversarial errors. Let $l_c$ be the number of rounds of the
recurrence method for entanglement distillation necessary to reach the $F = 0.82$ bound. This
number is independent of $N$, and depends only on the initial $F$. As described in the proof
of Theorem 9, each round of the recurrence method only requires a linear length message
in each direction. After this bound is reached, one last linear length classical message is
sufficient to generate a linear amount of entanglement through teleportation via an induced
depolarizing channel of quantum capacity $Q > 0$, so standard quantum error correction
techniques enables us to extract near perfect entanglement at this point. Once we have near
perfect entanglement, we can use techniques from the basic simulation protocol to perform
successful simulation of noiseless protocols, hence achieving our goals. The protocol described
above requires the communication of $2l_c + 1$ messages to distill near perfect entanglement,
independent of $N$, followed by an actual simulation phase, so the simulation protocol can
tolerate a constant error rate (though inversely proportional in $l_c$), requires a constant rate of
noisy entanglement consumption (though exponential in $l_c$ since each round of the recurrence
method consumes at least half of the noisy EPR pairs), and has a constant, positive rate of
communication (though inversely proportional in the amount of consumed noisy EPR pairs).
Details will appear in a future version of this work.

6.4.2 Random Errors on the Classical Communication

The case of noisy communication through binary symmetric channels once again is immediate
from the adversarial error case by a concentration argument. The communication rate is
inversely proportional in the classical capacity $C$, and also in the number of noisy EPR
pairs consumed. Details will appear in a future version of this work.

7 Conclusion: Discussion and Open Questions

We perform simulation of interactive quantum protocols over noisy channels with only a
linear dilation factor. In particular, our approach is to replace irreversible measurements by
reversible pseudo-measurements in the Cleve-Buhrman model (shared entanglement, classi-
cal communication), and then in the analogous noisy model to teleport the corresponding
quantum communication register to avoid losing the quantum information it contains over
the noisy channel. With this approach, we have been able to prove that it is possible to
simulate the evolution of quantum protocols designed for noiseless quantum channels over
noisy channels with only a linear dilation factor. Moreover, in the case of adversarial channel
errors in which parties are allowed to pre-share a linear amount of entanglement, we were
able to prove that the error rate of $\frac{1}{2} - \varepsilon$ our simulation protocol can sustain is optimal
unless we generalize the noisy communication model such that the order in which the parties
take turn in the protocol can be adapted to the errors. But in a noisy setting, restricting
to non-adaptive protocols is natural. Otherwise, depending on the particular view of each
party on the evolution of the protocol due to previous errors, they could disagree on whose
turn it is to speak, and this would result in protocols that are not well defined.

To simplify the exposition, we chose not to optimize different parameters, such as com-
munication and entanglement consumption rates and communication register size. It is
possible to modify our results in a straightforward manner to transmit larger (noiseless
protocol) communication registers in each round, hence decreasing the amount of interac-
tion while still tolerating high error rates. It is also possible to adapt our findings to a
random error model in which parties are allowed to share entanglement but communicate
over binary symmetric channels of capacity $C > 0$, and then we obtain communication
rates proportional to $C$. It can be shown that, up to a hidden constant, this is optimal
for some family of distributed binary functions, for example the inner product functions
$IP_n : \{0,1\}^n \times \{0,1\}^n \rightarrow \{0,1\}, IP_n(x,y) = \oplus_{i=1}^n x_i \cdot y_i$. Our findings can also be adapted
to obtain similar (though not optimal) results for the quantum model (the noisy version of

44
Yao’s model), in which our simulation protocols run in two phases: a first phase in which a linear amount of entanglement is distributed with standard techniques from quantum Shannon theory for random noise and from quantum coding theory for adversarial noise, and then an actual simulation part in which the parties perform actions similar to those of the shared entanglement model. In an adversarial noise model, we show that we can tolerate an error rate of $\frac{1}{6} - \varepsilon$ in the quantum model, while for a random noise model in which the parties communicate over depolarizing channels of capacity $Q > 0$, we obtain rates proportional to $Q$. We also show that the use of depolarizing channels in both directions enables the simulation to succeed even for some quantum channels of unassisted forward quantum capacity $Q = 0$, and we extend our ideas to perform simulation with noisy entanglement as well as noisy classical channels.

Further directions for this research program would be to try to obtain better communication rates in all of the models discussed. In particular, we would like to study the interactive capacity of the depolarizing channel with depolarizing parameter $\varepsilon'$. The question of interactive capacity for the binary symmetric channel was raised in the classical context by Braverman in a survey article on the topic of interactive coding [11], and there has been recent developments in providing lower and upper bounds for this quantity [29]. In the classical setting, a particular problem with worst case interaction of one bit transmissions to which all classical interactive protocols can be mapped was proposed to study such a quantity. Since every interactive quantum protocol can be mapped onto our general problem, we propose to study such a quantity in the quantum domain. Would the interactive capacity of the binary symmetric channel (with entanglement assistance) for quantum protocols be the same as for classical protocols [29], up to a factor of two for teleportation? We will show in upcoming works that for bit flip probability $\varepsilon$, the lower bound of $\frac{1}{2} - O(\sqrt{H(\varepsilon)})$ holds, but do the techniques developed in [29] adapt to the quantum setting to obtain a matching upper bound of $\frac{1}{2} - \Omega(\sqrt{H(\varepsilon)})$? Another question that remains open is that of the highest tolerable adversarial error rate that can be withstood in the quantum model. To study this question, we would like to develop a fully quantum approach to our problem, and to do so new kinds of quantum codes might need to be developed. In particular, ideas from fault-tolerant quantum computation might need to be borrowed, due to the nature of quantum information. Another important question in the quantum setting is what would happen in a shared entanglement setting if along with the noisy classical communication, the entanglement provided were also noisy; we only investigate this question for a simple noise model for the entanglement, but other interesting models would be interesting to study. In particular, what about adversarial noise on the shared EPR pairs above the $\frac{1}{8}$ binary error rate limit? Note that below that bound, standard quantum error correction for qubits with teleportation can be used for distillation. Finally, the question of efficient simulation also is an interesting one, and we will show in upcoming works how to adapt the techniques developed by Brakerski and Kalai [9] to our setting to efficiently process the classical communication in our simulation protocols.
Acknowledgement  The authors are grateful to Louis Salvail, Benno Salwey and Mark M. Wilde for useful discussions. G.B. is supported in part by the Natural Sciences and Engineering Research Council of Canada, the Canada Research Chair program and the Canadian Institute for Advanced Research. A.T. is supported by NSERC and CIFAR. D.T. is supported by a Fonds de Recherche Québec-Nature et Technologies B2 Doctoral research scholarship.

References

[1] Dorit Aharonov, Alexei Kitaev, and Noam Nisan. Quantum Circuits with Mixed States. Proceedings of the 30th Annual ACM Symposium on Theory of Computing (1997).

[2] Alexei Ashikhim, and Emanuel Knill. Non-binary quantum stabilizer codes. Information Theory, IEEE Transactions on 47.7 (2001): 3065-3072.

[3] Charles H. Bennett, Gilles Brassard, Claude Crépeau, Richard Jozsa, Asher Peres, and William K. Wootters. Teleporting an Unknown Quantum State via Dual Classical and Einstein-Podolsky-Rosen Channels. Physical Review Letters 70.13 (1993): 1895.

[4] Charles H. Bennett, Gilles Brassard, Sandu Popescu, Benjamin Schumacher, John A. Smolin, and William K. Wootters. Purification of Noisy Entanglement and Faithful Teleportation via Noisy Channels. Physical Review Letters 76.5 (1996): 722.

[5] Charles H. Bennett, David P. DiVincenzo, John A. Smolin, and William K. Wootters. Mixed State Entanglement and Quantum Error Correction. Physical Review A 54.5 (1996): 3824.

[6] Charles H. Bennett, Peter W. Shor, John A. Smolin, and Ashish V. Thapliyal. Entanglement-assisted classical capacity of noisy quantum channels. Physical Review Letters 83.15 (1999): 3081.

[7] Charles H. Bennett, Peter W. Shor, John A. Smolin, and Ashish V. Thapliyal. Entanglement-assisted capacity of a quantum channel and the reverse Shannon theorem. Information Theory, IEEE Transactions on 48.10 (2002): 2637-2655.

[8] Garry Bowen. Quantum Feedback Channels. Information Theory, IEEE Transactions on 50.10 (2004): 2429-2434.

[9] Zvika Brakerski, and Yael Tauman Kalai. Efficient Interactive Coding Against Adversarial Noise. Foundations of Computer Science (FOCS), 2012 IEEE 53rd Annual Symposium on (2012).

[10] Zvika Brakerski, and Moni Naor. Fast Algorithms for Interactive Coding. Proceedings of the 24th Annual ACM-SIAM Symposium on Discrete Algorithms (2013).
[11] Mark Braverman. *Coding for Interactive Computation: Progress and Challenges*. 50th Annual Allerton Conference on Communication, Control, and Computing (2012).

[12] Mark Braverman, and Anup Rao. *Towards Coding for Maximum Errors in Interactive Communication*. Proceedings of the 43rd Annual ACM symposium on Theory of Computing (2011).

[13] Harry Buhrman, Richard Cleve, and Avi Wigderson. *Quantum vs. classical communication and computation*. Proceedings of the 30th Annual ACM Symposium on Theory of Computing (1998).

[14] A. Robert Calderbank, Eric M. Rains, Peter W. Shor, and Neil. J. A. Sloane. *Quantum error correction via codes over GF(4)*. Information Theory, IEEE Transactions on 44.4 (1998): 1369-1387.

[15] A. Robert Calderbank, and Peter W. Shor. *Good Quantum Error-Correcting Codes Exist*. Physical Review A 54.2 (1996): 1098.

[16] Richard Cleve, and Harry Buhrman. *Substituting quantum entanglement for communication*. Physical Review A 56.2 (1997): 1201.

[17] Richard Cleve, Wim van Dam, Michael Nielsen, and Alain Tapp. *Quantum Entanglement and the Communication Complexity of the Inner Product Function*. Quantum Computing and Quantum Communications. Springer Berlin Heidelberg (1999): 61-74.

[18] Claude Crépeau, Daniel Gottesman, and Adam Smith. *Approximate Quantum Error-Correcting Codes and Secret Sharing Schemes*. Advances in Cryptology-EUROCRYPT 2005. Springer Berlin Heidelberg (2005): 285-301.

[19] Igor Devetak. *The Private Classical Capacity and Quantum Capacity of a Quantum Channel*. Information Theory, IEEE Transactions on 51.1 (2005): 44-55.

[20] Keqin Feng, and Zhi Ma. *A finite Gilbert-Varshamov bound for pure stabilizer quantum codes*. Information Theory, IEEE Transactions on 50.12 (2004): 3323-3325.

[21] Matthew Franklin, Ran Gelles, Rafail Ostrovsky, and Leonard Schulman. *Optimal Coding for Streaming Authentication and Interactive Communication*. Advances in Cryptology-CRYPTO 2013. Springer Berlin Heidelberg (2013): 1-20.

[22] Ran Gelles, Ankur Moitra, and Amit Sahai. *Efficient and Explicit Coding for Interactive Communication*. Foundations of Computer Science (FOCS), 2011 IEEE 52nd Annual Symposium on (2011).

[23] Daniel Gottesman. *A Class of Quantum Error-Correcting Codes Saturating the Quantum Hamming Bound*. Physical Review A 54.3 (1996): 1862.
[24] Hartmut Klauck, Ashwin Nayak, Amnon Ta-Shma, and David Zuckerman. *Interaction in quantum communication*. Information Theory, IEEE Transactions on 53.6 (2007): 1970-1982.

[25] Alexander S. Kholevo. *Towards the mathematical theory of quantum communication channels*. Probl. Peredachi Inform. 8 (1972): 63-71 (rus).

[26] Alexander S. Holevo. *Bounds for the Quantity of Information Transmitted by a Quantum Communication Channel*. Problems of Information Transmission 9 (1973): 177-183 (rus).

[27] Alexander S. Holevo. *The Capacity of the Quantum Channel with General Signal States*. Information Theory, IEEE Transactions on 44.1 (1998): 269-273.

[28] Aram W. Harrow. *Coherent Communication of Classical Messages*. Physical Review Letters 92.9 (2004): 097902.

[29] Gillat Kol, and Ran Raz. *Interactive Channel Capacity*. Proceedings of the 45th Annual ACM Symposium on Theory of Computing (2013).

[30] Ilan Kremer. *Quantum Communication*. Master’s thesis, Hebrew University, Computer Science Department (1995).

[31] Michael Langberg. *Private codes or succinct random codes that are (almost) perfect*. Foundations of Computer Science (FOCS), 2011 IEEE 45th Annual Symposium on (2004).

[32] Seth Lloyd. *Capacity of the Noisy Quantum Channel*. Physical Review A 55.3 (1997): 1613.

[33] Michael A. Nielsen, and Isaac L. Chuang. *Quantum Computation and Quantum Information*. Cambridge University Press (2000).

[34] Noam Nisan, and Avi Widgerson. *Rounds in communication complexity revisited*. Proceedings of the 23rd Annual ACM Symposium on Theory of computing (1991).

[35] Eric M. Rains. *Nonbinary quantum codes*. Information Theory, IEEE Transactions on 45.6 (1999): 1827-1832.

[36] Leonard J. Schulman. *Communication on Noisy Channels: A Coding Theorem for Computation*. Foundations of Computer Science (FOCS), 1992 IEEE 33rd Annual Symposium on (1992).

[37] Leonard J. Schulman. *Deterministic Coding for Interactive Communication*. Proceedings of the 25th Annual ACM Symposium on Theory of Computing (1993).

[38] Leonard J. Schulman. *Coding for Interactive Communication*. Information Theory, IEEE Transactions on 42.6 (1996): 1745-1756.
[39] Benjamin Schumacher, and Michael D. Westmoreland. *Sending Classical Information via Noisy Quantum Channels*. Physical Review A 56 (1997): 131-138.

[40] Peter W. Shor. *Scheme for reducing decoherence in quantum computer memory*. Physical review A 52.4 (1995): R2493.

[41] Peter W. Shor. *The Quantum Channel Capacity and Coherent Information*. Lecture Notes, MSRI Workshop on Quantum Computation (2002).

[42] Andrew Steane. *Multiple Particle Interference and Quantum Error Correction*. Proceedings of the Royal Society of London. Series A: Mathematical, Physical and Engineering Sciences 452.1954 (1996): 2551-2577.

[43] Ran Raz. *Exponential Separation of Quantum and Classical Communication Complexity*. Proceedings of the 31st Annual ACM Symposium on Theory of Computing (1999).

[44] John Watrous. *Theory of Quantum Information*. Lecture notes from Fall 2011, https://cs.uwaterloo.ca/ watrous/LectureNotes.html (2011).

[45] Mark M. Wilde. *Quantum Information Theory*. Cambridge University Press (2013), Preliminary version available as: arXiv e-print quant-ph/1106.1445.

[46] Andrew C.-C. Yao. *Quantum circuit complexity*. Foundations of Computer Science (FOCS), 1993 IEEE 34rd Annual Symposium on (1993).