Université de Montréal

Interactive Quantum Information Theory

par
Dave Touchette

Département d'informatique et de recherche opérationnelle
Faculté des arts et des sciences

Thèse présentée à la Faculté des études supérieures
en vue de l'obtention du grade de Philosophiæ Doctor (Ph.D.)
en informatique

Avril 2015

RÉSUMÉ

La théorie de l'information quantique s'est développée à une vitesse fulgurante au cours des vingt dernières années, avec des analogues et extensions des théorèmes de codage de source et de codage sur canal bruité pour la communication unidirectionnelle. Pour la communication interactive, un analogue quantique de la complexité de la communication a été développé, pour lequel les protocoles quantiques peuvent performer exponentiellement mieux que les meilleurs protocoles classiques pour certaines tâches classiques. Cependant, l'information quantique est beaucoup plus sensible au bruit que l'information classique. Il est donc impératif d'utiliser les ressources quantiques à leur plein potentiel.

Dans cette thèse, nous étudions les protocoles quantiques interactifs du point de vue de la théorie de l'information et étudions les analogues du codage de source et du codage sur canal bruité. Le cadre considéré est celui de la complexité de la communication : Alice et Bob veulent faire un calcul quantique biparti tout en minimisant la quantité de communication échangée, sans égard au coût des calculs locaux. Nos résultats sont séparés en trois chapitres distincts, qui sont organisés de sorte à ce que chacun puisse être lu indépendemment.

Étant donné le rôle central qu'elle occupe dans le contexte de la compression interactive, un chapitre est dédié à l'étude de la tâche de la redistribution d'état quantique. Nous prouvons des bornes inférieures sur les coûts de communication nécessaires dans un contexte interactif. Nous prouvons également des bornes atteignables avec un seul message, dans un contexte d'usage unique.

Dans un chapitre subséquent, nous définissons une nouvelle notion de complexité de l'information quantique. Celle-ci caractérise la quantité d'information, plutôt que de communication, qu'Alice et Bob doivent échanger pour calculer une tâche bipartie. Nous prouvons beaucoup de propriétés structurelles pour cette quantité, et nous lui donnons une interprétation opérationnelle en tant que complexité de la communication quantique amortie. Dans le cas particulier d'entrées classiques, nous donnons une autre caractérisation permettant de quantifier le coût encouru par un protocole quantique qui oublie de l'information classique. Deux applications sont présentées : le premier résultat général

de somme directe pour la complexité de la communication quantique à plus d'une ronde, ainsi qu'une borne optimale, à un terme polylogarithmique près, pour la complexité de la communication quantique avec un nombre de rondes limité pour la fonction « ensembles disjoints ».

Dans un chapitre final, nous initions l'étude de la capacité interactive quantique pour les canaux bruités. Étant donné que les techniques pour distribuer de l'intrication sont bien étudiées, nous nous concentrons sur un modèle avec intrication préalable parfaite et communication classique bruitée. Nous démontrons que dans le cadre plus ardu des erreurs adversarielles, nous pouvons tolérer un taux d'erreur maximal de $1/2 - \varepsilon$, avec $\varepsilon > 0$ arbitrairement petit, et ce avec un taux de communication positif. Il s'ensuit que les canaux avec bruit aléatoire ayant une capacité positive pour la transmission unidirectionnelle ont une capacité positive pour la communication interactive quantique.

Nous concluons avec une discussion de nos résultats et des directions futures pour ce programme de recherche sur une théorie de l'information quantique interactive.

Mots clés : Théorie des codes, compression, complexité de la communication, somme directe, ensembles disjoints, complexité de l'information, théorie de l'information à usage unique, calcul et information quantiques, redistribution d'états quantiques.

ABSTRACT

Quantum information theory has developed tremendously over the past two decades, with analogues and extensions of the source coding and channel coding theorems for unidirectional communication. Meanwhile, for interactive communication, a quantum analogue of communication complexity has been developed, for which quantum protocols can provide exponential savings over the best possible classical protocols for some classical tasks. However, quantum information is much more sensitive to noise than classical information. It is therefore essential to make the best use possible of quantum resources.

In this thesis, we take an information-theoretic point of view on interactive quantum protocols and study the interactive analogues of source compression and noisy channel coding. The setting we consider is that of quantum communication complexity: Alice and Bob want to perform some joint quantum computation while minimizing the required amount of communication. Local computation is deemed free. Our results are split into three distinct chapters, and these are organized in such a way that each can be read independently.

Given its central role in the context of interactive compression, we devote a chapter to the task of quantum state redistribution. In particular, we prove lower bounds on its communication cost that are robust in the context of interactive communication. We also prove one-shot, one-message achievability bounds.

In a subsequent chapter, we define a new, fully quantum notion of information cost for interactive protocols and a corresponding notion of information complexity for bipartite tasks. It characterizes how much quantum information, rather than quantum communication, Alice and Bob must exchange in order to implement a given bipartite task. We prove many structural properties for these quantities, and provide an operational interpretation for quantum information complexity as the amortized quantum communication complexity. In the special case of classical inputs, we provide an alternate characterization of information cost that provides an answer to the following question about quantum protocols: what is the cost of forgetting classical information? Two applications are pre-

sented: the first general multi-round direct-sum theorem for quantum protocols, and a tight lower bound, up to polylogarithmic terms, for the bounded-round quantum communication complexity of the disjointness function.

In a final chapter, we initiate the study of the interactive quantum capacity of noisy channels. Since techniques to distribute entanglement are well-studied, we focus on a model with perfect pre-shared entanglement and noisy classical communication. We show that even in the harder setting of adversarial errors, we can tolerate a provably maximal error rate of $1/2 - \varepsilon$, for an arbitrarily small $\varepsilon > 0$, at positive communication rates. It then follows that random noise channels with positive capacity for unidirectional transmission also have positive interactive quantum capacity.

We conclude with a discussion of our results and further research directions in interactive quantum information theory.

Keywords: Coding Theory, Compression, Communication Complexity, Direct Sum, Disjointness, Information Complexity, One-Shot Information Theory, Quantum Computation and Information, Quantum State Redistribution.

CONTENTS

LIST OF FIGURES

NOTATION

Quantum Theory; Section 2.1.

$A, B, C, \ldots$	Quantum systems, and associated Hilbert spaces.
$\lvert A\rvert$	Dimension of quantum system A.
AB	Tensor product space $A \otimes B$.
R	Usually reserved for a purifying quantum system.
X, Y, Z	Classical systems.
$\mathscr{D}(A)$	Set of density operators on quantum system A.
$\mathscr{H}(A)$	Set of pure states on quantum system A.
$\mathscr{C}(A, B)$	Set of quantum channels from $\mathscr{D}(A)$ to $\mathscr{D}(B)$.
$\mathscr{C}(A)$	$\mathscr{C}(A, A)$.
$\mathscr{U}(A, B)$	Set of isometries from $\mathscr{D}(A)$ to $\mathscr{D}(B)$.
$\mathscr{U}(A)$	Set of unitaries on $\mathscr{D}(A)$.
$\rho^A, \sigma^A, \theta^A, \cdots$	Density operators in $\mathscr{D}(A)$.
$\mathrm{Tr}_B(\rho^{AB})$	Partial trace over B system. The resulting state is denoted ρ^A.
$\mathrm{Tr}_{\neg A}(\rho^{AB})$	$\mathrm{Tr}_B(\rho^{AB})$.
$\lvert\psi\rangle^A, \lvert\phi\rangle^A, \cdots$	Pure states in $\mathscr{H}(A)$.
$\lvert\rho\rangle^{AR}$	Purification of ρ^A.
$\mathscr{N}^{A\to B}, \mathscr{M}^{A\to B}$	Channels in $\mathscr{C}(A, B)$.
$\mathscr{N}_2 \circ \mathscr{N}_1$	Composition of channels $\mathscr{N}_1$ and $\mathscr{N}_2$.
U, V	Isometries.
I^A	Identity on system A.
$U_{\mathscr{N}}^{A\to BE}$	Isometric extension of channel $\mathscr{N}^{A\to B}$.
$O^\dagger$	Adjoint of operator O.
Δ_B	Measurement channel on $\mathscr{D}(B)$.
$\mathscr{T}_\varepsilon$	Depolarizing channel of parameter ε.
X, Z	Pauli operators.

Quantum Information Theory; Section 2.2.

$\|O\|_1$	Trace norm of operator O.
$\|\rho^A - \sigma^A\|_1$	Trace distance between ρ^A and σ^A.
$\|\mathscr{N} - \mathscr{M}\|_\diamond$	Diamond norm between $\mathscr{N}$ and $\mathscr{M}$.
log	Base 2 logarithm.
$H(A)_\rho$	von Neumann entropy of ρ^A. If A is classical, Shannon entropy.
$H(p)$	Binary Shannon entropy, with $p \in [0,1]$.
$H(A\|B)_\rho$	Conditional entropy of ρ^{AB}.
$I(A;B)_\rho$	Mutual information of ρ^{AB}. If A is classical, Holevo information.
$I(A;B\|C)$	Conditional quantum mutual information of ρ^{ABC}.

One-Shot Quantum Information Theory; Section 3.2.

$\mathscr{P}(A)$	Set of positive semi-definite operators on A.
$\mathscr{D}_\leq(A)$	Set of sub-normalized states on A.
$F(\rho,\sigma)$	Fidelity between ρ and σ.
$\bar{F}(\rho,\sigma)$	Generalized fidelity between ρ and σ.
$P(\rho,\sigma)$	Purified distance between ρ and σ.
$\mathscr{B}^\varepsilon(\rho)$	ε-ball around ρ.
$D(\rho\|\sigma)$	Relative entropy of ρ with respect to σ
$D_{max}(\rho\|\sigma)$	Max-relative entropy of ρ with respect to σ.
$H_{min}(A\|B)_\rho$	Conditional min-entropy of ρ^{AB}.
$H_{max}(A\|B)_\rho$	Conditional max-entropy of ρ^{AB}.
$I_{max}(A;B)_\rho$	Max-information of ρ^{AB}.
$D^\varepsilon_{max}(\rho\|\sigma)$	Smooth max-relative entropy of ρ with respect to σ.
$H^\varepsilon_{min}(A\|B)_\rho$	Smooth Conditional min-entropy of ρ^{AB}.
$H^\varepsilon_{max}(A\|B)_\rho$	Smooth Conditional max-entropy of ρ^{AB}.
$I^\varepsilon_{max}(A;B)_\rho$	Smooth max-information of ρ^{AB}.
$\delta_{I(A;B)}(\rho,\sigma)$	$\|I(A;B)_\rho - I(A;B)_\sigma\|$.
$\mathscr{R}$	Quantum state redistribution channel, with implicit A,B,C systems.

	Quantum Communication Complexity and Models of Communication; Sections 2.3 and 2.4.
T	Classical relation.
X, Y, Z_A, Z_B	Input and output sets for classical relations.
$A_{in}, B_{in}, A_{out}, B_{out}$	Input and output registers for quantum tasks and protocols.
Π	Quantum protocol, usually in the hybrid model. Also used to represent the channel in $\mathscr{C}(A_{in}B_{in}, A_{out}B_{out})$ implemented by the protocol Π.
Π_ν	Quantum protocol, usually in the randomized model and with underlying distribution ν.
Π_{CB}	Quantum protocol in the Cleve-Buhrman model.
μ	Input distribution on $X \times Y$.
ρ_μ	Representation of μ as a quantum state in $\mathscr{D}(A_{in}B_{in})$. Often, μ is left implicit.
ρ	General input state.
$\Pi(x,y)$	Output, in $\mathscr{D}(A_{out}B_{out})$, of protocol Π on input (x,y).
$P_e(\Pi, \mu)$	Probability of error of Π on μ, for some implicit relation T.
(T, μ, ε)	Distributional classical task of implementing relation T on input distribution μ with average error at most ε.
$\mathscr{T}(T, \mu, \varepsilon)$	Set of all protocols in the hybrid model implementing (T, μ, ε).
$\mathscr{T}^M(T, \mu, \varepsilon)$	Restriction of $\mathscr{T}(T, \mu, \varepsilon)$ to M-message protocols.
(T, ε)	Worst-case classical task of implementing relation T with error at most ε on all inputs.
$\mathscr{T}(T, \varepsilon)$	Set of all protocols in the randomized model implementing (T, ε).
$\otimes_i (T_i, \mu_i, \varepsilon_i)$	Product task of implementing each of $(T_i, \mu_i, \varepsilon_i)$ in parallel.
$(T, \mu, \varepsilon)^{\otimes n}$	$\otimes_{i=1}^n (T, \mu, \varepsilon)$.

	Quantum Communication Complexity and Models of Communication; Continued.
$(\mathscr{N},\rho,\varepsilon)$	Quantum task of implementing channel $\mathscr{N}$ on input state ρ with error at most ε.
$\mathscr{T}(\mathscr{N},\rho,\varepsilon)$	Set of all protocol in the hybrid model implementing $(\mathscr{N},\rho,\varepsilon)$.
A_i, B_i, C_i	Registers of a protocol after message i has been sent.
A', B'	Leftover registers, beside A_{out}, B_{out}, at the end of a protocol.
T_A, T_B	Register holding the pre-shared entanglement ψ in a protocol.
$U_1, \cdots, U_{M+1}$	Unitaries defining a protocol.
$QCC(\Pi)$	Quantum communication cost of protocol Π.
$QCC_{A\to B}(\Pi)$	Quantum communication cost from Alice to Bob.
$QCC(T,\mu,\varepsilon)$	Quantum communication complexity of task (T,μ,ε).
$QCC^M(T,\mu,\varepsilon)$	M-message quantum communication complexity of (T,μ,ε).
$AQCC(T,\mu,\varepsilon)$	Amortized quantum communication complexity of (T,μ,ε)

	Quantum Information Complexity; Section 4.2.
$QIC(\Pi,\rho)$	Quantum information cost of hybrid protocol Π on input state ρ.
$QIC_R(\Pi_v,\rho)$	Quantum information cost of randomized protcol Π_v on input state ρ.
$QIC_{CB}(\Pi_{CB},\rho)$	Quantum information cost of protcol Π_{CB} in the Cleve-Buhrman model on input state ρ.
$QIC(T,\mu,\varepsilon)$	Quantum information complexity of task (T,μ,ε).
$QIC^M(T,\mu,\varepsilon)$	M-message quantum information complexity of task (T,μ,ε).
$QIC(T,\varepsilon)$	Quantum information complexity of task (T,ε).
$QIC_D(T,\varepsilon)$	Max-distributional quantum information complexity of task (T,ε).
$QIC(\otimes_i(T_i,\varepsilon_i))$	Quantum information complexity of product task $\otimes_i(T_i,\varepsilon_i)$.
$QIC_\times(\otimes_i(T_i,\varepsilon_i))$	Product quantum information complexity of product task $\otimes_i(T_i,\varepsilon_i)$.

Interactive Quantum Capacity; Sections 5.2 and 5.3.

s_A, s_B	History of Alice and Bob.
ℓ	Magnitude of error for a guess about an history.
$L(s, s^i)$	Equals ℓ for a guess s^i about history s.
Σ	Alphabet of a tree code.
α	Distance parameter of a tree code.
ε_α	$1-\alpha$.
$\mathcal{E}, \mathcal{D}$	Encoding and decoding functions for a tree code.
$\bar{\mathcal{E}}$	Extension of $\mathcal{E}$ to strings.
$\Delta(e_1, e_2)$	Hamming distance between e_1 and e_2.
$\mathcal{B}_i$	Encoding function for the ith message of a blueberry code.
$\mathcal{B}$	Concatenation of the $\mathcal{B}_i$'s.
Γ	Alphabet of a blueberrry code.
β	Erasure parameter of a blueberry code.
ε_β	$1-\beta$.
$\mathscr{F}_{q,N}$	Basis for operators acting on N systems of dimension q.
$\mathscr{E}_{\delta,q,N}$	Elements of $\mathscr{F}_{q,N}$ of weight at most δN.
$\lvert\psi_{init}\rangle$	Input state for the protocol to be simulated.
$\lvert\psi'_{init}\rangle$	Version of $\lvert\psi_{init}\rangle$ input to the simulation protocol.
$\lvert\psi_{final}\rangle$	Final state for the protocol to be simulated.
$\mathscr{M}^\Pi$	Quantum instrument $\mathscr{M}$ with black-box access to protocol Π.
Q	Simulation protocol in the quantum model.
S	Simulation protocol in the shared entanglement model.
$\mathcal{A}$	Adversary.
$\mathscr{A}$	Set of adversaries.
$Q^\Pi(\mathcal{A}(\lvert\psi_{init}\rangle))$	Output of simulator Q run against adversary $\mathcal{A}$ on input $\lvert\psi_{init}\rangle$.
$Q^\Pi(\mathcal{A})$	Channel implemented by running Q against $\mathcal{A}$.
$\mathscr{A}^{\mathrm{Q}}_{\delta,q,N'}$	Class of adversaries in quantum model with error rate bounded by δ.
$\mathscr{A}^{\mathrm{S}}_{\delta,q,N'}$	Analogue of $\mathscr{A}^{\mathrm{Q}}_{\delta,q,N'}$ in the shared entanglement model.

REMERCIEMENTS

J'ai eu la chance d'être supervisé par Gilles Brassard et Alain Tapp au cours des dernières années. J'ai beaucoup appris de par leur manière d'approcher les problèmes scientifiques et de par leur attitude face à la recherche en général. De plus, ils m'ont toujours soutenu et ont toujours été disponibles pour mes diverses requêtes, souvent faites avec seulement un court préavis de ma part, et je tiens donc à profiter de cette occasion pour les remercier pour leur appui au fils de ces années.

Je tiens également à remercier Louis Salvail, de qui j'ai pu apprendre autant, de par nos toujours agréables discussions quotidiennes, au sujet de la cryptographie, tant classique et quantique. J'aimerais remercier Pierre McKenzie de qui j'ai pu apprendre autant au sujet de l'informatique théorique. Je veux également les remercier, ainsi qu'Harry Buhrman, pour avoir accepté d'être sur le comité d'évaluation de ma thèse.

Également, je tiens à remercier Kamil Brádler, Patrick Hayden et Mark Wilde pour m'avoir transmis leur passion pour la théorie de l'information quantique. En particulier, Patrick qui a accepté de m'accorder de son temps pour me transmettre ses connaissances alors que je ne connaissais rien en informatique quantique. Également, Mark qui a toujours trouvé du temps, depuis toute ces années, pour répondre à mes multiples questions tant en théorie de l'information que sur divers autres sujets.

J'aimerais remercier mes collaborateurs, avec qui ce fut un réel plaisir de travailler et de qui j'ai beaucoup appris : Mario Berta, Gilles Brassard, Mark Braverman, Matthias Christandl, Ankit Garg, Young Kun Ko, Mathieu Laurière, Jieming Mao, Ashwin Nayak, Alain Tapp, Falk Unger.

Je tiens également à remercier Louis Salvail, Benno Salwey et Mark Wilde avec qui j'ai eu des discussions intéressantes au sujet de la communication interactive quantique bruitée, ainsi que Gilles Brassard, Omar Fawzi, Ashwin Nayak et Alain Tapp avec qui j'ai eu des discussions intéressantes au sujet de la complexité de l'information quantique.

Ce parcours au Laboratoire d'informatique théorique et quantique (LITQ) a pu être autant agréable grâce à l'ambiance conviviale qui y règne. Je remercie donc tous les collègues qui ont fait partie du LITQ pendant mon passage : Abdulrahman, Adrien, Alexan-

dre, Benno, Charles, Claude, Fabio, Francois, Heinz, Hichem, Hugo, Jürg, Kassem, Marc, Maxime, Michael et Michaël, Olivier, Paul, Patrick, Philippe, Samuel, Sara, Serge-Olivier, Rebecca, Xavier, Yara.

Finalement, je tiens tout particulièrement à remercier famille et amis pour leur appui, leur amour ainsi que leur compréhension lors de mes nombreuses périodes de “rush”, et tout specialement Alysa, Léo et Valérie pour ces dernières semaines ! Je leur dédie cette thèse.

Le travail dans cette thèse a été partiellement financé par le CRSNG, le FRQNT et CryptoWorks21.

CHAPTER 1

INTRODUCTION

1.1 Information Theory

One of the cornerstones of the 20th century was the development of information theory by Shannon [118]. In a single paper, Shannon laid the ground for a revolution in communication technology. His two quintessential theorems, the noiseless and noisy coding theorems, changed forever the way we approached compression and error correction codes. The noiseless coding theorem considers the setting in which a source emits messages from some set, each with some *a priori* probability. The theorem then states that a single quantity, depending only on these probabilities and not on the underlying messages, characterizes the optimal asymptotic rate for communicating many emissions of that source. This optimal rate is the entropy of the source. Based on the entropy, we can define further quantities with operational relevance, a prominent example being the mutual information between two random variables. Shannon's noisy coding theorem then characterizes the maximum rate at which it is possible to communicate over a noisy channel in terms of such a mutual information: the maximum over all possible input distributions of the mutual information between this input distribution and the induced distribution at the channel's output. These give neat characterizations of operational tasks in terms of simple quantities derived from Shannon's entropy. An excellent introduction to the field of information theory is Ref. [49].

1.2 Communication Complexity

In modern days, with the advent of fast internet communication, mobile phones, and multi-processor personal computers, new challenges emerge for communication in interactive settings. Indeed, a lot of modern communication is highly interactive, often arising as an integral part of computational processes. Communication complexity was introduced by Abelson [2] for computation over reals and then adapted to computation

over booleans by Yao [138] in order to study questions in distributed computation and circuit complexity. It has found applications in proving lower bounds in various models of computation. The basic setting is the following: Alice and Bob want to compute a joint function of their respective inputs while minimizing the communication they must exchange in order to do so. It is an idealized setting in which local computation is deemed free. It is one of the few models of computation for which it is possible to prove unconditional lower bounds. A typical example, and probably the one most studied, of a function in this setting is the disjointness function. Viewing Alice's and Bob's input as subsets of $\{1,2,\cdots,n-1,n\}$, this function asks whether these sets are disjoint. Its bounded error classical communication complexity is linear [86]. In general, interaction plays an important role in communication complexity: for any given number of rounds, there exist functions for which a single additional round of interaction enables an exponential saving in communication [106]. An excellent introduction to the field of communication complexity is Ref. [92].

1.3 Interactive Information Theory

The works of Shannon and his successors consider the case of unidirectional communication, and this paradigm is only well-motivated when we can afford to wait and communicate large blocks of data at once. In the highly interactive regime, in which we might want to run long protocols, but only once, we do not want to use compression and error correction codes designed to work in the unidirectional setting, since these will in general have parameters that are far from optimal in this interactive setting, and might impose an intolerable lag. Recent years have seen a flurry of results in classical interactive information and coding theory, for both compression [10, 33, 36, 38] and error correction [27, 35, 90, 114]. In particular, it has been shown that even under a constant rate of adversarial error, a constant factor overhead is sufficient to robustly implement interactive protocols [114]. Even more recently, some works have addressed the question of interactive channel capacity: the highest communication rate attainable over a given noisy channel for robustly implementing interactive protocols [90]. Also, compression

results are directly related to a generalization of Shannon's entropy to the interactive setting, termed information complexity. For a given interactive task, it has been shown that the information complexity corresponds to the amortized communication complexity of the task, i.e. the minimum amount of interactive communication per copy required for implementing such a task many times in parallel [33, 36]. Moreover, this information complexity paradigm can be used to show powerful communication complexity lower bounds. For example, it has recently been used to tightly characterize, up to second order in the input length, the communication complexity of the disjointness function [38]. Such a tight characterization seemed out of reach not so long ago, and witness the power of information-theoretic techniques to solve problems in communication complexity. Other problems for which information complexity has found many applications are the direct sum and direct product questions in communication complexity. The direct sum question asks if solving n copies of a task in parallel can be done more efficiently than by solving it n times sequentially, while the related direct product question asks whether the probability of successfully solving all n copies of a task in parallel, when given n times the resources required to solve it once, decays exponentially in n. An excellent introduction to the field of interactive information theory is Ref. [34].

1.4 Quantum Information Theory

As much as the work of Shannon led to what might be called the information age in the second half of the 20th century, it might well be that the 21st century will be the quantum age. Indeed, with the invention of quantum key distribution by Bennett and Brassard [13], promising unconditionally secure cryptography, and the discovery by Shor [120] of a quantum algorithm capable of breaking the computational hardness assumptions used in most modern day cryptographic systems, there has been a lot of interest and developments in recent years toward harnessing the power of quantum information processing. A quantum theory of information has been developed [18, 19, 53, 73, 99, 115, 116, 122], with many exciting results, in which the counterpart to Shannon's entropy is the von Neumann entropy. This quantum information theory is still

far less understood than its classical counterpart [56, 69], with such surprising effects as the superactivation of quantum capacity for channels that each have zero quantum capacity when used separately [123]. Recently, a lot of research effort has focused on so-called one-shot quantum information theory, in which we are interested in the amount of resources required to implement a single copy of a task (see Refs [111, 126] and reference therein). An excellent introduction to the field of quantum information theory is Ref. [133].

1.4.1 Quantum State Redistribution

One task in quantum information theory is particularly relevant in this thesis: quantum state redistribution. In quantum state redistribution as considered in Refs [54, 100, 140], there are 4 systems of interest: the A system held by Alice, the B system held by Bob, the C system that is to be transmitted from Alice to Bob, and the R system that holds a purification of the state in the ABC registers. The goal is to transmit the C register from Alice to Bob, up to some small overall error on the global state, while minimizing communication. It is the most general protocol for noiseless quantum coding. This protocol is also tightly linked to the notion of quantum information cost of interactive protocols, discussed in Section 1.6.1. While bounds on the asymptotic cost to perform this task have been known since the work of Devetak and Yard [54, 140], until recently no interesting bounds were known for the cost of performing state redistribution in a one-shot setting. We study this question in this thesis.

1.5 Quantum Communication Complexity

What happens in the setting of communication complexity if we allow Alice and Bob to exchange quantum rather than classical bits? Such a quantum model of communication complexity was defined by Yao [139], some fifteen years after the introduction of the classical model. He used it as a tool to prove lower bounds on quantum circuit complexity. A different model for quantum communication complexity was introduced by Cleve and Buhrman [47], in which we allow Alice and Bob to pre-share a large

entangled state of their choice, but in which, once the protocol starts, they can only exchange classical bits. They proved that their model is more powerful than Yao's classical model, at least by one classical bit. Due to teleportation [15], this model is at least as powerful as Yao's quantum model, up to a multiplicative factor of two, but it is an important open question to determine whether the two models are essentially equivalent. For both quantum models, it is known that there exists partial functions for which the quantum communication complexity can be exponentially lower than the classical communication complexity [42], even in the bounded error setting [109]. However, it is still unknown whether such a gap exists for a total function. Note that at least a polynomial gap exists in such a case, a typical example being the disjointness function for which the quantum communication complexity is $\Theta(\sqrt{n})$ [1, 42, 110], a quadratic improvement over the classical communication complexity. Concerning the power of interaction, a single more round of quantum interaction can also enable exponential savings [88]. An excellent introduction to the field of quantum communication complexity is Ref. [30].

1.6 Interactive Quantum Information Theory

With the flurry of important results in classical communication complexity due to interactive information theory, we would expect similar developments for an interactive theory of quantum information. However, until recently, it was unclear whether results similar to those for classical interactive information theory would hold in the quantum setting, where there is no direct analog of a protocol transcript. In particular, all classical error correction procedures for interactive protocols that achieve constant communication rate make heavy use of the transcript in multiple ways. Perhaps even worse, for noiseless coding, even the very definition of the classical information complexity is based on the underlying protocol transcript. Two recent results find ways around such problems, and constitute the core of this thesis.

1.6.1 Quantum Information Complexity

As said above, the definition of classical information complexity is directly related to the notion of the transcript of a protocol: a conditional mutual information between the transcript and the inputs. Since there is no direct analogue to the notion of a transcript for quantum protocols, it is not even clear *a priori* whether there exists a meaningful quantum generalization of this notion (see Open Problem 9 in Ref. [33]). Some previous attempts were made to define sensible notions of quantum information complexity [79, 81], with applications to quantum communication complexity lower bounds. In particular, Ref. [81] obtains a beautiful proof of the bounded round complexity of disjointness, however with a quadratic gap remaining for the round dependence compared with the best known upper bound [1]. An important drawback of these previous notions of quantum information complexity is that they have an explicit round dependence before yielding a communication lower bound, and as such they were probably limited to applications in a bounded round setting. The new notion of quantum information complexity avoids these problems by providing a novel interpretation of the classical information complexity, linking it to the task of noisy channel simulation with feedback and side information [19, 100]. A quantum generalization follows by making the link to the fully quantum analog of this task, which is equivalent to state redistribution [100, 134]. This definition is the first to apply to fully quantum inputs, and might find applications in studying quantum communication complexity of fully quantum tasks. It gets an operational interpretation as the amortized quantum communication complexity, i.e. the optimal rate of quantum communication for implementing many copies of a task, in the asymptotic limit. Moreover, it directly provides a lower bound on communication, independent of the number of rounds. It also provides an answer to the following important question about reversible quantum protocols: what is the quantum cost of forgetting classical information? This new notion of quantum information complexity has already found an application in proving the first general multi-round direct sum theorem in quantum communication complexity [129]. It has also been used to provide a tight lower bound (up to polylogarithmic terms) of the bounded round quantum com-

munication complexity of disjointness [40]. Both of these had been long standing open problems [81, 82].

1.6.2 Interactive Quantum Capacity

For highly interactive protocols, standard error correction techniques do not apply. We cannot wait for large blocks of data to accumulate before transmission, and if we do employ standard error correcting codes to transmit each message separately, then the code length must increase with the length of the protocol and rates of communication go to 0 asymptotically. The techniques developed classically for interactive coding do not seem to be applicable in the quantum setting due to the no-cloning theorem [55, 136], which forbids keeping a copy of previous states of the protocol to then go back to them if some quantum information is destroyed by noise. Moreover, in protocols in the Cleve-Buhrman model, even though the communication is classical, the irreversibility of quantum measurements performed to produce this classical information also seem to prevent the parties from backtracking to an earlier point in the protocol if noisy communication is detected after a measurement has been performed based on such erroneous information. Thus, for this problem also it is not even clear *a priori* that it is possible to simulate interactive quantum protocols with positive communication rate in the low noise regime, let alone in the very noisy regime.

By taking the approach to map every protocol into a reversible form, for example by using pseudo-measurements instead of actual measurements, it is possible to develop a representation for noisy quantum protocols that avoids these problems, and achieves positive communication rate while tolerating positive adversarial error rate [31]. A corollary is that for high random noise, any channel with positive capacity for quantum data transmission also has positive interactive quantum capacity. That is, even the hardest interactive quantum tasks can be implemented over noisy channels with only constant overhead (constant in the protocol length, not in the level of noise). This shows that quantum communication complexity is robust under most natural models of noisy quantum communication, another long standing open problem [104].

1.7 Outline of this Thesis

The aim of this thesis is to present in a coherent way a theory of quantum information for interactive protocols. The problems of noiseless compression and noisy channel coding for interactive quantum protocols are studied in separate chapters. Also, since it plays such an important role in the definition of quantum information cost of interactive protocols, we devote a separate chapter to the study of quantum state redistribution. Each of these three chapters contains mostly new contributions. These chapters are organized in such a way that each can be read independently.

In more detail, the next chapter provides the necessary preliminaries for the remainder of the thesis. After a brief introduction to quantum theory mainly to set the notation, we recall necessary notions from quantum information theory before providing precise definitions for the models of interactive quantum communication that we study.

In Chapter 3, quantum state redistribution is studied at length. A formal definition for the task is provided. Note that this is the only chapter requiring notions of one-shot quantum information theory, so the necessary notions are presented in it. The main results in that chapter are from a collaboration with Mario Berta and Matthias Christandl [23]. We give upper and lower bounds on the amount of quantum communication required to perform the task of quantum state redistribution in a one-shot setting. Our bounds are in terms of smooth conditional min- and max- entropies, and the smooth max-information. The protocol for the upper bound has a clear structure, building on the work of Oppenheim [107]: it decomposes the quantum state redistribution task into two simpler quantum state merging tasks by introducing a coherent relay. There remains a gap between our upper and lower bounds. This gap vanishes in the independent and identical (iid) asymptotic limit and the remaining terms can then be rewritten as a quantum conditional mutual information, thus yielding an alternative proof of optimality of this communication rate for iid asymptotic quantum state redistribution. Other new results of interest in this chapter are a proof that the lower bound on the communication required from Alice to Bob is robust in a multi-round scenario, along with a weaker upper bound on the one-shot communication cost in terms of conditional quantum mutual information.

This last bound is sufficient to prove the first general multi-round direct sum theorem for quantum communication complexity in Chapter 4, and is from Ref. [129]

In Chapter 4, we introduce new, fully quantum notions of quantum information cost and complexity. These are the first such notions to simultaneously be direct lower bound on communication while also being additive. They are the quantum analogues to the quantities that have found multiple applications recently in classical communication complexity. We prove many important properties for these quantities, and provide an operational interpretation for quantum information complexity as the amortized quantum communication complexity in a distributional setting. That is, the quantum information complexity of a task is exactly equal to the asymptotic quantum communication complexity per copy for this task. This is a result from Ref. [129]. By providing a general protocol compression result and using Yao's Min-Max theorem, we obtain the first general multi-round direct sum theorem for quantum communication complexity, a result from Ref. [129]. In the case of classical inputs, we also present an alternative characterization of quantum information cost that quantifies the cost for a protocol to forget classical information. This result is part of work in progress with Mathieu Laurière [94]. Finally, a high level overview is presented on another application of these new notions to provide a tight lower bound, up to polylogarithmic terms, on the bounded round quantum communication complexity of disjointness. This is joint work with Mark Braverman, Ankit Garg, Young Kun Ko and Jieming Mao [40]. Note that many of the properties of quantum information complexity proved in that chapter are due to this collaboration.

Finally, in Chapter 5, we initiate the study of the interactive quantum capacity of noisy channels. Most of the material in that chapter is from a collaboration with Gilles Brassard, Ashwin Nayak, Alain Tapp and Falk Unger [31]. We focus on the noisy analogue to the Cleve-Buhrman model of quantum communication complexity, with perfect pre-shared entanglement and noisy classical communication. We can then extend our results to other models of noisy communication, since techniques to distribute entanglement are well-studied. Our simulation strategy has a far higher communication rate than a naive one that encodes separately each particular round of communication to achieve

comparable success. Such a strategy would have a communication rate going to 0 in the worst interaction case as the length of the protocols increases. This contrasts with our strategy, which has a communication rate proportional to the capacity of the channel used. To avoid problems due to irreversibility of quantum measurements, we first have Alice and Bob purify all of their actions. Then, the idea that we exploit is to have Alice and Bob teleport a virtual communication register to each other, and keep track of errors in the communication of the teleportation measurement outcomes. We show how to use this technique along with ideas from classical interactive coding to tolerate a maximal fraction $1/2 - \varepsilon$, for an arbitrarily small $\varepsilon > 0$, of adversarial error while achieving strictly positive communication rates. This requires the development of new bounds on classical interactive codes. Note that in this model, the naive strategy would not work for any constant fraction of errors. A corollary is that in the very noisy regime, channels with non-zero capacity for unidirectional transmission have non-zero capacity for interactive communication. Thus, only constant overhead is required (constant in the amount of communication of protocols, not in the noise parameter), and quantum communication complexity is robust under noisy communication. Another finding that is perhaps surprising is that there exists quantum channels with zero forward capacity that can implement interactive communication without requiring assistance by pre-shared entanglement or a classical side-channel.

We conclude with a discussion of our results, and discuss further directions for this research program.

CHAPTER 2

PRELIMINARIES

In this chapter, we review the different notions of quantum theory, quantum information theory and quantum communication complexity that are required in the remainder of this thesis.

2.1 Quantum Theory

We 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 [105, 131, 133].

2.1.1 Quantum Systems

Let us first consider the case of general quantum systems; we restrict our study to finite dimensional systems. For every quantum system A, we associate a finite dimensional Hilbert space, which by a slight abuse of notation we also denote by A. The dimension of A is denoted $|A|$. The state of quantum system A is represented by a density operator ρ^A, a positive semi-definite operator over the Hilbert space A with unit trace. We denote by $\mathscr{D}(A)$ the set of all density operators representing states of system A. Composite quantum systems are associated with 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 $\mathscr{D}(A \otimes B)$. We use the shorthand AB for $A \otimes B$. Given a bipartite state $\rho^{AB} \in \mathscr{D}(AB)$, it is said to be separable if there exists a decomposition of the form $\rho^{AB} = \sum_i p(i) \cdot \sigma_i^A \otimes \theta_i^B$ for a probability distribution p and states $\sigma_i \in \mathscr{D}(A), \theta_i \in \mathscr{D}(B)$. If ρ^{AB} is not separable, it is said to be entangled. The evolution of a quantum system A is represented by a completely positive, trace preserving linear map (CPTP map) $\mathscr{N}^A$ such that if the state of the system was $\rho \in \mathscr{D}(A)$ before evolution through $\mathscr{N}^A$, the state of the system is $\mathscr{N}^A(\rho) \in \mathscr{D}(A)$ after. If the system A

is clear from context, we might drop the superscript. We refer to such maps as quantum channels, and to the set of all channels acting on A as $\mathscr{C}(A)$. An important quantum channel, which we consider in our study of noisy interactive quantum communication, is the qubit depolarizing channel $\mathscr{T}_\varepsilon$ with depolarizing parameter $0 \leq \varepsilon \leq 1$: it takes as input a qubit ρ, and outputs a qubit $\mathscr{T}_\varepsilon(\rho) = (1-\varepsilon)\rho + \varepsilon \frac{\mathrm{I}}{2}$, i.e., with probability $1-\varepsilon$ it outputs ρ and with complementary probability ε it outputs a completely mixed state $\frac{\mathrm{I}}{2}$, with I the identity operator.

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 $\mathscr{C}(A,B)$. An important operation on a composite system $A \otimes B$ is the partial trace $\mathrm{Tr}_B(\rho^{AB})$, which effectively derives the *reduced* or marginal state of the A subsystem from the quantum state ρ^{AB}. Fixing an orthonormal basis $\{|i\rangle\}$ for B, the partial trace is given by $\mathrm{Tr}_B(\rho^{AB}) = \sum_i (\mathrm{I} \otimes \langle i|)\rho(\mathrm{I} \otimes |i\rangle)$, and this is a valid quantum channel in $\mathscr{C}(A \otimes B, A)$. Note that the action of Tr_B is independent of the choice of basis chosen to represent it, so we unambiguously write $\rho^A = \mathrm{Tr}_B(\rho^{AB})$. We might also use the notation $\mathrm{Tr}_{\neg A} = \mathrm{Tr}_B$ to express that we want to keep only the A register. For channels $\mathscr{N}_1 \in \mathscr{C}(A,B), \mathscr{N}_2 \in \mathscr{C}(B,C)$, we denote their composition as $\mathscr{N}_2 \circ \mathscr{N}_1 \in \mathscr{C}(A,C)$, with action $(\mathscr{N}_2 \circ \mathscr{N}_1)(\rho) = \mathscr{N}_2(\mathscr{N}_1(\rho))$ on any state $\rho \in \mathscr{D}(A)$. We might drop the $\circ$ symbol if the composition is clear from context. An important subset of $\mathscr{C}(A)$ is the set of unitary channels $\mathscr{U}(A)$, the set of all maps $U \in \mathscr{C}(A)$ with an adjoint map $U^\dagger \in \mathscr{C}(A)$ such that $U^\dagger \circ U = U \circ U^\dagger = I^A$, with I^A the identity channel on A. More generally, if $|B| \geq |A|$, we denote by $\mathscr{U}(A,B)$ the set of isometric channels, i.e. the set of all maps $V \in \mathscr{C}(A,B)$ with an adjoint map $V^\dagger \in \mathscr{C}(B,A)$ such that $V^\dagger \circ V = I^A$.

An important special case for quantum states are the 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) upon which the density operator projects. We denote by $\mathscr{H}(A)$ the set of all such unit vectors (up to equivalence of global phase) in system A. For isometric spaces A, A', a maximally entangled state $\phi^{AA'}$ is a pure state satisfying $\mathrm{Tr}_{A'}(\phi^{AA'}) = \frac{\mathrm{I}^A}{|A|}$. A particularly important characteristic of pure states is the fact

that they do not share any correlations, classical or quantum, with any external system: if ρ^A is a pure state and σ^{AB} is any extension of it, i.e. any state such that $\mathrm{Tr}_B(\sigma^{AB}) = \rho^A$, then $\sigma^{AB} = \rho^A \otimes \sigma^B$. Pure state evolution is represented by an isometry $U^{A\to C}$ acting on $|\psi\rangle^A$, denoted $U|\psi\rangle^A \in \mathscr{H}(C)$. Evolution of the B register of a state $|\psi\rangle^{AB}$ under the action of an isometry $U^{B\to C}$ is represented by $(\mathrm{I}^A \otimes U^{B\to C})|\psi\rangle^{AB}$, for I^A representing the identity operator acting on the A system, and is denoted by the shorthand $U^{B\to C}|\psi\rangle^{AB}$ for convenience. We occasionally drop the superscripts when the systems are clear from context. The evolution under consecutive action of unitaries U_j's is denoted by:

$$\left(\prod_{j=1}^{\ell} U_j\right)|\psi\rangle = U_\ell \cdots U_1 |\psi\rangle . \tag{2.1.1}$$

For a state $\rho^A \in \mathscr{D}(A)$, a purification is a pure state $\rho^{AR} \in \mathscr{D}(A \otimes R)$ satisfying $\mathrm{Tr}_R(\rho^{AR}) = \rho^A$. We then say that R is a purifying system for ρ^A. If R has dimension at least that of A, then such a purification always exists. For a given R, all purifications are equivalent up to a unitary on R, and more generally, if $|R'| \geq |R|$ and $\rho_1^{AR}, \rho_2^{AR'}$ are two purifications of ρ^A, then there exists an isometry $V_\rho^{R\to R'}$ such that $\rho_2^{AR'} = V_\rho(\rho_1^{AR})$. We also consider purifications of channels: for a channel $\mathscr{N} \in C(A,B)$, an isometric extension is an isometry $U_{\mathscr{N}} \in \mathscr{U}(A, BE)$ with $\mathrm{Tr}_E(U_{\mathscr{N}}(\rho^A)) = \mathscr{N}(\rho^A)$ for all ρ^A. Such an extension always exists provided E is of dimension at least $|A|^2$.

2.1.2 Classical Systems

We now consider the special case of classical systems. We represent a classical random variable X with probability distribution p_X by a density operator σ^X that is diagonal in a fixed (orthonormal) basis $\{|x\rangle\}_{x\in\mathscr{X}}$: $\sigma^X = \sum_{x\in\mathscr{X}} p_X(x)|x\rangle\langle x|^X$. More generally, subsystem B of ρ^{BA} is said to be classical if we can write $\rho^{BA} = \sum_b p_B(b) \cdot |b\rangle\langle b|^B \otimes \rho_b^A$ for some $\rho_b^A \in \mathscr{D}(A)$ and an orthonormal basis $|b\rangle^B$. The state ρ^{BA} is then said to be a classical-quantum state. An important example of a channel mapping a quantum system to a classical one is the measurement channel Δ_B, defined as $\Delta_B(\rho) = \sum_b \langle b|\rho|b\rangle \cdot |b\rangle\langle b|^B$ for any $\rho \in \mathscr{D}(B)$. An isometric extension is given by $U_\Delta = \sum_b |b\rangle^{B'} |b\rangle^B \langle b|^B$.

Note that for any state $\rho \in \mathscr{D}(B_1B_2CR)$ of the form

$$|\rho\rangle^{B_1B_2CR} = \sum_b \sqrt{p_B(b)} \cdot |b\rangle^{B_1} |b\rangle^{B_2} |\rho_b\rangle^{CR}, \tag{2.1.2}$$

we have $\mathrm{Tr}_{B_2}(\rho^{B_1B_2CR}) = \sum_b p_B(b) \cdot |b\rangle\langle b|^{B_1} \otimes |\rho_b\rangle\langle\rho_b|^{CR}$ and $\mathrm{Tr}_{B_2R}(\rho^{B_1B_2CR}) = \sum_b p_B(b) \cdot |b\rangle\langle b|^{B_1} \otimes \rho_b^C$, with the state on B_1 classical in both cases. Often, $A, B, C, \ldots$ will be used to discuss general systems, while X, Y, Z will be reserved for classical systems, or quantum systems like B_1 and B_2 above that are classical once one of them is traced out, and can be thought of as containing a quantum copy of the classical content of one another. 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 quantum systems, quantum instruments can be viewed as a special case of quantum channels.

2.1.3 Teleportation and Pseudo-Measurements

Especially in the context of noisy interactive quantum communication, we make heavy use of the teleportation protocol between Alice and Bob [15], which uses the following resource state shared by Alice and Bob, called a Bell state: $|\Phi^+\rangle^{T_\mathsf{A}T_\mathsf{B}} = \frac{1}{\sqrt{2}}(|00\rangle + |11\rangle)$, 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 a Bell state, she applies a joint Bell measurement, which can perfectly distinguish the Bell states $\{|\Phi_{xz}\rangle = \frac{1}{\sqrt{2}}(|0x\rangle + (-1)^z |1\bar{x}\rangle)\}_{x,z\in\{0,1\}}$, 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 $\mathrm{X}^x\mathrm{Z}^z|\psi\rangle$, 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 $(\mathrm{X}^x\mathrm{Z}^z)^{-1} = \mathrm{Z}^z\mathrm{X}^x$. Note that X and Z anticommute: $\mathrm{XZ} = -\mathrm{ZX}$. 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). We call the coherent version of a measurement a *pseudo-measurement*. Without loss in generality, it suffices to consider the measurement of a single qubit in the standard basis $\{|0\rangle, |1\rangle\}$. This measurement corresponds to the instrument $\mathscr{N}$ defined by $\mathscr{N}(\rho) = \langle 0|\rho|0\rangle\ |0\rangle\langle 0| + \langle 1|\rho|1\rangle\ |1\rangle\langle 1|$. We replace this with the action of the CNOT operation $|0\rangle\langle 0| \otimes \mathrm{I} + |1\rangle\langle 1| \otimes \mathrm{X}$ on the qubit and a fresh ancillary qubit prepared in state $|0\rangle$, i.e., with the CPTP map $\mathscr{N}'$ defined by $\mathscr{N}'(\rho) = U(\rho \otimes |0\rangle\langle 0|)U^\dagger$, where U is the CNOT operation. The ancilla qubit may now be transmitted instead of sending the classical outcome of the measurement $\mathscr{N}$. Provided all further operations on the two qubits are only controlled unitary operations (in which the two qubits may only be control qubits), each separately behaves like the classical measurement outcome. The advantage of this substitution is that unlike measurements, pseudo-measurements are *reversible*. If it is later realized that a qubit should not have been measured, the pseudo-measurement can be undone.

2.2 Quantum Information Theory

We need to compare quantum states and channels throughout this thesis, thus we introduce a notion of distance for these. We are also interested in quantifying the information content of different quantum systems, so we also introduce different measures of information.

2.2.1 Distance Measures

In order to compare quantum states, the notion of distance we use is the trace distance $\|\rho^A - \sigma^A\|_1$ between two arbitrary states ρ^A and σ^A, in which

$$\|O^A\|_1 = \mathrm{Tr}((O^\dagger O)^{\frac{1}{2}}) \tag{2.2.1}$$

is the trace norm for operators on system A. We might drop the A superscript 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 ρ^A and σ^A, given a single unknown copy of one of these two states. We use the following results about trace distance. First, it is a metric, so it is symmetric in ρ, σ, non-negative, evaluate to 0 if and only if $\rho = \sigma$ and it satisfies the triangle inequality. Moreover, it is monotone under noisy channels: for any $\rho_1, \rho_2 \in \mathscr{D}(A)$ and $\mathscr{N} \in \mathscr{C}(A,B)$,

$$\|\mathscr{N}(\rho_1) - \mathscr{N}(\rho_2)\|_1 \leq \|\rho_1 - \rho_2\|_1. \tag{2.2.2}$$

For isometries, the inequality becomes an equality, a property called isometric invariance of the trace distance. Hence, for any ρ_1, $\rho_2 \in \mathscr{D}(A)$ and any $U \in \mathscr{U}(A,B)$, we have

$$\|U(\rho_1) - U(\rho_2)\|_1 = \|\rho_1 - \rho_2\|_1. \tag{2.2.3}$$

Also, the trace distance cannot change by adjoining an uncorrelated system: for any ρ_1, $\rho_2 \in \mathscr{D}(A)$, $\sigma \in \mathscr{D}(B)$

$$\|\rho_1 \otimes \sigma - \rho_2 \otimes \sigma\|_1 = \|\rho_1 - \rho_2\|_1. \tag{2.2.4}$$

The trace distance obeys a joint linearity property: for a classical system X and two states $\rho_1^{XA} = p_X(x) \cdot |x\rangle\langle x|^X \otimes \rho_{1,x}^A$ and $\rho_2^{XA} = p_X(x) \cdot |x\rangle\langle x|^X \otimes \rho_{2,x}^A$,

$$\|\rho_1 - \rho_2\|_1 = \sum_x p_X(x) \|\rho_{1,x} - \rho_{2,x}\|_1. \tag{2.2.5}$$

To distinguish between quantum channels with arbitrary input, we first consider the induced norm for quantum channels $\mathscr{N} \in \mathscr{C}(A,B)$: $\|\mathscr{N}\| = \max\{\|\mathscr{N}(\sigma)\|_1 : \sigma \in \mathscr{D}(A)\}$. Correlations with another quantum system can help distinguish between quantum channels, so an appropriate norm to use to account for this is the diamond norm [4]: $\|\mathscr{N}\|_\diamond = \|\mathscr{N} \otimes \mathrm{I}^R\|$ for some reference system R of the same dimension as the input system A. For two quantum channels $\mathscr{N}$, $\mathscr{M} \in \mathscr{C}(A,B)$, $\|\mathscr{N} - \mathscr{M}\|_\diamond$ has a useful operational interpre-

tation; it is (four times) the best possible bias with which we can identify a uniformly random (unknown) channel out of the two, when allowed only one use of the channel.

A further measure of distance between quantum states, the purified distance, is used in our discussion of one-shot quantum state redistribution in Chapter 3. It is based on the fidelity. Since both of these notions appear only in Chapter 3, they are defined in that chapter.

2.2.2 Information Measures

In Chapter 3 and especially in Chapter 4, we make use of the following information measures. The basic measure of information that we use is the von Neumann entropy, defined for any state $\rho \in \mathscr{D}(A)$ as

$$H(A)_\rho = -\operatorname{Tr}(\rho \log \rho),$$

in which we take the convention that $0 \log 0 = 0$, justified by a continuity argument. The logarithm log is taken in base 2. Note that H is invariant under isometries applied on ρ. If the state to be evaluated is clear from context, we might drop the subscript. Also note that if system A is classical, then we recover the Shannon entropy. Hence, for $p \in [0,1]$, we denote the binary Shannon entropy of p as $H(p)$. Conditional entropy for a state $\rho^{ABC} \in \mathscr{D}(ABC)$ is then defined as

$$H(A|B)_\rho = H(AB) - H(B),$$

mutual information as

$$I(A;B)_\rho = H(A) - H(A|B),$$

and conditional mutual information as

$$I(A;B|C)_\rho = H(A|C) - H(A|BC).$$

Note that mutual information and conditional mutual information are symmetric in interchange of A, B, and invariant under a local isometry applied to A, B or C. If X is a classical system, $I(X;B)$ is also called the Holevo information. For any pure bipartite state $\rho^{AB} \in \mathscr{D}(AB)$, the entropy on each subsystem is the same:

$$H(A) = H(B). \tag{2.2.6}$$

For any tripartite pure state $\rho^{ABC} \in \mathscr{D}(ABC)$, the conditional entropy satisfies a duality relation:

$$H(A|B) = -H(A|C). \tag{2.2.7}$$

For any four-partite pure state $\rho^{ABCR} \in \mathscr{D}(ABCR)$, the conditional mutual information satisfies the following relation:

$$I(C;R|B) = I(C;R|A). \tag{2.2.8}$$

Since all purifications are equivalent up to an isometry on the purification registers, we get that for any two pure states $|\phi\rangle^{ABCR'}$ and $|\psi\rangle^{ABCR}$ such that $\phi^{ABC} = \psi^{ABC}$,

$$I(C;R'|B)_{\phi} = I(C;R|B)_{\psi}. \tag{2.2.9}$$

For a system A and any $\rho \in \mathscr{D}(ABC)$, we have the bounds

$$0 \leq H(A) \leq \log|A|, \tag{2.2.10}$$

$$-H(A) \leq H(A|B) \leq H(A), \tag{2.2.11}$$

$$0 \leq I(A;B) \leq 2H(A), \tag{2.2.12}$$

$$0 \leq I(A;B|C) \leq H(A) + H(A|C). \tag{2.2.13}$$

If X is a classical system, we get the tighter bounds

$$0 \leq H(X|B), \tag{2.2.14}$$

$$0 \leq H(A|X), \tag{2.2.15}$$

$$I(X;B) \leq H(X), \tag{2.2.16}$$

$$I(X;B) \leq H(B), \tag{2.2.17}$$

$$I(X;B|C) \leq H(X|C). \tag{2.2.18}$$

For general quantum systems A, B, C and D, the conditional mutual information satisfies a chain rule: for any $\rho \in \mathscr{D}(ABCD)$,

$$I(AB;C|D) = I(A;C|D) + I(B;C|AD). \tag{2.2.19}$$

For any product state $\rho^{A_1B_1C_1A_2B_2C_2} = \rho_1^{A_1B_1C_1} \otimes \rho_2^{A_2B_2C_2}$, entropy is additive,

$$H(A_1A_2) = H(A_1) + H(A_2), \tag{2.2.20}$$

and so conditional mutual information between product systems vanishes,

$$I(A_1;A_2|B_1B_2) = 0, \tag{2.2.21}$$

and conditioning on a product system is useless,

$$I(A_1;B_1|C_1A_2) = I(A_1;B_1|C_1). \tag{2.2.22}$$

More generally,

$$I(A_1A_2;B_1B_2|C_1C_2) = I(A_1;B_1|C_1) + I(A_2;B_2|C_2). \tag{2.2.23}$$

Two important properties of the conditional mutual information are non-negativity and the data processing inequality, both equivalent to a deep result in quantum information

theory known as strong subadditivity [98]. For any $\rho \in \mathscr{D}(ABC)$ and $\mathscr{N} \in \mathscr{C}(B, B')$, denote $\sigma = \mathscr{N}(\rho)$ and then

$$I(A;B|C)_\rho \geq 0, \tag{2.2.24}$$

$$I(A;B|C)_\rho \geq I(A;B'|C)_\sigma. \tag{2.2.25}$$

It is also continuous [5]: for any two states ρ_1, $\rho_2 \in \mathscr{D}(ABC)$ with $\|\rho_1 - \rho_2\|_1 \leq \varepsilon \leq 1$, it holds that

$$|H(A|B)_{\rho_1} - H(A|B)_{\rho_2}| \leq 4\varepsilon \log \dim(A) + 2H(\varepsilon), \tag{2.2.26}$$

$$|I(A:B|C)_{\rho_1} - I(A:B|C)_{\rho_2}| \leq 8\varepsilon \log \dim(A) + 4H(\varepsilon), \tag{2.2.27}$$

in which $H(\varepsilon)$ is the binary entropy function. For classical systems, conditioning is equivalent to taking an average: for any $\rho^{ABCX} = \sum_x p_X(x) \cdot |x\rangle\langle x|^X \otimes \rho_x^{ABC}$ with a classical system X and some appropriate $\rho_x \in \mathscr{D}(ABC)$,

$$H(A|BX)_\rho = \sum_x p_X(x) \cdot H(A|B)_{\rho_x}, \tag{2.2.28}$$

$$I(A;B|CX)_\rho = \sum_x p_X(x) \cdot I(A;B|C)_{\rho_x}. \tag{2.2.29}$$

For proofs of the statements in this section, we refer the reader to Ref. [133]. Further measures of information, appropriate to study one-shot quantum information theory, are used in Chapter 3. Since these only appear in Chapter 3, they are defined in that chapter.

2.3 Quantum Communication Complexity

In the Yao model for quantum communication complexity [139], Alice is given a classical input $x \in X$, Bob a classical input $y \in Y$, and they want to compute a classical relation $T \subset X \times Y \times Z_A \times Z_B$, with $z_A \in Z_A$ and $z_B \in Z_B$ the output of Alice and Bob, respectively. These should satisfy the relation, i.e. $(x, y, z_A, z_B) \in T$. They want to do so by communicating as few quantum bits as possible, without regard to the local computation

cost. A special case of particular interest is that of computing a function $f : X \times Y \to Z$ of their joint input (often $X = Y = \{0,1\}^n$, $Z = \{0,1\}$). Often, we are only interested in $x \in X$, $y \in Y$ satisfying some promise $P : X \times Y \to \{0,1\}$. 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 Ref. [91]. Note that in this thesis, we are interested in more general models of communication, defined is Section 2.4.1, which is not reflected in the paragraph below.

A length M protocol in the Yao model is defined by a sequence of unitaries $U_1, \cdots, U_{M+1}$ 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 is not taken into account in the cost of the protocol. The complexity of the U_i's is also immaterial, since local computation is deemed free. However, the number of qubits c in the C register is important and is taken into account in the communication cost, which is $M \cdot c$. The outcome of the protocol is obtained by measuring an appropriate number of qubits of the registers A and B of Alice and Bob, respectively, after the application of U_{M+1}. The protocol succeeds if the outcomes of both measurements satisfy the relation, i.e. if the outcomes z_A, z_B satisfy $(x, y, z_A, z_B) \in T$. In the special case of a binary function f, it succeeds if the outcomes of both measurements equal $f(x,y)$ with good probability, usually required to be a constant greater than $1/2$, for any x, y satisfying the promise.

Another model for quantum communication complexity is the one introduced by Cleve and Buhrman [47]. In their model, communication is classical, but parties are allowed to pre-share an arbitrary entangled quantum state at the outset of the protocol. Note that a further variant on this model would be to allow for an unlimited supply of entanglement; we do not elaborate on such a variant in this thesis. We can view protocols in the Cleve-Buhrman 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 the Yao model (not including each party's classical input). In the context of noisy interactive quantum communication, our simulation protocols consider general unitary local processing but do not assume any particular form for the initial state. They then work on this slight adaptation of the Cleve-Buhrman model as well as for Yao's model of quantum communication complexity.

We will mostly work in the hybrid model, which allows both pre-shared entanglement and quantum communication. This model is almost equivalent to the Cleve-Buhrman model, at least in the case of a fixed order of communication. The fact that quantum communication is used instead of classical communication could lead to an improvement up to a factor of two of the communication complexity, due to superdense coding [14], but no more, due to the teleportation protocol [15]. In the context of quantum information complexity, this shared entanglement model is the natural analogue of the framework for classical information complexity in which parties are allowed shared randomness for free. It is thus natural to use this shared entanglement model rather than the model introduced by Yao [139], in which parties locally initialize their registers. In the context of noisy interactive communication, the hybrid model with quantum communication is better suited as a noiseless model due to the fact that protocols can be defined without irreversible measurements, and thus it is possible to backtrack to earlier points in the protocol if required. However, for simulation over noisy channel, the use of teleportation in a Cleve-Buhrman like model will arise as a natural model to overcome noise.

We describe precisely the models of communication that we consider in Section 2.4.1, but first let us describe what kind of tasks, classical and quantum, that we want to im-

plement with such protocols, by considering only their input-output behaviour for now, i.e. by viewing protocols as channels from their bipartite input to their bipartite output.

2.3.1 Classical Tasks

We consider the quantum communication complexity of two different kind of classical tasks: distributional tasks with fixed input distribution, and tasks with worst-case error. In both, for a given bipartite relation $T \subset X \times Y \times Z_A \times Z_B$, Alice and Bob are given input registers A_{in}, B_{in} containing their classical input $x \in X, y \in Y$ at the outset of the protocol, respectively, and they output registers A_{out}, B_{out} containing their classical output $z_A \in Z_A$, $z_B \in Z_B$ at the end of the protocol, respectively, which should satisfy the relation T. We generally allow for some small error ε in the output, which is formalized below.

2.3.1.1 Distributional Tasks

In the distributional communication complexity setting, the inputs x and y are distributed according to some joint input distribution μ on $X \times Y$. This is represented by a classical input state

$$\rho_\mu = \sum_{x \in X,\ y \in Y} \mu(x,y) \cdot |x\rangle\langle x|^{A_{in}} \otimes |y\rangle\langle y|^{B_{in}}, \tag{2.3.1}$$

in which we might drop the subscript μ if it is clear from context. Similarly the output of the protocol Π on each input (x,y) is a classical state

$$\Pi(x,y) = \sum_{z_A \in Z_A,\ z_B \in Z_B} p_{Z_A Z_B | X,Y}(z_A, z_B | x, y) \cdot |z_A\rangle\langle z_A|^{A_{out}} \otimes |z_B\rangle\langle z_B|^{B_{out}}, \tag{2.3.2}$$

and the error when implementing the relation T corresponds to the probability of failure $P_e(\Pi, \mu) = \sum_{x,y} \mu(x,y) \cdot \Pr_\Pi[\big(x, y, \Pi(x,y)\big) \notin T]$. Note that we require the quantum protocol to implement a classical channel, i.e. we ask that the channel implemented by the protocol be invariant under application of a measurement before and after the protocol, so that $\Pi = \Delta_{A_{out}B_{out}} \circ \Pi \circ \Delta_{A_{in}B_{in}}$.

To keep track of correlations in this distributional setting, we introduce a purification register R. For a classical input $\rho^{A_{in}B_{in}}$ like the one we consider here, we can always take this purification to be of the form

$$|\rho\rangle^{A_{in}B_{in}R} = \sum_{x,y} \sqrt{\mu(x,y)} \cdot |x\rangle^{A_{in}} |y\rangle^{B_{in}} |xy\rangle^{R_1} |xy\rangle^{R_2}, \tag{2.3.3}$$

for an appropriately chosen partition of R into R_1, R_2. We can think of $R = R_1R_2$ as containing quantum copies of the actual joint input to the protocol. Notice that if R_2 is traced out, then R_1 is classical and contains a copy of the joint input. It can then be used to compare input-output behaviour, and we can write more succinctly $P_e(\Pi,\mu) = \Pr_{\mu,\Pi}[\Pi(\rho^{A_{in}B_{in}R_1}) \in T]$. We say that a protocol Π for implementing relation T on input μ has average error at most $\varepsilon \in [0,1]$ if $P_e(\Pi,\mu) \leq \varepsilon$. Note that the idea to purify classical inputs has already appeared in previous works in quantum complexity theory, e.g. Ref. [85].

We are mostly interested in protocols in the hybrid model, as defined in Section 2.4.1.1, to solve distributional classical tasks. We say that a protocol implements the task (T,μ,ε) if it implements relation T on input distribution μ with error at most ε. The set of all protocols in the hybrid model implementing (T,μ,ε) is denoted $\mathscr{T}(T,\mu,\varepsilon)$. When restricting this set to bounded round protocols with at most M messages, this is denoted $\mathscr{T}^M(T,\mu,\varepsilon)$.

2.3.1.2 Worst-Case Tasks

We also consider tasks for worst-case inputs: the task (T,ε) is similar to the task (T,μ,ε), but instead of requiring average error ε with respect to the input distribution μ, we require that for all inputs $(x,y) \in X \times Y$ satisfying some promise $P : X \times Y \to \{0,1\}$, the error is bounded by ε, i.e. $P_e(\Pi,(x,y)) \leq \varepsilon$ for each pair (x,y) satisfying $P(x,y) = 1$, with

$$P_e(\Pi,(x,y)) = \Pr_{\Pi}[\big(x,y,\Pi(x,y)\big) \notin T]. \tag{2.3.4}$$

From now on, we will leave implicit the fact that we only consider pairs (x,y) satisfying some underlying promise. For technical reasons, for tasks with worst-case inputs, we allow Alice and Bob to jointly sample which protocol they want to use. The corresponding model of communication is called the randomized model, and is defined in Section 2.4.1.2. For this, they will share a finite random string r distributed according to a probability distribution ν, and sample a protocol Π_r according to ν. Such a randomized protocol Π_ν will implement the average channel

$$\Pi_\nu = \sum_r \nu(r)\Pi_r. \tag{2.3.5}$$

Correspondingly, the average error on any (x,y) is

$$P_e(\Pi_\nu,(x,y)) = \sum_r \nu(r)\mathrm{Pr}_{\Pi_r}[(x,y,\Pi_r(x,y)) \notin T]. \tag{2.3.6}$$

The worst case error of a protocol Π_ν is $P_e^w(\Pi_\nu) = \max_{(x,y)} P_e(\Pi_\nu,(x,y))$. We say that a protocol for implementing relation T has worst-case error at most $\varepsilon \in [0,1]$ if $P_e^w(\Pi_\nu) \leq \varepsilon$.

As said above, we are mostly interested in the randomized model, as defined in Section 2.4.1.2, to solve classical tasks with worst-case error. We say that a protocol implements the task (T,ε) if it implements relation T with worst-case error at most ε. The set of all protocols in the randomized model implementing (T,ε) is denoted $\mathscr{T}(T,\varepsilon)$. When restricting this set to bounded round protocols with at most M messages, this is denoted $\mathscr{T}^M(T,\varepsilon)$.

Note that allowing for such randomized protocols would be unnecessary in the distributional setting. To see this, suppose that for some distribution ν on protocols and μ on inputs, the probability of error satisfies $\sum_{x,y,r} \mu(x,y)\cdot\nu(r)\mathrm{Pr}_{\Pi_r}[(x,y,\Pi_r(x,y)) \notin T] = \varepsilon$. Then there must exists some r^* such that $\sum_{x,y} \mu(x,y)\mathrm{Pr}_{\Pi_{r^*}}[(x,y,\Pi_{r^*}(x,y)) \notin T] \leq \varepsilon$, and it is sufficient to pick the corresponding protocol Π_{r^*} to achieve good average error under this particular μ.

2.3.2 Quantum Tasks

We also want to study in full generality the quantum communication complexity of bipartite quantum channels on fixed input states. This is the generalization of distributional communication complexity of classical functions to the fully quantum setting, and contains as a special case the distributional quantum communication complexity of classical functions as well as that of implementing bipartite unitary transformations. An example of such a quantum task that was studied in Ref. [104] is the communication complexity for implementing the distributed quantum Fourier transform. Another example that will be of central importance in the study of quantum information complexity in Chapter 4 is the communication task of quantum state redistribution, which is discussed at length in Chapter 3.

The model for communication complexity of quantum tasks that we consider is the following. For a given bipartite channel $\mathcal{N} \in \mathscr{C}(A_{in}B_{in}, A_{out}B_{out})$ and input state $\rho \in \mathscr{D}(A_{in}B_{in})$, Alice and Bob are given input registers A_{in}, B_{in} at the outset of the protocol, respectively, and they output registers A_{out}, B_{out} at the end of the protocol, respectively, which should be in state $\mathcal{N}(\rho)$. We generally allow some small error ε in the output, which is formalized as follows. A protocol Π is said to implement channel $\mathcal{N}$ on input $\rho^{A_{in}B_{in}}$ with error at most $\varepsilon \in [0,2]$ if, for a purifying register R,

$$||\Pi(\rho^{A_{out}B_{out}R}) - \mathcal{N}(\rho^{A_{out}B_{out}R})||_1 \leq \varepsilon. \tag{2.3.7}$$

The introduction of the reference system R is essential to ensure that the protocol preserves any correlation the input state might have with any external systems as well as the channel it is supposed to implement.

We are mostly interested in protocols in the hybrid model to solve such quantum tasks. We say that a protocol implements the task $(\mathcal{N}, \rho, \varepsilon)$ if it implements channel $\mathcal{N}$ on input ρ with error at most ε. The set of all protocols in the hybrid model implementing $(\mathcal{N}, \rho, \varepsilon)$ is denoted $\mathscr{T}(\mathcal{N}, \rho, \varepsilon)$. When restricting this set to bounded round protocols with at most M messages, this is denoted $\mathscr{T}^M(\mathcal{N}, \rho, \varepsilon)$. Note that it will always be clear from context whether we are discussing classical or quantum tasks.

2.4 Interactive Quantum Protocols

We now formally introduce the different models of communication that are used throughout this thesis, along with corresponding notions of quantum communication complexity for classical and quantum tasks.

2.4.1 Models of Communication

We consider three different models of quantum communication: the hybrid model, the randomized model, and the Cleve-Buhrman model.

2.4.1.1 Hybrid Model

Most of our discussions will concern the hybrid model with quantum communication and pre-shared entanglement. In this model, an M-message protocol Π for a given task from input registers A_{in}, B_{in} to output registers A_{out}, B_{out} is defined by a sequence of isometries $U_1, \cdots, U_{M+1}$ along with a pure state $\psi \in \mathscr{D}(T_A T_B)$ shared between Alice and Bob, for arbitrary finite dimensional registers T_A, T_B: the pre-shared entanglement. We need $M+1$ isometries in order to have M messages since a first isometry is applied before the first message is sent and a last one after the final message is received. In the case of even M, for appropriate finite dimensional quantum memory registers A_1, $A_3, \cdots, A_{M-1}, A'$ held by Alice, $B_2, B_4, \cdots, B_{M-2}, B'$ held by Bob, and quantum communication registers $C_1, C_2, C_3, \cdots, C_M$ exchanged by Alice and Bob, we have $U_1 \in \mathscr{U}(A_{in}T_A, A_1C_1)$, $U_2 \in \mathscr{U}(B_{in}T_BC_1, B_2C_2)$, $U_3 \in \mathscr{U}(A_1C_2, A_3C_3)$, $U_4 \in \mathscr{U}(B_2C_3, B_4C_4)$, $\cdots, U_M \in \mathscr{U}(B_{M-2}C_{M-1}, B_{out}B'C_M)$, $U_{M+1} \in \mathscr{U}(A_{M-1}C_M, A_{out}A')$: see Figure 2.1. We adopt the convention that, at the outset, $A_0 = A_{in}T_A$, $B_0 = B_{in}T_B$, for odd i with $1 \leq i < M$, $B_i = B_{i-1}$, for even i with $1 < i \leq M$, $A_i = A_{i-1}$ and also $B_M = B_{M+1} = B_{out}B'$, and $A_{M+1} = A_{out}A'$. In this way, after application of U_i, Alice holds register A_i, Bob holds register B_i and the communication register is C_i. In the case of an odd number of message M, the registers corresponding to U_M, U_{M+1} are changed accordingly. We slightly abuse notation and also write Π to denote the channel in $\mathscr{C}(A_{in}B_{in}, A_{out}B_{out})$ implemented by

the protocol, i.e. for any $\rho \in \mathscr{D}(A_{in}B_{in})$,

$$\Pi(\rho) = \mathrm{Tr}_{A'B'}(U_{M+1}U_M \cdots U_2U_1(\rho \otimes \psi)). \tag{2.4.1}$$

Note that the A' and B' registers are the final memory registers that are being discarded at the end of the protocol by Alice and Bob, respectively.

We have the following definition.

Definition 2.4.1. *For a protocol* Π *defined as above, we define the* quantum communication cost *of* Π *as*

$$QCC(\Pi) = \sum_i \log |C_i|.$$

Note that in general we do not require that $|C_i| = 2^k$ for some $k \in \mathbb{N}$, as is often done. This will not affect our definition on information cost and complexity, but might affect the quantum communication complexity by at most a factor of two, without affecting the round complexity.

We sometimes distinguish between the communication from Alice to Bob and vice versa. We then use the following definitions.

Definition 2.4.2. *For a protocol* Π *defined as above, we define the* quantum communication cost from Alice to Bob *of* Π *as*

$$QCC_{A\to B}(\Pi) = \sum_{i \text{ odd}} \log |C_i|,$$

and the quantum communication cost from Bob to Alice *of* Π *as*

$$QCC_{B\to A}(\Pi) = \sum_{i \text{ even}} \log |C_i|.$$

2.4.1.2 Randomized Model

We use a further variation of the hybrid model in which Alice and Bob are allowed to pre-share randomness to jointly sample which protocol they use. Importantly, the order

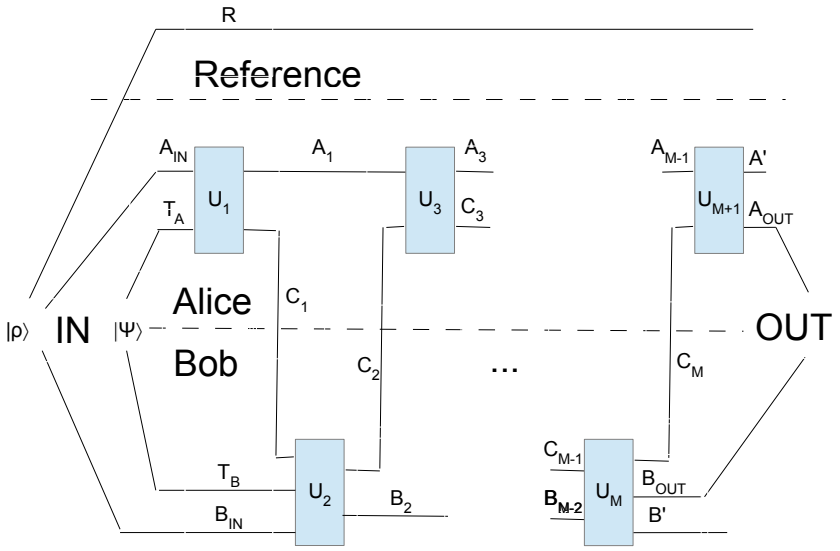

Figure 2.1: A protocol in the hybrid quantum communication model

of speech can depend on this randomness. This will be used in order to solve tasks on worst case inputs. If they sample according to some distribution ν with finite support, we denote the average protocol as Π_ν. The channel implemented by Π_ν is

$$\Pi_\nu = \sum_r \nu(r) \cdot \Pi_r, \tag{2.4.2}$$

its quantum communication cost is defined as

$$QCC(\Pi_\nu) = \max_{r:\nu(r)>0} QCC(\Pi_r), \tag{2.4.3}$$

and the number of message M is defined as the maximum over the number of messages M_r in any protocol Π_r with $\nu(r) > 0$. The choice to have worst-case instead of average-case countability over the shared randomness for the communication cost of randomized protocols is made in accordance with the standard corresponding classical definitions; see Ref. [92] for a related discussion of why this does not have a substantial qualitative

effect for a wide variety of tasks. As said above, we allow the order of speech to depend on this public sampling. This is in order to allow for a bounded-round quantum analogue of Yao's Min-Max theorem without having to place a restriction on the communication pattern for the bounded-round protocols; see Section 4.6.3.1. Note that in the special case in which all protocols considered have the same communication pattern, that is, if the communication registers C_i of all protocols have the same dimensions, then we could implement the corresponding randomized protocol with an equivalent one in the hybrid model by using shared entanglement to simulate the shared randomness. We will see that in the context of quantum information complexity, we can implement any randomized protocol in the hybrid model, without such a restriction on the communication patterns; see Section 4.3.1.4 for a precise statement.

2.4.1.3 Cleve-Buhrman Model

One last model that we also consider is the one introduced by Cleve and Buhrman [47], in which Alice and Bob pre-share entanglement but can only communicate classically once the protocol starts. In this model, an M-message protocol Π for a given task from input registers $A_{in}B_{in}$ to output registers $A_{out}B_{out}$ is defined by a sequence of quantum instruments $\mathscr{M}_1, \cdots, \mathscr{M}_{M+1}$ along with a pure state $\psi \in \mathscr{D}(T_A T_B)$ shared between Alice and Bob, for arbitrary finite dimensional registers T_A, T_B: the pre-shared entanglement. In the case of even M, for appropriate finite dimensional quantum memory registers A_1, A_3, $\cdots$, A_{M-1}, A' held by Alice, B_2, B_4, $\cdots$, B_{M-2}, B' held by Bob, classical memory registers M_1^A, M_3^A, $\cdots$, M_{M-3}, M_{M-1}^A, M_M^A held by Alice, M_2^B, M_4^B, $\cdots$, M_{M-2}^B, M_M^B held by Bob, and classical communication registers C_1, C_2, C_3, $\cdots$, C_M exchanged by Alice and Bob, we have $\mathscr{M}_1 \in \mathscr{C}(A_{in}T_A, A_1 M_1^A C_1)$, $\mathscr{M}_2 \in \mathscr{C}(B_{in}T_B C_1, B_2 M_2^B C_2)$, $\mathscr{M}_3 \in \mathscr{C}(A_1 M_1^A C_2, A_3 M_3^A C_3)$, $\mathscr{M}_4 \in \mathscr{C}(B_2 M_2^B C_3, B_4 M_4^B C_4)$, $\cdots$, $\mathscr{M}_M \in \mathscr{C}(B_{M-2} M_{M-2}^B C_{M-1}, B_{out} B' M_M^B C_M)$, $\mathscr{M}_{M+1} \in \mathscr{C}(A_{M-1} M_{M-1}^A C_M, A_{out} A' M_M^A)$, in which the registers M_j^A, M_j^B contain classical copies of the classical messages C_1 through C_j. In the case of an odd number of message M, the registers corresponding to $\mathscr{M}_M, \mathscr{M}_{M+1}$ are changed accordingly.

2.4.2 Classical Tasks

We consider the quantum communication complexity of many different classical tasks: distributional tasks, worst-case tasks, and product tasks, both distributional and worst-case.

2.4.2.1 Distributional Tasks

Recall that in the distributional setting, a classical task (T, μ, ε) consists of implementing a relation $T \subset X \times Y \times Z_A \times Z_B$ on input distribution μ on $X \times Y$ with average error at most ε, and we denote the set of all protocols in the hybrid model implementing the classical task (T, μ, ε) as $\mathscr{T}(T, \mu, \varepsilon)$. If we restrict this set to bounded-round protocols with at most M messages, we denote it as $\mathscr{T}^M(T, \mu, \varepsilon)$. We use the following definitions.

Definition 2.4.3. *For a classical task (T, μ, ε) and a bound $M \in \mathbb{N}$ on the number of messages, we define the ε-error* quantum communication complexity *of T on input μ as*

$$QCC(T, \mu, \varepsilon) = \min_{\Pi \in \mathscr{T}(T, \mu, \varepsilon)} QCC(\Pi),$$

and the M-message, ε-error quantum communication complexity *of T on input μ as*

$$QCC^M(T, \mu, \varepsilon) = \min_{\Pi \in \mathscr{T}^M(T, \mu, \varepsilon)} QCC(\Pi).$$

Note that these quantities are in general discontinuous in their parameters μ and ε for a fixed relation T. Also note that no good bound is known on the size of the entangled state that might be required to achieve these minima. See Ref. [101] for a recent discussion on related issues in a different setting. We make the trivial remarks that quantum communication complexity decreases as the error parameter increases, that it vanishes for $\varepsilon = 1$, that it is bounded by $\log |A_{in}| + \log |A_{out}|$, and that it also vanishes for any fixed input (x, y). At $\varepsilon = 1$, it is because the error is saturated and so we can consider a protocol that outputs anything without communication, while for fixed input

it is because, in this distributional setting, if the input is known then a corresponding output also is. In the special case of a single message, additional restrictions on A_{out} might be needed in order for the quantum communication complexity to be well-defined. Also, the quantum communication complexity without any restriction on the number of messages is at most as large as the bounded round quantum communication complexity for any bound M on the number of messages.

Remark 2.4.1. *For any T, μ, ε, ε_1, ε_2, M, with $0 \leq \varepsilon_1 \leq \varepsilon_2 \leq 1$, the following holds:*

$$\begin{aligned} QCC(T,\mu,\varepsilon) &\leq QCC^M(T,\mu,\varepsilon), \\ QCC(T,\mu,\varepsilon_2) &\leq QCC(T,\mu,\varepsilon_1), \\ QCC^M(T,\mu,\varepsilon_2) &\leq QCC^M(T,\mu,\varepsilon_1), \\ QCC^M(T,\mu,1) &= 0. \end{aligned}$$

If $M \geq 2$,

$$QCC^M(T,\mu,0) \leq \log|A_{in}| + \log|A_{out}|.$$

Also, for any $\varepsilon \in [0,1]$, the following holds for any distribution μ with support of size one:

$$QCC^M(T,\mu,\varepsilon) = 0.$$

2.4.2.2 Worst-Case Tasks

To solve a classical task with worst-case error (T,ε), which consists of implementing a relation $T \subset X \times Y \times Z_A \times Z_B$ with worst-case error at most ε, we also allow Alice and Bob to jointly sample which protocol they use, i.e. they can use protocols in the randomized model. We denote by $\mathscr{T}(T,\varepsilon)$ the set of all randomized protocols implementing task (T,ε), and by $\mathscr{T}^M(T,\varepsilon)$ if we restrict this set to M-message protocols. We get the following definitions.

Definition 2.4.4. *For a classical task (T,ε) and a bound $M \in \mathbb{N}$ on the number of messages, we define the ε-error* quantum communication complexity *of T as*

$$QCC(T,\varepsilon) = \min_{\Pi \in \mathscr{T}(T,\varepsilon)} QCC(\Pi),$$

and the M-message, ε-error quantum communication complexity *of T as*

$$QCC^M(T,\varepsilon) = \min_{\Pi \in \mathscr{T}^M(T,\varepsilon)} QCC(\Pi).$$

We denote by $\mathscr{D}_{XY}$ the set of all distributions over $X \times Y$. Note that it follows from discussions in Section 2.3.1 and the above definitions that for any relation T, error ε and number of message M, the following holds:

$$\max_{\mu \in \mathscr{D}_{XY}} QCC(T,\mu,\varepsilon) \leq QCC(T,\varepsilon), \tag{2.4.4}$$

$$\max_{\mu \in \mathscr{D}_{XY}} QCC^M(T,\mu,\varepsilon) \leq QCC^M(T,\varepsilon). \tag{2.4.5}$$

The reason why we allow Alice and Bob to use protocols in the randomized model to solve tasks with worst-case error is to be able to prove the reverse inequalities (possibly up to a discontinuity in the error parameter). See Section 4.6.3.1, which gives a quantum analogue to Yao's Min-Max theorem. We also get similar remarks as in the distributional case.

Remark 2.4.2. *For any T, ε, ε_1, ε_2, M, with $0 \leq \varepsilon_1 \leq \varepsilon_2 \leq 1$, the following holds:*

$$\begin{aligned} QCC(T,\varepsilon) &\leq QCC^M(T,\varepsilon), \\ QCC(T,\varepsilon_2) &\leq QCC(T,\varepsilon_1), \\ QCC^M(T,\varepsilon_2) &\leq QCC^M(T,\varepsilon_1), \\ QCC^M(T,1) &= 0. \end{aligned}$$

If $M \geq 2$,

$$QCC^M(T,0) \leq \log|A_{in}| + \log|A_{out}|.$$

2.4.2.3 Product Tasks

We are also interested in the quantum communication complexity of implementing multiple classical tasks in parallel. A protocol Π_n is said to compute the n-fold product relation $T_1 \otimes T_2 \otimes \cdots \otimes T_n$ on input $\mu^n = \mu_1 \times \mu_2 \times \cdots \times \mu_n$, each with corresponding error ε_i if, for each $i \in [n]$, the probability of failure for task i is at most ε_i. That is, if we denote by (x_i, y_i) the ith coordinate of the input and by Π_n^i the ith coordinate of the output of Π_n, corresponding to T_i, then it holds that $P_e^i(\Pi_n, \mu^n) \leq \varepsilon_i$ for each i, with $P_e^i(\Pi_n, \mu^n) = \sum_{x^n,y^n} \mu^n(x^n,y^n) \cdot \Pr_{\Pi_n}[\big(x_i, y_i, \Pi_n^i(x^n,y^n)\big) \notin T_i]$. This error criterion corresponds to the one achieved when sequentially implementing the n tasks $(T_i, \mu_i, \varepsilon_i)$ and, even for the case of $\varepsilon_i = \varepsilon$ for each i, this is weaker than demanding to simulate them with overall error ε. Indeed, asking for overall error ε could correspond to a much harder task. In particular, the direct product question, considering the particular case $T_i = T$ for each i, asks if this overall error goes to 1 exponentially fast in n if we do not allow sufficiently more resources than for sequential implementation. Hence, if we want to study amortized communication complexity with nontrivial error, we have to settle for such a success parameter. We call $\otimes_i(T_i, \mu_i, \varepsilon_i)$ a product classical task, and denote $\mathscr{T}\big(\otimes_i (T_i, \mu_i, \varepsilon_i)\big)$ the set of all protocols achieving the above goal of having error ε_i for each corresponding task. When restricting this set to M-message protocols, we denote it as $\mathscr{T}^M\big(\otimes_i (T_i, \mu_i, \varepsilon_i)\big)$. We also specialize the definition to the special case in which we are interested in implementing n times the same task, and in such a case we simply write $(T, \mu, \varepsilon)^{\otimes n}$ for $\otimes_{i=1}^n (T, \mu, \varepsilon)$. We have the following definitions.

Definition 2.4.5. *For a product classical task* $\otimes_i(T_i, \mu_i, \varepsilon_i)$ *and a bound* $M \in \mathbb{N}$ *on the number of messages, we define the* n-fold quantum communication complexity *of*

$\otimes_i(T_i,\mu_i,\varepsilon_i)$ *as*

$$QCC\big(\otimes_i(T_i,\mu_i,\varepsilon_i)\big) = \min_{\Pi_n\in\mathscr{T}(\otimes_i(T_i,\mu_i,\varepsilon_i))} QCC(\Pi_n),$$

and the M-message, n-fold quantum communication complexity *of* $\otimes_i(T_i,\mu_i,\varepsilon_i)$ *as*

$$QCC^M\big(\otimes_i(T_i,\mu_i,\varepsilon_i)\big) = \min_{\Pi_n\in\mathscr{T}^M(\otimes_i(T_i,\mu_i,\varepsilon_i))} QCC(\Pi_n).$$

Note that for all n,

$$QCC\big(\otimes_i(T_i,\mu_i,\varepsilon_i)\big) \leq \sum_i QCC(T_i,\mu_i,\varepsilon_i), \tag{2.4.6}$$

$$QCC^M\big(\otimes_i(T_i,\mu_i,\varepsilon_i)\big) \leq \sum_i QCC^M(T_i,\mu_i,\varepsilon_i), \tag{2.4.7}$$

as is made clear by running in parallel the n protocols achieving the minimum in the definition of quantum communication complexity of $(T_i,\mu_i,\varepsilon_i)$. Restricting to performing the same task, we have $QCC((T,\mu,\varepsilon)^{\otimes n}) \leq nQCC(T,\mu,\varepsilon)$. We define the amortized quantum communication complexity $AQCC(T,\mu,\varepsilon)$ as the asymptotic cost per copy for computing (T,μ,ε) n times in parallel.

Definition 2.4.6. *For a classical task* (T,μ,ε) *and a bound* $M\in\mathbb{N}$ *on the number of messages, we define the* ε*-error* amortized quantum communication complexity *of* T *on input* μ *as*

$$AQCC(T,\mu,\varepsilon) = \lim_{n\to\infty}\frac{1}{n}QCC((T,\mu,\varepsilon)^{\otimes n}),$$

and the M-message, ε*-error* amortized quantum communication complexity *of* T *on input* μ *as*

$$AQCC^M(T,\mu,\varepsilon) = \lim_{n\to\infty}\frac{1}{n}QCC^M((T,\mu,\varepsilon)^{\otimes n}).$$

These limits are well-defined, and we will provide an information-theoretic charac-

terization for them in Chapter 4.

We also have corresponding definitions for product classical tasks wih worst-case error. With similar notation as above, the product task $\otimes_i(T_i, \varepsilon_i)$ requires that for each i, for each global input (x^n, y^n), the error of a randomized protocol Π_ν satisfies $\sum_r \nu(r) \cdot \Pr_{\Pi_r}[\big(x_i, y_i, \Pi_r^i(x^n, y^n)\big) \notin T_i] \leq \varepsilon_i$. We denote $\mathscr{T}(\otimes_i(T_i, \varepsilon_i))$ the set of all protocols in the randomized model achieving the above goal of having error ε_i for each corresponding task. When restricting this set to M-message protocols, we denote it as $\mathscr{T}^M(\otimes_i(T_i, \varepsilon_i))$. We have the following definitions.

Definition 2.4.7. *For a product classical task $\otimes_i(T_i, \varepsilon_i)$ and a bound $M \in \mathbb{N}$ on the number of messages, we define the* n-fold quantum communication complexity *of $\otimes_i(T_i, \varepsilon_i)$ as*

$$QCC\big(\otimes_i (T_i, \varepsilon_i)\big) = \min_{\Pi_n \in \mathscr{T}(\otimes_i(T_i, \varepsilon_i))} QCC(\Pi_n),$$

and the M-message, n-fold quantum communication complexity *of $\otimes_i(T_i, \varepsilon_i)$ as*

$$QCC^M\big(\otimes_i (T_i, \varepsilon_i)\big) = \min_{\Pi_n \in \mathscr{T}^M(\otimes_i(T_i, \varepsilon_i))} QCC(\Pi_n).$$

Note that for all n,

$$QCC\big(\otimes_i (T_i, \varepsilon_i)\big) \leq \sum_i QCC(T_i, \varepsilon_i), \tag{2.4.8}$$

$$QCC^M\big(\otimes_i (T_i, \varepsilon_i)\big) \leq \sum_i QCC^M(T_i, \varepsilon_i), \tag{2.4.9}$$

as is made clear by running in parallel the n protocols achieving the minimum in the definition of quantum communication complexity of (T_i, ε_i).

2.4.3 Quantum Tasks

We also study the quantum analogue of distributional classical tasks, along with the product version of such tasks: quantum tasks with fixed input state.

2.4.3.1 Tasks with Fixed Input State

Recall that a quantum task consists of implementing a channel $\mathcal{N} \in \mathcal{C}(A_{in}B_{in}, A_{out}B_{out})$ on input $\rho \in \mathcal{D}(A_{in}B_{in})$ with error at most $\varepsilon \in [0,2]$. It is denoted $(\mathcal{N}, \rho, \varepsilon)$, and we denote the set of all protocols in the hybrid model implementing the quantum task $(\mathcal{N}, \rho, \varepsilon)$ as $\mathcal{T}(\mathcal{N}, \rho, \varepsilon)$. If we want to restrict this set to bounded round protocols with M messages, we write $\mathcal{T}^M(\mathcal{N}, \rho, \varepsilon)$. The notion of quantum communication complexity of a channel is defined as follows.

Definition 2.4.8. *For a quantum task $(\mathcal{N}, \rho, \varepsilon)$ and a bound $M \in \mathbb{N}$ on the number of messages, we define the ε-error* quantum communication complexity *of $\mathcal{N}$ on input ρ as*

$$QCC(\mathcal{N}, \rho, \varepsilon) = \min_{\Pi \in \mathcal{T}(\mathcal{N}, \rho, \varepsilon)} QCC(\Pi),$$

and the M-message, ε-error quantum communication complexity *of $\mathcal{N}$ on input ρ as*

$$QCC^M(\mathcal{N}, \rho, \varepsilon) = \min_{\Pi \in \mathcal{T}^M(\mathcal{N}, \rho, \varepsilon)} QCC(\Pi).$$

Similarly to the case for quantum communication complexity of classical tasks, these quantities are discontinuous in their parameters $\mathcal{N}, \rho, \varepsilon$, and other analogous remarks hold. Also note that for pure states, the quantum communication vanishes. This is because, in this quantum analogue to the distributional setting, pure states have no correlation with the outside world, so we can consider a protocol that is given, as entanglement for the protocol, the output of the channel acting on the pure state, and outputs it without communication.

Remark 2.4.3. *For any $\mathcal{N}$, ρ, ε, ε_1, ε_2, M, with $0 \leq \varepsilon_1 \leq \varepsilon_2 \leq 2$, the following holds:*

$$\begin{aligned} QCC(\mathcal{N},\rho,\varepsilon) &\leq QCC^M(\mathcal{N},\rho,\varepsilon), \\ QCC(\mathcal{N},\rho,\varepsilon_2) &\leq QCC(\mathcal{N},\rho,\varepsilon_1), \\ QCC^M(\mathcal{N},\rho,\varepsilon_2) &\leq QCC^M(\mathcal{N},\rho,\varepsilon_1), \\ QCC^M(\mathcal{N},\rho,2) &= 0. \end{aligned}$$

If $M \geq 2$,

$$QCC^M(\mathcal{N},\rho,0) \leq \log|A_{in}| + \log|A_{out}|.$$

Also, for any $\varepsilon \in [0,2]$, the following holds for any pure state ρ:

$$QCC^M(\mathcal{N},\rho,\varepsilon) = 0.$$

2.4.3.2 Product Tasks

We are also interested in the quantum communication complexity of implementing multiple quantum tasks in parallel. For channels $\mathcal{N}_1, \cdots, \mathcal{N}_n$, input states $\rho_1, \cdots, \rho_n$ and error parameters $\varepsilon_1, \cdots, \varepsilon_n$, we call $\otimes_i(\mathcal{N}_i, \rho_i, \varepsilon_i)$ a product quantum task. A protocol Π_n is said to implement the product quantum task $\otimes(\mathcal{N}_i, \rho_i, \varepsilon_i)$ if, for all $i \in [n]$,

$$\| \operatorname{Tr}_{\neg(A^i_{out}B^i_{out}R^i)} \circ \Pi_n(\otimes_j \rho_j^{A^j_{in}B^j_{in}R^j}) - \mathcal{N}_i(\rho_i^{A^i_{in}B^i_{in}R^i}) \|_1 \leq \varepsilon_i. \tag{2.4.10}$$

We have implicitly used the fact that it is possible to find a purification of $\otimes_j \rho_j^{A^j_{in}B^j_{in}}$ with a decomposition of the purifying register $R = R_1 \otimes \cdots \otimes R_n$, and with ρ_i purified by register R_i. Similarly to the classical setting, this error criterion corresponds to the one achieved when sequentially implementing the n tasks $(\mathcal{N}_i, \rho_i, \varepsilon_i)$, and is weaker than demanding to simulate it n times with overall error $\varepsilon = \min_i \varepsilon_i$. Once again, the reason for this is that asking for overall error ε could be a much harder task. Indeed, consider n times the same task with a purified input state that is exactly ε away, for

some $\varepsilon \in (0,1)$ in trace distance to a state which is product with respect to the $A_{in}B_{in} - R$ bipartite cut. Then, since the trace distance is monotone under noisy channels, the parties can simulate the channel at zero communication cost and achieve error ε by taking, as part of the entanglement of their protocol, the $A_{out}B_{out}$ registers of the channel acting on that product state. Thus the quantum communication complexity is zero. We can then also achieve the task of amortized quantum communication complexity with ε error in each input at zero communication. However, using the operational interpretation of the trace distance as the best bias in a distinguishability experiment, the amortized quantum communication task in which we ask for overall error ε cannot be achieved at zero communication cost, since having access to many instances of the output state leads to exponentially better distinguishability whenever starting with distinguishability greater than zero between the actual input and the product state [11]. Hence, if we want to study amortized quantum communication complexity with non-trivial error, we have to settle for such a success parameter.

We denote $\mathscr{T}(\otimes_i(\mathscr{N}_i, \rho_i, \varepsilon_i))$ the set of all protocols achieving the above goal of having ε_i error in each corresponding output. When restricting this set to M-message protocols, we denote it as $\mathscr{T}^M(\otimes_i(\mathscr{N}_i, \rho_i, \varepsilon_i))$. If, for all i, $(\mathscr{N}_i, \rho_i, \varepsilon_i) = (\mathscr{N}, \rho, \varepsilon)$, we may use the notation $\otimes_i(\mathscr{N}_i, \rho_i, \varepsilon_i) = (\mathscr{N}, \rho, \varepsilon)^{\otimes n}$. We have the following definitions.

Definition 2.4.9. *For a product quantum task $\otimes(\mathscr{N}_i, \rho_i, \varepsilon_i)$ and a bound M on the number of messages, we define the* n-fold quantum communication complexity *of $\otimes_i(\mathscr{N}_i, \rho_i, \varepsilon_i)$ as*

$$QCC(\otimes_i(\mathscr{N}_i, \rho_i, \varepsilon_i)) = \min_{\Pi_n \in \mathscr{T}(\otimes_i(\mathscr{N}_i, \rho_i, \varepsilon_i))} QCC(\Pi_n),$$

and the M-message n-fold quantum communication complexity *of $\otimes_i(\mathscr{N}_i, \rho_i, \varepsilon_i)$ as*

$$QCC^M(\otimes_i(\mathscr{N}_i, \rho_i, \varepsilon_i)) = \min_{\Pi_n \in \mathscr{T}^M(\otimes_i(\mathscr{N}_i, \rho_i, \varepsilon_i))} QCC(\Pi_n),$$

Definition 2.4.10. *For a quantum task $(\mathscr{N}, \rho, \varepsilon)$ and a bound M on the number of messages, we define the ε-error* amortized quantum communication complexity *of $\mathscr{N}$ on*

input ρ *as*

$$AQCC(\mathcal{N},\rho,\varepsilon) = \lim_{n\to\infty}\frac{1}{n}QCC((\mathcal{N},\rho,\varepsilon)^{\otimes n}),$$

and the M-message, ε*-error* amortized quantum communication complexity *of* $\mathcal{N}$ *on input* ρ *as*

$$AQCC^{M}(\mathcal{N},\rho,\varepsilon) = \lim_{n\to\infty}\frac{1}{n}QCC^{M}((\mathcal{N},\rho,\varepsilon)^{\otimes n}).$$

Note that for all n, $QCC^{M}((\mathcal{N},\rho,\varepsilon)^{\otimes n}) \leq nQCC^{M}(\mathcal{N},\rho,\varepsilon)$, as is made clear by running n times in parallel a protocol achieving the minimum in the definition of the quantum communication complexity. Hence, the amortized quantum communication complexity is bounded by the quantum communication complexity. The aim of the direct sum question is to provide the reverse inequality.

CHAPTER 3

QUANTUM STATE REDISTRIBUTION

3.1 Introduction

Quantum state redistribution is the most general noiseless coding task in unidirectional quantum information theory. It is also of fundamental importance in the context of quantum information complexity. We study in this chapter the amount of resources required to implement quantum state redistribution in a one-shot setting.

3.1.1 The Information Processing Task

In the task of quantum state redistribution, we are interested in the amounts of quantum communication and entanglement that are required to transmit part of the system of one party to another party who possesses some side information about this system. It is required that all correlations, including those with any external system, are maintained. More precisely, consider two parties Alice and Bob, with Alice initially holding the A and C registers, and Bob holding the B register. The goal is then for Alice to transmit the C register to Bob. If we consider a reference register R holding a purification of the ABC systems, then the global state on $ABCR$ is uncorrelated with any other external system, and it is sufficient to insure that correlations are maintained across these systems. Please refer to Section 3.2.3 for a formal description.

In the independent and identical (iid) asymptotic version, Alice and Bob want to perform this task on blocks of n identical states, for n large, and we are interested in the best asymptotic rates achievable. Luo and Devetak [100] proved a converse theorem in the iid asymptotic regime, stating that the quantum communication rate q and the sum of the entanglement consumption rate e and the quantum communication rate q must be at least

$$q \geq \frac{1}{2} I(C;R|B) \quad \text{and} \quad e + q \geq H(C|B) \, . \tag{3.1.1}$$

Subsequently, Devetak and Yard [54, 140] proved that these rates are also achievable and hence fully characterize the achievable rate region for iid asymptotic quantum state redistribution. Note that since the overall state ρ_{ABCR} for quantum state redistribution is pure, we have the symmetry in A–B,

$$I(C;R|B)_\rho = I(C;R|A)_\rho \ . \tag{3.1.2}$$

Later, the achievability proofs for quantum state redistribution were significantly simplified by Oppenheim [107] and independently by Ye, Bai and Wang [141]. State redistribution can be seen as the most general bipartite noiseless coding problem, and indeed, other noiseless quantum coding primitives such as Schumacher source coding [115], quantum state merging [74, 75] (including fully quantum Slepian-Wolf [3]), and state splitting [3] can be obtained by considering the case of trivial AB, A or B system, respectively. Quantum state redistribution can also be understood as the fully quantum analogue of the tensor power input reverse Shannon theorem [20, 22] with feedback to the sender and side information at the receiver [100, 134].

In recent years, there has been some effort on finding good bounds for the one-shot version of these results (see, e.g., [21, 22, 43, 51, 58, 59] and references therein). In the one-shot setting, instead of being interested in iid asymptotic rates, we are interested in the cost of achieving these tasks when only a single copy of the input state is available. Useful bounds are often stated in terms of so-called smooth conditional entropies (see the theses of Renner [111] and Tomamichel [126], as well as references therein).

3.1.2 Link to Interactive Protocols

In our study of quantum information complexity in Chapter 4, we take the following view on interactive protocols in the hybrid communication model. The isometries U_i the parties apply locally correspond to channels, the output of which is to be transmitted to the other party. The task of compressing messages to their information content is then naturally associated with the task of optimally simulating the corresponding channel. But as said above, channel simulation with side-information at the receiver and coherent

feedback to the sender is equivalent to quantum state redistribution, so that this task will be of fundamental importance in our study of compression of interactive protocols.

3.1.3 Overview of Results

Most of the results in this chapter are based on a collaboration with Mario Berta and Matthias Christandl [23]. We are interested in finding good bounds for the quantum communication cost of one-shot quantum state redistribution in terms of smooth conditional entropies. Our main result states that it is possible to implement quantum state redistribution for a pure quantum state ρ^{ABCR} up to error ε, for $\varepsilon > 0$, with quantum communication cost at most

$$\frac{1}{2}\left(H_{\max}^{\varepsilon}(C|B)_{\rho} - H_{\min}^{\varepsilon}(C|BR)_{\rho}\right) + O\left(\log(1/\varepsilon)\right) , \tag{3.1.3}$$

when free entanglement assistance is available. Note that both the conditional min- and max-entropy terms appearing in (3.1.3) are smoothed, notwithstanding the fact that it is in general unknown how to simultaneously smooth marginals of overlapping quantum systems (see, e.g., [57] and references therein). For the special case of iid asymptotic resources, this allows us to recover the optimal quantum communication rate (3.1.1) by means of the fully quantum asymptotic equipartition property for smooth conditional entropies [127].

We also state lower bounds in terms of smooth conditional min- and max-entropies, and smooth max-information. However, our achievability bound (3.1.3) only matches these lower bounds in an iid asymptotic scenario. We then also speculate on how to improve the bound (3.1.3) with the help of embezzling entangled quantum states [130] along with some of the ideas in [22].

Using the substate theorem of Jain, Radhakrishnan and Sen [78, 80, 83], we also obtain a one-shot bound in terms of conditional quantum mutual information, however with a $1/\varepsilon^2$ dependence on the allowed error ε. This bound finds an application in Section 4.6 to prove the first multi-round general direct sum theorem in quantum communication complexity, through the notion of quantum information complexity. This

result is from Ref. [129].

Since quantum state redistribution plays a fundamental role in the definition of quantum information cost of interactive protocols, we also show that the lower bounds extend to multi-round protocols. Thus, the definition of quantum information cost is robust for interactive communication in this sense.

Organization: This chapter is structured as follows. In Section 3.2, we introduce our notation and state the definitions of the relevant smooth entropy measures and of the quantum state redistribution task. We then present our converse bounds in Section 3.3. Our achievability bounds are presented in Section 3.4, which begins with a discussion of the work of Oppenheim [107], arguing that quantum state redistribution can be optimally decomposed into two applications of quantum state merging. We end with some conclusions.

3.2 One-Shot Quantum Information Theory

In order to study one-shot quantum state redistribution, we require new notions of distance and information more appropriate in the one-shot setting. We also state some useful properties for these, and provide a formal definition for one-shot quantum state redistribution.

3.2.1 Information and Distance Measures

For one-shot information measures, we also require smooth versions, optimized over an ε-ball around the state under consideration. We later define an appropriate notion of distance in order to do so, but first we extend the definition of quantum states to allow for subnormalized states, in order to define ε-ball of sub-normalized states. The set of linear, nonnegative operators is denoted by $\mathscr{P}(A)$. We denote by $\mathscr{D}_{\leq}(A)$ the set of sub-normalized states on A, i.e., the set of operators $\rho^A \in \mathscr{P}(A)$ that are positive semi-definite, denoted $\rho \geq 0$, and have trace at most one. The set of normalized states is still denoted by $\mathscr{D}(A)$. Given a multipartite state $\rho^{AB} \in \mathscr{D}_{\leq}(AB)$, we write $\rho^A = \mathrm{Tr}_B[\rho^{AB}]$ for the reduced state on the system A. A purification of state $\rho \in \mathscr{D}_{\leq}(A)$ is a rank one

operator ρ^{AR} such that $\mathrm{Tr}_R(\rho^{AR}) = \rho^A$.

The relative entropy of $\rho \in \mathscr{D}_{\leq}(A)$ with respect to $\sigma \in \mathscr{P}(A)$ is defined as

$$D(\rho\|\sigma) = \mathrm{Tr}(\rho \log \rho) - \mathrm{Tr}(\rho \log \sigma). \tag{3.2.1}$$

Note that we can rewrite the conditional entropy of A given B for $\rho \in \mathscr{D}(AB)$ as

$$H(A|B)_\rho = -D(\rho^{AB}\|I^A \otimes \rho^B) \tag{3.2.2}$$

$$= -\inf_{\sigma^B \in \mathscr{D}(B)} D(\rho^{AB}\|I^A \otimes \sigma^B). \tag{3.2.3}$$

The max-relative entropy of $\rho \in \mathscr{D}_{\leq}(A)$ with respect to $\sigma \in \mathscr{P}(A)$ is defined as

$$D_{\max}(\rho\|\sigma) = \inf\left\{\lambda \in \mathbb{R} : 2^\lambda \sigma \geq \rho\right\} . \tag{3.2.4}$$

The conditional min-entropy of A given B for $\rho^{AB} \in \mathscr{D}_{\leq}(AB)$ is defined as

$$H_{\min}(A|B)_\rho = -\inf_{\sigma \in \mathscr{D}(B)} D_{\max}\left(\rho^{AB}\|I^A \otimes \sigma^B\right) . \tag{3.2.5}$$

The conditional max-entropy of A given B for $\rho^{AB} \in \mathscr{D}_{\leq}(AB)$, with purification $\rho^{ABR} \in \mathscr{D}_{\leq}(ABR)$ for some system R, is defined as

$$H_{\max}(A|B)_\rho = -H_{\min}(A|R)_\rho . \tag{3.2.6}$$

Note that this definition does not depend on the choice of the purification. The max-information that B has about A for $\rho^{AB} \in \mathscr{D}_{\leq}(AB)$ is defined as

$$I_{\max}(A:B)_\rho = \inf_{\sigma \in \mathscr{D}(B)} D_{\max}\left(\rho^{AB}\|\rho^A \otimes \sigma^B\right) . \tag{3.2.7}$$

To define smooth entropy measures, an optimization over a set of nearby states is performed. The distance measure used is the purified distance [128], defined for $\rho, \sigma \in$

$\mathscr{D}_{\leq}(A)$ as

$$P(\rho,\sigma) = \sqrt{1-\bar{F}^2(\rho,\sigma)}\,, \tag{3.2.8}$$

in which the generalized fidelity $\bar{F}$ is defined in terms of the fidelity F, with $F(\rho,\sigma) = \|\sqrt{\rho}\sqrt{\sigma}\|_1$, as

$$\bar{F}(\rho,\sigma) = F(\rho,\sigma) + \sqrt{(1-\mathrm{Tr}(\rho))(1-\mathrm{Tr}(\sigma))}\,. \tag{3.2.9}$$

We then define an ε-ball around $\rho \in \mathscr{D}_{\leq}(A)$ as

$$\mathscr{B}^{\varepsilon}(\rho) = \{\bar{\rho} \in \mathscr{D}_{\leq}(A) : P(\rho,\bar{\rho}) \leq \varepsilon\}\,. \tag{3.2.10}$$

For $\varepsilon \geq 0$, the smooth max-relative entropy of $\rho \in \mathscr{D}_{\leq}(A)$ with respect to $\sigma \in \mathscr{P}(A)$ is then defined as

$$D^{\varepsilon}_{\max}(\rho\|\sigma) = \inf_{\bar{\rho}\in\mathscr{B}^{\varepsilon}(\rho)} D_{\max}(\rho\|\sigma)\,. \tag{3.2.11}$$

For $\varepsilon \geq 0$, the smooth conditional min-entropy of A given B for $\rho^{AB} \in \mathscr{D}_{\leq}(AB)$ is then defined as

$$H^{\varepsilon}_{\min}(A|B)_{\rho} = \sup_{\bar{\rho}\in\mathscr{B}^{\varepsilon}(\rho^{AB})} H_{\min}(A|B)_{\bar{\rho}}\,, \tag{3.2.12}$$

and the smooth conditional max-entropy as

$$H^{\varepsilon}_{\max}(A|B)_{\rho} = \inf_{\bar{\rho}\in\mathscr{B}^{\varepsilon}(\rho^{AB})} H_{\max}(A|B)_{\bar{\rho}}\,. \tag{3.2.13}$$

The smooth max-information that B has about A for $\rho^{AB} \in \mathscr{D}_{\leq}(AB)$ is defined as

$$I^{\varepsilon}_{\max}(A:B)_{\rho} = \inf_{\bar{\rho}\in\mathscr{B}^{\varepsilon}(\rho^{AB})} I_{\max}(A:B)_{\bar{\rho}}\,. \tag{3.2.14}$$

3.2.2 Properties

We make use of the following properties of the purified distance and the above information measures.

The properties of the purified distance that we use are proved in [126, 128]. First, it is a metric, so it is symmetric in ρ, σ, non-negative, evaluate to 0 if and only if $\rho = \sigma$ and it satisfies the triangle inequality. Moreover, it is monotone under noisy channels: for any ρ_1, $\rho_2 \in \mathscr{D}_{\leq}(A)$ and $\mathscr{N} \in \mathscr{C}(A,B)$,

$$P(\mathscr{N}(\rho_1), \mathscr{N}(\rho_2)) \leq P(\rho_1, \rho_2). \tag{3.2.15}$$

For isometries, the inequality becomes an equality, a property called isometric invariance of the purified distance. Also, the purified distance remains invariant when adjoining an uncorrelated system: for any ρ_1, $\rho_2 \in \mathscr{D}_{\leq}(A)$, $\sigma \in \mathscr{D}_{\leq}(B)$

$$P(\rho_1 \otimes \sigma, \rho_2 \otimes \sigma) = P(\rho_1, \rho_2). \tag{3.2.16}$$

The purified distance is related to the trace distance through a generalization of the Fuchs-van de Graaf inequalities [63]: for any ρ_1, $\rho_2 \in \mathscr{D}_{\leq}(A)$, it holds that

$$\frac{1}{2}\|\rho_1 - \rho_2\|_1 \leq P(\rho_1, \rho_2) \leq \sqrt{2\|\rho_1 - \rho_2\|_1}. \tag{3.2.17}$$

In the case that $\mathrm{Tr}(\rho_1) = \mathrm{Tr}(\rho_2)$, we can strengthen this to $P(\rho_1, \rho_2) \leq \sqrt{\|\rho_1 - \rho_2\|_1}$. We also make use of the following variant of Uhlmann's theorem.

Lemma 3.2.1. *Let ρ_1, $\rho_2 \in \mathscr{D}_{\leq}(A)$ have purifications $\rho_1^{AR_1}$, $\rho_2^{AR_2}$, with $|R_1| \leq |R_2|$. Then, there exists an isometry $V^{R_1 \to R_2}$ such that*

$$P\left(\rho_1^A, \rho_2^A\right) = P\left(V\left(\rho_1^{AR_1}\right), \rho_2^{AR_2}\right). \tag{3.2.18}$$

The following bound holds on the smooth max-information [22, Lemma B.9].

Lemma 3.2.2. *Let $\varepsilon \geq 0$ and $\rho^{ABC} \in \mathscr{D}(ABC)$. Then, we have*

$$I^{\varepsilon}_{\max}(A:BC)_{\rho} \leq I^{\varepsilon}_{\max}(A:B)_{\rho} + 2\log|C| \ . \tag{3.2.19}$$

We also need the same type of bound for smooth conditional min-entropy.

Lemma 3.2.3. *Let $\varepsilon \geq 0$ and $\rho^{ABC} \in \mathscr{D}(ABC)$. Then, we have*

$$H^{\varepsilon}_{\min}(A|B)_{\rho} \leq H^{\varepsilon}_{\min}(A|BC)_{\rho} + 2\log|C| \ . \tag{3.2.20}$$

Note that by duality, a similar result holds for smooth conditional max-entropy. The max-information and the min-entropy are monotone under local operations [22, 126]. That is, they satisfy a data processing inequality: for any $\varepsilon \geq 0$, $\rho \in \mathscr{D}_{\leq}(AB)$ and $\mathscr{N} \in \mathscr{C}(B,C)$,

$$H^{\varepsilon}_{min}(A|B)_{\rho} \leq H^{\varepsilon}_{min}(A|C)_{\mathscr{N}(\rho)}, \tag{3.2.21}$$

$$I^{\varepsilon}_{max}(A;B)_{\rho} \geq I^{\varepsilon}_{max}(A;C)_{\mathscr{N}(\rho)}. \tag{3.2.22}$$

They are left invariant under appending an uncorrelated system to one subsystem: for any $\varepsilon \geq 0, \rho \in \mathscr{D}_{\leq}(AB)$ and $\sigma \in \mathscr{D}_{\leq}(C)$,

$$H^{\varepsilon}_{min}(A|B)_{\rho} = H^{\varepsilon}_{min}(A|BC)_{\rho\otimes\sigma}, \tag{3.2.23}$$

$$I^{\varepsilon}_{max}(A;B)_{\rho} = I^{\varepsilon}_{max}(A;BC)_{\rho\otimes\sigma}. \tag{3.2.24}$$

For the von Neumann mutual information, we introduce the following parameter for any two states ρ_1, $\rho_2 \in \mathscr{D}(AB)$ that are close:

$$\delta_{I(A;B)}(\rho_1,\rho_2) = |I(A;B)_{\rho_1} - I(A;B)_{\rho_2}|. \tag{3.2.25}$$

By continuity, the following holds when $\varepsilon = \|\rho_1 - \rho_2\|_1$ is small enough:

$$\delta_{I(A;B)}(\rho_1,\rho_2) \leq 8\varepsilon\log|A| + 4H_2(\varepsilon). \tag{3.2.26}$$

3.2.3 Definition of Quantum State Redistribution

The fully quantum task of state redistribution plays a central role in our developments, and can be formulated as a quantum task. We have the following definition.

Definition 3.2.1. *We say that the bipartite channel $\mathscr{R} \in \mathscr{C}(A_{in}B_{in}, A_{out}B_{out})$ implements* state redistribution *on input $\rho^{A_{in}B_{in}}$, with $A_{in} = AC$, $B_{in} = B$, $A_{out} = A$, $B_{out} = BC$: it implements the identity channel on such a state and such a partition of the input-output registers, i.e. it transfers the C part of ρ from Alice to Bob. We say that a protocol Π is an ε-error state redistribution protocol for ρ^{ABC} if $\Pi \in \mathscr{T}(\mathscr{R}, \rho, \varepsilon)$.*

When only EPR pairs are consumed as pre-shared entanglement, and possibly some EPR pairs are generated, it makes sense to also speak of the net entanglement cost of a protocol. This is the difference between the number of pairs consumed vs. the number of pairs generated. The net cost can be negative if a protocol generates more pairs than it consumes. For a protocol Π consuming e_c EPR pairs as pre-shared entanglement and generating e_g EPR pairs, up to error ε, when run on input ρ, we denote the net entanglement consumption cost as $e(\Pi, \rho, \varepsilon) = e_c - e_g$.

3.3 Converse Bounds

We first state lower bounds on the amount of quantum communication required to implement quantum state redistribution. These are called *converse* bounds in the information theory literature. We first prove such bounds for one-message protocols, and then prove that these bounds also hold for multi-round protocols, in an interactive setting.

3.3.1 Single-Round Protocols

We provide lower bounds on the amount of communication required for one-shot state redistribution. They do not match the upper bound given in the direct coding theorem in general, but in the asymptotic regime they also simplify to the conditional mutual information $I(C;R|B)_\rho$.

Proposition 3.3.1. *Let $\varepsilon_1,\ \varepsilon_2 \geq 0$ and $\rho \in \mathscr{D}(ABC)$ with purifying register R. Then, for every one-message protocol $\Pi \in \mathscr{T}(\mathscr{R}, \rho, \varepsilon_1)$, the quantum communication cost is lower bounded by*

$$QCC(\Pi) \geq \frac{1}{2} I_{\max}^{\varepsilon_1+\varepsilon_2}(R;BC)_\rho - \frac{1}{2} I_{\max}^{\varepsilon_2}(R;B)_\rho, \tag{3.3.1}$$

$$QCC(\Pi) \geq \frac{1}{2} H_{\min}^{\varepsilon_2}(R|B)_\rho - \frac{1}{2} H_{\min}^{\varepsilon_1+\varepsilon_2}(R|BC)_\rho, \tag{3.3.2}$$

$$QCC(\Pi) \geq \frac{1}{2} H_{\max}^{\varepsilon_1+\varepsilon_2}(R|B)_\rho - \frac{1}{2} H_{\max}^{\varepsilon_2}(R|BC)_\rho, \tag{3.3.3}$$

$$QCC(\Pi) \geq \frac{1}{2} I(R;C|B)_\rho - \frac{1}{2} \max_{\sigma \in \mathscr{B}^{\varepsilon_1}(\rho)} \delta_{I(R;BC)}(\rho, \sigma), \tag{3.3.4}$$

and the same bounds hold for B replaced with A.

Note that the first bound is optimal in the case of a trivial B register, for state splitting, while the corresponding bound with A replacing B is optimal in the case of a trivial A register, for state merging. Also note that, in contrast to the direct coding bound, the time-reversal symmetry between the A, B systems is not apparent here. Finally, note that these bounds hold irrespective of the kind of entanglement used.

Proof. (Proposition 3.3.1) Similar to the proof of the optimal bound on state splitting in [22], we look at the correlations between Bob and the reference register. To be able to use Lemma 3.2.2, we look at the max-information that Bob has about R at the end of any protocol for quantum state redistribution. A one-message protocol for state redistribution necessarily has the following structure: local operations on Alice's side, followed by communication from Alice to Bob, and then local operations on Bob's side. In more detail:

General protocol Π for input ρ^{ABCR} using entanglement $\phi^{T_A^{in} T_B^{in}}$

1. Alice holds the A, C, T_A^{in} systems at the outset, and Bob the B, T_B^{in} systems.

2. Alice applies a local operation on the ACT_A^{in} registers. Her registers are then $T_A^{out}A'Q$. The joint state is $\sigma^{T_A^{out}A'QBT_B^{in}R}$.
3. Alice transmits the Q register to Bob.
4. Bob applies a local operation on QBT_B^{in}. His registers are then $T_B^{out}B'C'$. The joint state is $\theta^{T_A^{out}T_B^{out}A'B'C'R}$.

– The requirement is that the $A'B'C'R$ part is ε_1-close to $\rho^{A'B'C'R} = I^{ABC\to A'B'C'}(\rho^{ABCR})$ in purified distance.

For the bound in terms of max-information, consider a state $\hat{\theta}^{A'B'C'R} \in D_{\leq}(A'B'C'R)$ such that $P(\theta^{A'B'C'R}, \hat{\theta}^{A'B'C'R}) \leq \varepsilon_2$ and $I_{\max}^{\varepsilon_2}(R;B'C')_\theta = I_{\max}(R;B'C')_{\hat{\theta}}$. Such a state must exist by the definition of smoothing and the properties of the purified distance. Then $P(\rho^{A'B'C'R}, \hat{\theta}^{A'B'C'R}) \leq \varepsilon_1 + \varepsilon_2$ by the triangle inequality since $\theta^{A'B'C'R}$ must be ε_1-close to $\rho^{A'B'C'R}$. We get the following chain of inequalities

$$
\begin{aligned}
I_{\max}^{\varepsilon_1+\varepsilon_2}(R;BC)_\rho &\leq I_{\max}(R;B'C')_{\hat{\theta}} && (3.3.5)\\
&= I_{\max}^{\varepsilon_2}(R;B'C')_\theta && (3.3.6)\\
&\leq I_{\max}^{\varepsilon_2}(R;QBT_B^{in})_\sigma && (3.3.7)\\
&\leq I_{\max}^{\varepsilon_2}(R;BT_B^{in})_\sigma + 2\log|Q| && (3.3.8)\\
&= I_{\max}^{\varepsilon_2}(R;BT_B^{in})_{\rho\otimes\phi} + 2\log|Q| && (3.3.9)\\
&= I_{\max}^{\varepsilon_2}(R;B)_\rho + 2\log|Q|\ , && (3.3.10)
\end{aligned}
$$

in which the first inequality follows by definition of smooth max-information and monotonicity of purified distance, since $\theta^{A'B'C'R}$ is within distance ε_1 of $\rho^{A'B'C'R}$, the first equality is by the choice of $\hat{\theta}$, the second inequality is because the max-information is monotone under local operations, the third inequality follows by Lemma 3.2.2, the second equality is because local operations of Alice do not change the max-information of Bob about the reference, and the last is because $\phi^{T_A^{in}T_B^{in}}$ is uncorrelated to ρ^{ABCR}.

For the bound in terms of conditional min-entropy, we similarly get, by taking an appropriate $\hat{\theta}$ and using an unlockability property of min-entropy (Lemma 3.2.3),

$$\begin{aligned}
H_{\min}^{\varepsilon_1+\varepsilon_2}(R|BC)_\rho &\geq H_{\min}(R|B'C')_{\hat{\theta}} && (3.3.11)\\
&= H_{\min}^{\varepsilon_2}(R;B'C')_\theta && (3.3.12)\\
&\geq H_{\min}^{\varepsilon_2}(R|QBT_B^{in})_\sigma && (3.3.13)\\
&\geq H_{\min}^{\varepsilon_2}(R|BT_B^{in})_\sigma - 2\log|Q| && (3.3.14)\\
&= H_{\min}^{\varepsilon_2}(R|BT_B^{in})_{\rho\otimes\phi} - 2\log|Q| && (3.3.15)\\
&= H_{\min}^{\varepsilon_2}(R|B)_\rho - 2\log|Q|\ . && (3.3.16)
\end{aligned}$$

For the bound in terms of the conditional max-entropy, we obtain the bound with the A system instead of B by using the duality relation of conditional min- and max-entropy. For the bound in terms of von Neumann mutual information, we get, similarly to the derivation for max-information and using the definition of $\delta_{I(R;BC)}(\theta,\rho)$ along with the chain rule for conditional quantum mutual information,

$$\begin{aligned}
I(R;BC)_\rho - \delta_{I(R;B'C')}(\theta,\rho) &\leq I(R;BC)_\theta && (3.3.17)\\
&\leq I(R;QBT_B^{in})_\sigma && (3.3.18)\\
&= I(R;BT_B^{in})_\sigma + I(R;Q|BT_B^{in})_\sigma && (3.3.19)\\
&\leq I(R;BT_B^{in})_\sigma + 2\log|Q| && (3.3.20)\\
&= I(R;BT_B^{in})_{\rho\otimes\phi} + 2\log|Q| && (3.3.21)\\
&= I(R;B)_\rho + 2\log|Q|. && (3.3.22)
\end{aligned}$$

We then get the remaining bounds by interchanging the A and B systems in those already proved, and by using the symmetry of state redistribution under time reversal. □

3.3.2 Multi-Round Protocols

Note that asymptotic quantum state redistribution composes perfectly. That is, given any decomposition $C = D_1 D_2 \cdots D_d$, the total asymptotic cost for transmitting C in a

single message versus transmitting it in d successive messages from Alice to Bob is the same: $I(C;R|B) = I(D_1;R|B) + I(D_2;R|BD_1) + \cdots + I(D_d;R|BD_1 \cdots D_{d-1})$. This follows from the chain rule for conditional quantum mutual information. By allowing back-communication, we could hope to improve on this. This is impossible: we show that even if there is back-communication from Bob to Alice, multiple messages cannot decrease the total asymptotic cost of communication from Alice to Bob.

Proposition 3.3.2. *Let $\varepsilon_1,\ \varepsilon_2 \geq 0$ and $\rho \in \mathscr{D}(ABC)$ with purification register R. Then, for every M-message protocol $\Pi \in \mathscr{T}(\mathscr{R}, \rho, \varepsilon_1)$, the quantum communication cost from Alice to Bob is lower bounded by*

$$QCC_{A\to B}(\Pi) \geq \frac{1}{2} I_{\max}^{\varepsilon_1+\varepsilon_2}(R;BC)_\rho - \frac{1}{2} I_{\max}^{\varepsilon_2}(R;B)_\rho, \tag{3.3.23}$$

$$QCC_{A\to B}(\Pi) \geq \frac{1}{2} H_{\min}^{\varepsilon_2}(R|B)_\rho - \frac{1}{2} H_{\min}^{\varepsilon_1+\varepsilon_2}(R|BC)_\rho, \tag{3.3.24}$$

$$QCC_{A\to B}(\Pi) \geq \frac{1}{2} H_{\max}^{\varepsilon_1+\varepsilon_2}(R|B)_\rho - \frac{1}{2} H_{\max}^{\varepsilon_2}(R|BC)_\rho, \tag{3.3.25}$$

$$QCC_{A\to B}(\Pi) \geq \frac{1}{2} I(R;C|B)_\rho - \frac{1}{2} \max_{\sigma \in \mathscr{B}^{\varepsilon_1}(\rho)} \delta_{I(R;BC)}(\rho, \sigma), \tag{3.3.26}$$

and the same bounds hold for B replaced with A.

Note that there is no dependence on the number M of messages in these lower bounds.

Proof. The proof is similar to the one in the single round case. Hence, we only write down the details for the bound in terms of max-information. We consider an M-message protocol in the hybrid model, for the case of even M (the case of odd M follows similarly). We use the following notation: $\rho_0 = \rho^{ABCR} \otimes \phi^{T_A T_B}$, $\rho_1 = U_1(\rho_0)$, $\rho_2 = U_2(\rho_1)$, $\cdots$, $\rho_{M+1} = U_{M+1}(\rho_M)$. It must hold that $P(\rho_{M+1}^{ABCR}, \rho^{ABCR}) \leq \varepsilon_1$. Consider a state $\hat{\theta}^{ABCR} \in D_{\leq}(ABCR)$ such that $P(\rho_{M+1}^{ABCR}, \hat{\theta}^{ABCR}) \leq \varepsilon_2$ and $I_{\max}^{\varepsilon_2}(R;BC)_{\rho_{M+1}} = I_{\max}(R;BC)_{\hat{\theta}}$. Such a state must exist by the definition of smoothing and the properties of the purified distance. Then $P(\rho^{ABCR}, \hat{\theta}^{ABCR}) \leq \varepsilon_1 + \varepsilon_2$ by the triangle inequality. We get the follow-

ing chain of inequalities:

$$
\begin{aligned}
I_{\max}^{\varepsilon_1+\varepsilon_2}(R;BC)_\rho &\leq I_{\max}(R;BC)_{\hat{\theta}} \\
&= I_{\max}^{\varepsilon_2}(R;BC)_{\rho_{M+1}} \\
&= I_{\max}^{\varepsilon_2}(R;BC)_{\rho_M} \\
&\leq I_{\max}^{\varepsilon_2}(R;C_{M-1}B_{M-2})_{\rho_{M-1}} \\
&\leq I_{\max}^{\varepsilon_2}(R;B_{M-2})_{\rho_{M-1}} + 2\log|C_{M-1}| \\
&= I_{\max}^{\varepsilon_2}(R;B_{M-2})_{\rho_{M-2}} + 2\log|C_{M-1}| \\
&\leq I_{\max}^{\varepsilon_2}(R;C_{M-3}B_{M-4})_{\rho_{M-3}} + 2\log|C_{M-1}| \\
&\leq \cdots \\
&\leq I_{\max}^{\varepsilon_2}(R;C_1B_0)_{\rho_1} + 2\sum_{i\geq 1}\log|C_{2i+1}| \\
&\leq I_{\max}^{\varepsilon_2}(R;B_0)_{\rho_1} + 2\sum_{i\geq 0}\log|C_{2i+1}| \\
&= I_{\max}^{\varepsilon_2}(R;BT_B)_{\rho_0} + 2QCC_{A\to B}(\Pi) \\
&= I_{\max}^{\varepsilon_2}(R;B)_\rho + 2QCC_{A\to B}(\Pi).
\end{aligned}
$$

The first inequality follows by definition of smooth max-information and monotonicity of purified distance. The first equality is by the choice of $\hat{\theta}$, and the second because U_{M+1} is applied on Alice's side. The second inequality is because the max-information is monotone under local operations, and the third inequality follows by Lemma 3.2.2. The third equality is because local operations of Alice do not change the max-information of Bob about the reference, and the following sequence of inequality follows by applying the last few ones repeatedly. The last inequality follows by Lemma 3.2.2, the following equality is by definition of B_0 and $QCC_{A\to B}(\Pi)$, and the last is because $\phi^{T_A^{in}T_B^{in}}$ is uncorrelated to ρ^{ABCR}. □

3.4 Achievability Bounds

We now give upper bounds on the amount of quantum communication required to implement quantum state redistribution. The achievability part of a coding theorem is called the *direct coding* theorem in the information theory literature. We first present the approach we take, then present our smooth entropy bounds, and finally use these to derive bounds in terms of conditional quantum mutual information.

3.4.1 Decoupling Approach to State Redistribution

We want to make use of the following observation of Oppenheim [107]: quantum state redistribution can be optimally decomposed into two applications of quantum state merging by introducing a coherent relay, through which all communication is relayed and possibly modified, and applying an ebit repackaging sub-protocol. In more detail, we consider four distinct parties, each holding a register. Charlie holds register C, that he wants to transmit to Bob, who holds register B, and to do so he may use help from Alice acting as a coherent relay, who holds register A. The state ρ in registers ABC is purified by state $|\rho\rangle^{ABCR}$ with the R register held by some reference party. The goal is to transmit C to Bob while minimizing the communication from Alice to Bob, and while keeping the overall correlation with the reference. No direct communication between Charlie and Bob is allowed. We might also keep track of communication between Charlie and Alice, as well as of the entanglement consumption and generation between both Charlie and Alice, and Alice and Bob, but here our main focus is the communication between Alice and Bob. A key observation is that applying a single decoupling unitary at Charlie's side suffices to generate two hypothetical state merging protocols. Firstly, the state merging protocol that directly transmits the C register to Bob while considering both the A and R registers as reference. In an iid asymptotic setting, this state merging protocol requires quantum communication rate of $\frac{1}{2}I(C;AR)$ and generates ebits between Charlie and Bob at a rate $\frac{1}{2}I(C;B)$. Secondly, if we instead consider the state merging protocol that transmits the C register to Alice, this requires communication of $\frac{1}{2}I(C;RB)$ qubits between Charlie and Alice, and generates $\frac{1}{2}I(C;A)$ ebits between Charlie and Alice. As

Oppenheim noted, this pure state entanglement between Alice and Charlie should not be communicated to Bob. The state redistribution protocol that uses Alice as a coherent relay then runs as follows.

Charlie merges his state with Alice's, generating $\frac{1}{2}I(C;A)$ ebits between them. Alice then replaces these ebits by some pre-shared ebits between her and Bob. This is the ebit repackaging sub-protocol, which effectively acts as a communication of $I(C;A)$ qubits between Charlie and Bob in the direct merging protocol. Alice then transmits the remaining qubits required to complete the direct merging protocol between Charlie and Bob. A communication of

$$\frac{1}{2}I(C;AR) - \frac{1}{2}I(C;A) = \frac{1}{2}I(C;R|B) \tag{3.4.1}$$

is required to achieve this, which is asymptotically optimal (3.1.1). We formalize this idea below while using it in a one-shot setting and expressing the relevant bounds in terms of smooth conditional entropies.

Following the decoupling approach to quantum information theory [58, 59, 70], quantum state merging is conveniently understood in terms of decoupling theorems. Here we first restate the central decoupling theorem of [22] in terms of smooth conditional min-entropy.

Theorem 3.4.1. *[22, Theorem III.1] For $\varepsilon > 0$, $\rho^{AR} \in \mathscr{D}_{\leq}(AR)$, and any decomposition $A = A_1A_2$, if*

$$\log|A_1| \leq \frac{1}{2}\log|A| + \frac{1}{2}H_{\min}(A|R)_\rho - \log\frac{1}{\varepsilon}\,, \tag{3.4.2}$$

then

$$\int_{\mathscr{U}(A)} \left\| \mathrm{Tr}_{A_2}\left[U^{A\to A_1A_2}\left(\rho^{AR}\right)\right] - \pi^{A_1}\otimes\rho^R \right\|_1 dU \leq \varepsilon\,, \tag{3.4.3}$$

where dU is the Haar measure over the unitaries on system A, normalized to $\int dU = 1$, and π^{A_1} is the completely mixed state on A_1.

For our purpose we need the following bi-decoupling result in terms of smooth conditional entropies, a direct generalization of Theorem 3.4.1. [1]

Corollary 3.4.1. *For any* $\varepsilon_1,\ \varepsilon_2 > 0$, $\rho_1^{CR_1} \in \mathscr{D}_{\leq}(CR_1), \rho_2^{CR_2} \in \mathscr{D}_{\leq}(CR_2)$ *and any decomposition* $C = C_1C_2C_3$, *if*

$$\log|C_1| \leq \frac{1}{2}\log|C| + \frac{1}{2}H_{\min}(C|R_1)_{\rho_1} - \log\frac{1}{\varepsilon_1} \tag{3.4.4}$$

and

$$\log|C_2| \leq \frac{1}{2}\log|C| + \frac{1}{2}H_{\min}(C|R_2)_{\rho_2} - \log\frac{1}{\varepsilon_2}\ , \tag{3.4.5}$$

then there exists a unitary $U^{C\to C_1C_2C_3}$ *such that*

$$\left\|\mathrm{Tr}_{C_2C_3}\left[U\left(\rho_1^{CR_1}\right)\right] - \pi^{C_1}\otimes\rho_1^{R_1}\right\|_1 \leq 3\varepsilon_1 \tag{3.4.6}$$

and

$$\left\|\mathrm{Tr}_{C_1C_3}\left[U\left(\rho_2^{CR_2}\right)\right] - \pi^{C_2}\otimes\rho_2^{R_2}\right\|_1 \leq 3\varepsilon_2\ . \tag{3.4.7}$$

Proof. By Markov's inequality, if the condition on $|C_1|$, $|C_2|$ are satisfied, then Theorem 3.4.1 says that the probability over the Haar measure on $\mathscr{U}(C)$ that $\|\mathrm{Tr}_{C_2C_3}[U(\rho_1^{CR_1})] - \pi^{C_1}\otimes\rho_1^{R_1}\|_1 \geq 3\varepsilon_1$ is at most $\frac{1}{3}$, and similarly for $\|\mathrm{Tr}_{C_1C_3}[U(\rho_2^{CR_2})] - \pi^{C_2}\otimes\rho_2^{R_2}\|_1 \geq 3\varepsilon_2$, so by the union bound theorem there is at least probability $\frac{1}{3}$ that none of these is satisfied, and then the condition of the corollary are satisfied for all corresponding U's. □

3.4.2 Smooth Entropy Bounds

We obtain the following direct coding theorem for one-shot quantum state redistribution.

1. A similar bi-decoupling result appears in [141], with bounds in terms of register dimensions instead of smooth conditional entropies. It would be possible to apply ideas similar to theirs to obtain a different coding theorem achieving the same achievability bound (3.1.3) for one-shot quantum state redistribution.

Theorem 3.4.2. *Let ε_1, $\varepsilon_2 \geq 0$, ε_3, $\varepsilon_4 > 0$, and $\rho^{ABC} \in \mathscr{D}(ABC)$ purified by ρ^{ABCR} with purifying register R. Then, there exists a one-message protocol $\Pi \in \mathscr{T}(\mathscr{R}, \rho, \varepsilon')$, with $\varepsilon' = 8\varepsilon_1 + 2\varepsilon_2 + 4\sqrt{3\varepsilon_3} + \sqrt{3\varepsilon_4}$, satisfying*

$$QCC(\Pi) \leq \frac{1}{2} H_{\max}^{\varepsilon_1}(C|B)_\rho - \frac{1}{2} H_{\min}^{\varepsilon_2}(C|BR)_\rho + \log \frac{1}{\varepsilon_3} + \log \frac{1}{\varepsilon_4} + 2 \, . \tag{3.4.8}$$

Moreover, Π only uses EPR states as pre-shared entanglement and also generates EPR pairs. The net entanglement consumption cost $e(\Pi, \rho, \varepsilon')$ satisfies

$$e(\Pi, \rho, \varepsilon') \leq \frac{1}{2} H_{\max}^{\varepsilon_1}(C|B)_\rho + \frac{1}{2} H_{\min}^{\varepsilon_2}(C|BR)_\rho - \log \frac{1}{\varepsilon_3} + \log \frac{1}{\varepsilon_4} + 1 \, . \tag{3.4.9}$$

Proof. We first prove the theorem for the special case $\varepsilon_1 = \varepsilon_2 = 0$. In Corollary 3.4.1, we take $R_1 = BR$, $R_2 = AR$, $\rho_1 = \rho^{CBR}$, $\rho_2 = \rho^{CAR}$,

$$\log |C_1| = \left\lfloor \frac{1}{2} \log |C| + \frac{1}{2} H_{\min}(C|BR)_\rho - \log \frac{1}{\varepsilon_3} \right\rfloor , \tag{3.4.10}$$

$$\log |C_2| = \left\lfloor \frac{1}{2} \log |C| + \frac{1}{2} H_{\min}(C|AR)_\rho - \log \frac{1}{\varepsilon_4} \right\rfloor , \tag{3.4.11}$$

and then there exists a unitary $U^{C \to C_1 C_2 C_3}$ satisfying

$$\left\| \mathrm{Tr}_{C_2 C_3} \left[U \left(\rho^{CBR} \right) \right] - \pi^{C_1} \otimes \rho^{BR} \right\|_1 \leq 3\varepsilon_3 \quad \text{and} \quad \left\| \mathrm{Tr}_{C_1 C_3} \left[U \left(\rho^{CAR} \right) \right] - \pi^{C_2} \otimes \rho^{AR} \right\|_1 \leq 3\varepsilon_4 \, . \tag{3.4.12}$$

We transform these in purified distance bounds using the generalized Fuchs-van der Graaf inequality:

$$P\Big(\mathrm{Tr}_{C_2 C_3} \left[U \left(\rho^{CBR} \right) \right], \pi^{C_1} \otimes \rho^{BR} \Big) \leq \sqrt{3\varepsilon_3}, \tag{3.4.13}$$

$$P\Big(\mathrm{Tr}_{C_1 C_3} \left[U \left(\rho^{CAR} \right) \right], \pi^{C_2} \otimes \rho^{AR} \Big) \leq \sqrt{3\varepsilon_4} \, . \tag{3.4.14}$$

Let A', A'' be isomorphic to A, B''' be isomorphic to B, C', C''' be isomorphic to C, and C_2'', C_3'' be isomorphic to C_2, C_3, respectively. Then, Uhlmann's theorem tells us that

there exist isometries

$$V_1^{C_2C_3A\to A_1A'C'} \quad \text{and} \quad V_2^{C_1C_3B\to B_2B'''C'''} \tag{3.4.15}$$

satisfying

$$P\Big(V_1U\left(\rho^{ABCR}\right),|\phi_1\rangle\langle\phi_1|^{A_1C_1}\otimes I^{AC\to A'C'}\left(\rho^{ABCR}\right)\Big)=P\Big(\operatorname{Tr}_{C_2C_3}\left[U\left(\rho^{CBR}\right)\right],\pi^{C_1}\otimes\rho^{BR}\Big) \tag{3.4.16}$$

$$P\Big(V_2U\left(\rho^{ABCR}\right),|\phi_2\rangle\langle\phi_2|^{B_2C_2}\otimes I^{BC\to B'''C'''}\left(\rho^{ABCR}\right)\Big)=P\Big(\operatorname{Tr}_{C_1C_3}\left[U\left(\rho^{CAR}\right)\right],\pi^{C_2}\otimes\rho^{AR}\Big)\,. \tag{3.4.17}$$

Let T_A, T_B be isomorphic to A_1, C_1, respectively, and denote by

$$\hat{U}^{C\to T_BC_2''C_3''}, \quad \hat{V}_1^{C_2''C_3''A''\to T_AA'C'} \quad \text{and} \quad \hat{V}_2^{T_BC_3''B\to B_2B'''C'''} \tag{3.4.18}$$

a version of U that maps register C into registers $T_BC_2''C_3''$, a version of V_1 that maps registers $C_2''C_3''A''$ into registers $T_AA'C'$, and a version of V_2 that maps registers $T_BC_3''B$ into registers $B_2B'''C'''$, respectively. Also let $\mathscr{M}^{T_AA'C'}$ be a channel performing a projective measurement onto the image of $\hat{V}_1$, and mapping everything outside this image to some fixed state $\phi_{\mathscr{M}}^{T_AA'C'}$ in it. Then there exists an inverse $\hat{V}_1^\dagger$ on the image satisfying $\hat{V}_1^\dagger\hat{V}_1 = \hat{V}_1^\dagger\mathscr{M}\hat{V}_1 = I^{C_2''C_3''A''}$. We can now define our one-shot state redistribution protocol Π (also see Figure 3.1):

Protocol Π for input ρ^{ABCR} using ebits $\phi_1^{T_AT_B}$

1. Charlie applies U on register C, keeps register C_1, and transmits the C_2, C_3 registers to Alice.
2. Alice applies V_1 on C_2C_3A, obtains registers $A_1A'C'$, and then uses T_A instead of A_1; she performs $\mathscr{M}$ on $T_AA'C'$ to apply $\hat{V}_1^\dagger$ and obtains registers $A''C_2''C_3''$.
3. Alice transmits the C_3'' register to Bob.
4. Bob applies $\hat{V}_2$ on $T_BC_3''B$ and obtains registers $B_2B'''C'''$.

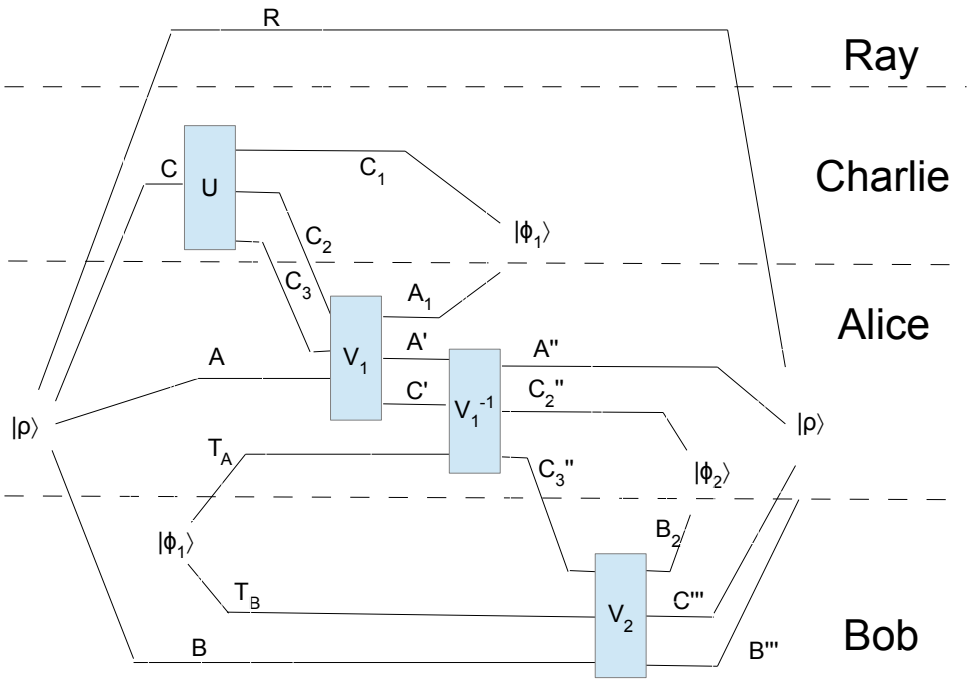

Figure 3.1: Protocol for quantum state redistribution from ebit repackaging.

– The B''', C''' output registers held by Bob correspond to the B, C input registers, respectively, while the A'' output register held by Alice corresponds to the A input register. Together with the untouched reference register R, these should be close to ρ^{ABCR}.

– The A_1C_1 registers should be close to the maximally entangled state $|\phi_1\rangle^{A_1C_1} = I^{T_AT_B\to A_1C_1}|\phi_1\rangle^{T_AT_B}$ shared between Alice and Charlie, while the $C_2''B_2$ registers should be close to the maximally entangled state $|\phi_2\rangle^{C_2''B_2}$ shared between Alice and Bob, with Alice holding the C_2'' share.

Note that Charlie only communicates with Alice, and the only register effectively transmitted between Alice and Bob is the C_3'' register, which is of the same size as the C_3 register. We then have the following bound on the communication:

$$QCC(\Pi) = \log|C_3| = \log|C| - \log|C_2| - \log|C_1| \tag{3.4.19}$$

$$\leq -\frac{1}{2}H_{\min}(C|AR)_\rho - \frac{1}{2}H_{\min}(C|BR)_\rho + \log\frac{1}{\varepsilon_4} + \log\frac{1}{\varepsilon_3} + 2 \tag{3.4.20}$$

$$= \frac{1}{2}H_{\max}(C|B)_\rho - \frac{1}{2}H_{\min}(C|BR)_\rho + \log\frac{1}{\varepsilon_3} + \log\frac{1}{\varepsilon_4} + 2\,. \tag{3.4.21}$$

Note that this protocol is EPR-based, i.e., the only pre-shared entanglement it uses are EPR pairs. The consumption and generation of EPR pairs can also be easily computed from the above dimensions. The consumption is $\log|T_A| = \log|C_1| \leq \frac{1}{2}\log|C| + \frac{1}{2}H_{\min}(C|BR)_\rho - \log\frac{1}{\varepsilon_3}$ EPR pairs, and the number of EPR pairs generated is $\log|C_2''| = \log|C_2| \geq \frac{1}{2}\log|C| + \frac{1}{2}H_{\min}(C|AR)_\rho - \log\frac{1}{\varepsilon_4} - 1$. The net entanglement cost $e(\Pi, \rho, \varepsilon')$

is then bounded by

$$
\begin{aligned}
e(\Pi,\rho,\varepsilon') &= \log|C_1| - \log|C_2| && (3.4.22)\\
&\leq \frac{1}{2}H_{\min}(C|BR)_\rho - \log\frac{1}{\varepsilon_3} - \frac{1}{2}H_{\min}(C|AR)_\rho + \log\frac{1}{\varepsilon_4} + 1 && (3.4.23)\\
&= \frac{1}{2}H_{\max}(C|B)_\rho + \frac{1}{2}H_{\min}(C|BR)_\rho - \log\frac{1}{\varepsilon_3} + \log\frac{1}{\varepsilon_4} + 1\ . && (3.4.24)
\end{aligned}
$$

It is left to verify that the final state is close enough to ρ^{ABCR}. We prove a stronger result, that the global final state is close to $\rho^{ABCR}\otimes\phi_1^{A_1C_1}\otimes\phi_2^{C_2''B_2}$. This is the criteria normally used in EPR-based state-redistribution. We first use the triangle inequality to obtain the following 4 terms:

$$
\begin{aligned}
&P\Big(\hat{V}_2\hat{V}_1^{-1}\mathscr{M}V_1U\big(\rho^{ABCR}\otimes\phi_1^{T_AT_B}\big),\\
&\qquad I^{ABC\to A''B'''C'''}\big(\rho^{ABCR}\otimes\phi_1^{A_1C_1}\otimes\phi_2^{C_2''B_2}\big)\Big)\\
&\leq P\Big(\hat{V}_2\hat{V}_1^{-1}\mathscr{M}V_1U\big(\rho^{ABCR}\otimes\phi_1^{T_AT_B}\big),\\
&\qquad \hat{V}_2\hat{V}_1^{-1}\mathscr{M}I^{AC\to A'C'}\big(\rho^{ABCR}\otimes\phi_1^{A_1C_1}\otimes\phi_1^{T_AT_B}\big)\Big)\\
&\quad + P\Big(\hat{V}_2\hat{V}_1^{-1}\mathscr{M}I^{AC\to A'C'}\big(\rho^{ABCR}\otimes\phi_1^{A_1C_1}\otimes\phi_1^{T_AT_B}\big),\\
&\qquad \hat{V}_2\hat{V}_1^{-1}I^{AC\to A'C'}\big(\rho^{ABCR}\otimes\phi_1^{A_1C_1}\otimes\phi_1^{T_AT_B}\big)\Big)\\
&\quad + P\Big(\hat{V}_2\hat{V}_1^{-1}I^{AC\to A'C'}\big(\rho^{ABCR}\otimes\phi_1^{A_1C_1}\otimes\phi_1^{T_AT_B}\big),\\
&\qquad \hat{V}_2SWAP_{C_1\leftrightarrow T_B}I^{AC_2C_3T_A\to A''C_2''C_3''A_1}U\big(\rho^{ABCR}\otimes\phi_1^{T_AT_B}\big)\Big)\\
&\quad + P\Big(\hat{V}_2SWAP_{C_1\leftrightarrow T_B}I^{AC_2C_3T_A\to A''C_2''C_3''A_1}U\big(\rho^{ABCR}\otimes\phi_1^{T_AT_B}\big),\\
&\qquad I^{ABC\to A''B'''C'''}\big(\rho^{ABCR}\otimes\phi_1^{A_1C_1}\otimes\phi_2^{C_2''B_2}\big)\Big)\ . \qquad (3.4.25)
\end{aligned}
$$

To bound the first term, we have

$$
\begin{aligned}
&P\Big(\hat{V}_2\hat{V}_1^{-1}\mathscr{M}V_1U\big(\rho^{ABCR}\otimes\phi_1^{T_AT_B}\big),\hat{V}_2\hat{V}_1^{-1}\mathscr{M}I^{AC\to A'C'}\big(\rho^{ABCR}\otimes\phi_1^{A_1C_1}\otimes\phi_1^{T_AT_B}\big)\Big)\\
&\leq P\Big(V_1U\big(\rho^{ABCR}\otimes\phi_1^{T_AT_B}\big),I^{AC\to A'C'}\big(\rho^{ABCR}\otimes\phi_1^{A_1C_1}\otimes\phi_1^{T_AT_B}\big)\Big) &(3.4.26)\\
&= P\Big(V_1U\big(\rho^{ABCR}\big),I^{AC\to A'C'}\big(\rho^{ABCR}\otimes\phi_1^{A_1C_1}\big)\Big) &(3.4.27)\\
&\leq \sqrt{3\varepsilon_3}\,. &(3.4.28)
\end{aligned}
$$

The first inequality is by monotonicity of the purified distance, the first equality is because appending an uncorrelated system does not change the distance, and finally the last inequality is by combining (3.4.13) and (3.4.16). For the second term, we have

$$
\begin{aligned}
&P\Big(\hat{V}_2\hat{V}_1^{-1}\mathscr{M}I^{AC\to A'C'}\big(\rho^{ABCR}\otimes\phi_1^{A_1C_1}\otimes\phi_1^{T_AT_B}\big),\\
&\qquad\hat{V}_2\hat{V}_1^{-1}I^{AC\to A'C'}\big(\rho^{ABCR}\otimes\phi_1^{A_1C_1}\otimes\phi_1^{T_AT_B}\big)\Big)\\
&\leq P\Big(\mathscr{M}I^{AC\to A'C'}\big(\rho^{ABCR}\otimes\phi_1^{A_1C_1}\otimes\phi_1^{T_AT_B}\big),I^{AC\to A'C'}\big(\rho^{ABCR}\otimes\phi_1^{A_1C_1}\otimes\phi_1^{T_AT_B}\big)\Big)\\
&\leq P\Big(\mathscr{M}I^{AC\to A'C'}\big(\rho^{ABCR}\otimes\phi_1^{T_AT_B}\big),\mathscr{M}\hat{V}_1\hat{U}I^{A\to A''}\big(\rho^{ABCR}\big)\Big)\\
&\quad+P\Big(\mathscr{M}\hat{V}_1\hat{U}I^{A\to A''}\big(\rho^{ABCR}\big),I^{AC\to A'C'}\big(\rho^{ABCR}\otimes\phi_1^{T_AT_B}\big)\Big)\\
&\leq P\Big(I^{AC\to A'C'}\big(\rho^{ABCR}\otimes\phi_1^{T_AT_B}\big),\hat{V}_1\hat{U}I^{A\to A''}\big(\rho^{ABCR}\big)\Big)\\
&\quad+P\Big(\hat{V}_1\hat{U}I^{A\to A''}\big(\rho^{ABCR}\big),I^{AC\to A'C'}\big(\rho^{ABCR}\otimes\phi_1^{T_AT_B}\big)\Big)\\
&\leq 2\sqrt{3\varepsilon_3}\,. &(3.4.29)
\end{aligned}
$$

The first inequality is by monotonicity of the purified distance, the second by the triangle inequality and because appending an uncorrelated system does not change the distance, the third by monotonicity and because $\mathscr{M}\hat{V}_1=\hat{V}_1$, and finally the last inequality is by combining (3.4.13) and (3.4.16) twice after relabelling systems. For the third term, we

have

$$
\begin{aligned}
&P\Big(\hat{V}_2\hat{V}_1^{-1} I^{AC\to A'C'}\big(\rho^{ABCR}\otimes\phi_1^{A_1C_1}\otimes\phi_1^{T_AT_B}\big),\\
&\quad \hat{V}_2 SWAP_{C_1\leftrightarrow T_B} I^{AC_2C_3T_A\to A''C_2''C_3''A_1} U\big(\rho^{ABCR}\otimes\phi_1^{T_AT_B}\big)\Big)\\
&= P\Big(SWAP_{A_1C_1\leftrightarrow T_AT_B} I^{AC\to A'C'}\big(\rho^{ABCR}\otimes\phi_1^{A_1C_1}\otimes\phi_1^{T_AT_B}\big),\\
&\qquad SWAP_{A_1C_1\leftrightarrow T_AT_B}\hat{V}_1 SWAP_{C_1\leftrightarrow T_B} I^{AC_2C_3T_A\to A''C_2''C_3''A_1} U\big(\rho^{ABCR}\otimes\phi_1^{T_AT_B}\big)\Big) \qquad (3.4.30)\\
&= P\Big(I^{AC\to A'C'}\big(\rho^{ABCR}\otimes\phi_1^{A_1C_1}\otimes\phi_1^{T_AT_B}\big), V_1 U\big(\rho^{ABCR}\otimes\phi_1^{T_AT_B}\big)\Big) \qquad (3.4.31)\\
&= P\Big(I^{AC\to A'C'}\big(\rho^{ABCR}\otimes\phi_1^{A_1C_1}\big), V_1 U\big(\rho^{ABCR}\big)\Big) \qquad (3.4.32)\\
&\leq \sqrt{3\varepsilon_3}\,. \qquad (3.4.33)
\end{aligned}
$$

The first equality is by isometric invariance, the second is because $SWAP_{A_1C_1\leftrightarrow T_AT_B}$ leaves the first state invariant and also because

$$
\begin{aligned}
&SWAP_{A_1C_1\leftrightarrow T_AT_B}\hat{V}_1^{C_2''C_3''A''\to T_AA'C'} SWAP_{C_1\leftrightarrow T_B} I^{AC_2C_3T_A\to A''C_2''C_3''A_1}\\
&= I^{T_AT_BC_1} V_1^{C_2C_3A\to A_1A'C'} \qquad (3.4.34)
\end{aligned}
$$

the next is because appending uncorrelated systems does not change the distance, and finally the last inequality is by combining (3.4.13) and (3.4.16). For the fourth term, we

have

$$
\begin{aligned}
&P\Big(\hat{V}_2 SWAP_{C_1\leftrightarrow T_B} I^{AC_2C_3T_A\to A''C_2''C_3''A_1} U\left(\rho^{ABCR}\otimes\phi_1^{T_AT_B}\right),\\
&\quad I^{ABC\to A''B'''C'''}\left(\rho^{ABCR}\otimes\phi_1^{A_1C_1}\otimes\phi_2^{C_2''B_2}\right)\Big)\\
&= P\Big(I^{A''C_2''\to AC_2}\hat{V}_2 SWAP_{C_1\leftrightarrow T_B} I^{AC_2C_3T_A\to A''C_2''C_3''A_1} U\left(\rho^{ABCR}\otimes\phi_1^{T_AT_B}\right),\\
&\qquad I^{A''C_2''\to AC_2} I^{ABC\to A''B'''C'''}\left(\rho^{ABCR}\otimes\phi_1^{A_1C_1}\otimes\phi_2^{C_2''B_2}\right)\Big) \qquad (3.4.35)\\
&= P\Big(I^{T_AT_B\to A_1C_1} V_2 U\left(\rho^{ABCR}\otimes\phi_1^{T_AT_B}\right),\\
&\qquad I^{BC\to B'''C'''}\left(\rho^{ABCR}\otimes\phi_1^{A_1C_1}\otimes\phi_2^{C_2B_2}\right)\Big) \qquad (3.4.36)\\
&= P\Big(V_2 U\left(\rho^{ABCR}\right),\\
&\qquad I^{BC\to B'''C'''}\left(\rho^{ABCR}\otimes\phi_2^{C_2B_2}\right)\Big) \qquad (3.4.37)\\
&\leq \sqrt{3\varepsilon_4}\,. \qquad (3.4.38)
\end{aligned}
$$

The first equality is just a system relabelling, the second is because

$$
\begin{aligned}
&I^{A''C_2''\to AC_2}\hat{V}_2^{T_BC_3''B\to B_2B'''C'''} SWAP_{C_1\leftrightarrow T_B} I^{AC_2C_3T_A\to A''C_2''C_3''A_1}\\
&= I^{AC_2} I^{T_AT_B\to A_1C_1} V_2^{C_1C_3B\to B_2B'''C'''} \qquad (3.4.39)
\end{aligned}
$$

the third is because appending an uncorrelated system does not change the distance and finally the last inequality is by combining (3.4.14) and (3.4.17). Putting these four bounds together, we get the stated bound for $\varepsilon_1,\varepsilon_2=0$, and this completes the proof for this case.

We can now prove the smooth entropy version of the theorem by extending the above argument to the states achieving the extremum in the smooth entropies. Let $\omega_1^{ABCR}\in\mathscr{D}_{\leq}(ABCR)$ be such that $P(\omega_1,\rho)\leq\varepsilon_1$ and $H_{\min}^{\varepsilon_1}(C|BR)_\rho = H_{\min}(C|BR)_{\omega_1}$. Similarly, let $\omega_2^{ABCR}\in\mathscr{D}_{\leq}(ABCR)$ be such that $P(\omega_2,\rho)\leq\varepsilon_2$ and $H_{\max}^{\varepsilon_2}(C|B)_\rho = H_{\max}(C|B)_{\omega_2}$, and consider a purification $\omega_2^{ABCRS_2}$. In Corollary 3.4.1, we take $R_1=BR$, $R_2=ARS_2$,

$\rho_1 = \omega_1^{CBR}$, $\rho_2 = \omega_2^{CARS_2}$,

$$\log|C_1| = \left\lfloor \frac{1}{2}\log|C| + \frac{1}{2}H_{\min}(C|BR)_{\omega_1} - \log\frac{1}{\varepsilon_3} \right\rfloor \tag{3.4.40}$$

$$\log|C_2| = \left\lfloor \frac{1}{2}\log|C| + \frac{1}{2}H_{\min}(C|ARS_2)_{\omega_2} - \log\frac{1}{\varepsilon_4} \right\rfloor \ , \tag{3.4.41}$$

and then there exists a unitary $U^{C\to C_1C_2C_3}$ satisfying

$$\left\| \mathrm{Tr}_{C_2C_3}\left[U\left(\omega_1^{CBR}\right)\right] - \pi^{C_1}\otimes\omega_1^{BR} \right\|_1 \leq 3\varepsilon_3 \tag{3.4.42}$$

$$\left\| \mathrm{Tr}_{C_1C_3}\left[U\left(\omega_2^{CARS_2}\right)\right] - \pi^{C_2}\otimes\omega_2^{ARS_2} \right\|_1 \leq 3\varepsilon_4 \ . \tag{3.4.43}$$

Transforming these in purified distance bounds, we get

$$P\Big(\mathrm{Tr}_{C_2C_3}\left[U\left(\omega_1^{CBR}\right)\right], \pi^{C_1}\otimes\omega_1^{BR} \Big) \leq \sqrt{3\varepsilon_3} \tag{3.4.44}$$

$$P\Big(\mathrm{Tr}_{C_1C_3}\left[U\left(\omega_2^{CAR}\right)\right], \pi^{C_2}\otimes\omega_2^{AR} \Big) \leq \sqrt{3\varepsilon_4} \ , \tag{3.4.45}$$

in which we also used monotonicity of the purified distance under partial trace of S_2 and the fact that in each purified distance, the two states have the same trace. Since

$$P(\omega_1^{ABCR}, \rho^{ABCR}) \leq \varepsilon_1 \quad \text{and} \quad P(\omega_2^{ABCR}, \rho^{ABCR}) \leq \varepsilon_2 \ , \tag{3.4.46}$$

the triangle inequality along with monotonicity of the purified distance and the fact that appending uncorrelated systems does not increase distance imply the bounds

$$P\Big(\mathrm{Tr}_{C_2C_3}\left[U\left(\rho^{CBR}\right)\right], \pi^{C_1}\otimes\rho^{BR} \Big) \leq \sqrt{3\varepsilon_3} + 2\varepsilon_1 \tag{3.4.47}$$

$$P\Big(\mathrm{Tr}_{C_1C_3}\left[U\left(\rho^{CAR}\right)\right], \pi^{C_2}\otimes\rho^{AR} \Big) \leq \sqrt{3\varepsilon_4} + 2\varepsilon_2 \ . \tag{3.4.48}$$

Considering systems A', A'', B''', C', C''', C_2'', C_3'', T_A, T_B as above, Uhlmann's theorem tells us that there exist partial isometries

$$V_1^{C_2C_3A\to A_1A'C'} \quad \text{and} \quad V_2^{C_1C_3B\to B_2B'''C'''} \tag{3.4.49}$$

satisfying

$$P\Big(V_1 U\left(\rho^{ABCR}\right), |\phi_1\rangle\langle\phi_1|^{A_1C_1} \otimes I^{AC\to A'C'}\left(\rho^{ABCR}\right)\Big) = P\Big(\operatorname{Tr}_{C_2C_3}\left[U\left(\rho^{CBR}\right)\right], \pi^{C_1}\otimes\rho^{BR}\Big) \tag{3.4.50}$$

$$P\Big(V_2 U\left(\rho^{ABCR}\right), |\phi_2\rangle\langle\phi_2|^{B_2C_2} \otimes I^{BC\to B'''C'''}\left(\rho^{ABCR}\right)\Big) = P\Big(\operatorname{Tr}_{C_1C_3}\left[U\left(\rho^{CAR}\right)\right], \pi^{C_2}\otimes\rho^{AR}\Big)\,. \tag{3.4.51}$$

Also consider

$$\hat{U}^{C\to T_B C_2'' C_3''}, \quad \hat{V}_1^{C_2'' C_3'' A''\to T_A A'C'} \quad \text{and} \quad \hat{V}_2^{T_B C_3'' B\to B_2 B'''C'''}, \tag{3.4.52}$$

the versions of U, V_1, V_2 acting on the corresponding registers, as in the ε_1, $\varepsilon_2 = 0$ case, as well as the channel $\mathscr{M}^{T_A A'C'}$ performing a projective measurement onto the image of $\hat{V}_1$ as above. We can then take the smooth version of our one-shot state redistribution protocol Π to be formally defined as the non-smooth version above, but using these U, V_1, $\mathscr{M}$, $\hat{V}_1$, $\hat{V}_2$ instead. We then have the following bound on the communication:

$$\begin{aligned} QCC(\Pi) = \log|C_3| &= \log|C| - \log|C_2| - \log|C_1| && (3.4.53)\\ &\leq -\frac{1}{2}H_{\min}(C|ARS_2)_{\omega_2} - \frac{1}{2}H_{\min}(C|BR)_{\omega_1} + \log\frac{1}{\varepsilon_4} + \log\frac{1}{\varepsilon_3} + 2 && (3.4.54)\\ &= \frac{1}{2}H_{\max}(C|B)_{\omega_2} - \frac{1}{2}H_{\min}(C|BR)_{\omega_1} + \log\frac{1}{\varepsilon_3} + \log\frac{1}{\varepsilon_4} + 2 && (3.4.55)\\ &= \frac{1}{2}H_{\max}^{\varepsilon_2}(C|B)_{\rho} - \frac{1}{2}H_{\min}^{\varepsilon_1}(C|BR)_{\rho} + \log\frac{1}{\varepsilon_3} + \log\frac{1}{\varepsilon_4} + 2\,. && (3.4.56) \end{aligned}$$

Similarly, we have the following bound on the net entanglement cost $e(\Pi, \rho, \varepsilon')$:

$$\begin{aligned} e(\Pi,\rho,\varepsilon') &= \log|C_1| - \log|C_2| && (3.4.57)\\ &\leq \frac{1}{2}H_{\max}^{\varepsilon_2}(C|B)_{\rho} + \frac{1}{2}H_{\min}^{\varepsilon_1}(C|BR)_{\rho} - \log\frac{1}{\varepsilon_3} + \log\frac{1}{\varepsilon_4} + 1\,. && (3.4.58) \end{aligned}$$

Is left to verify that the final state is close enough to ρ^{ABCR}. The analysis is the same

as in the $\varepsilon_1, \varepsilon_2 = 0$ case, with the bounds (3.4.13) and (3.4.14) replaced by (3.4.47) and (3.4.48), yielding the desired bound

$$P\Big(\hat{V}_2\hat{V}^{-1}\mathscr{M}V_1U\left(\rho^{ABCR}\otimes\phi_1^{T_AT_B}\right), I^{ABC\to A''B'''C'''}\left(\rho^{ABCR}\otimes\phi_1^{A_1C_1}\otimes\phi_2^{C_2''B_2}\right)\Big)$$
$$\leq 8\varepsilon_1 + 2\varepsilon_2 + 4\sqrt{3\varepsilon_3} + \sqrt{3\varepsilon_4}\,. \tag{3.4.59}$$

□

Note however that it is possible for the above bound to not be tight in general (at least if we allow arbitrary shared entanglement). This can be seen by considering the situation where the B register is trivial, which corresponds to state splitting, and for which it is known [22] that we can succeed with communication $I_{\max}^{\varepsilon}(C;R)_\rho$ using entanglement embezzling states [130]. Entanglement embezzling states are states from which a variable number of near-perfect EPR pairs can be extracted while only slightly modifying the state. The communication achieved in this way can be much smaller than the bound we provide for some states ρ. We provide an alternate protocol, using entanglement embezzling states rather than standard maximally entangled states, which achieves a communication rate that is upper bounded by the smooth max-information, up to small additive terms, in the case that either the A or the B register is trivial, and so this protocol has optimal communication for the special cases of state merging and state splitting.

The idea for the protocol with embezzling states is borrowed from [22], and is the following. At the outset of the protocol, before applying the above protocol as a subprotocol, we first perform a coherent projective measurement in the eigenbasis of the C system, and discard the portion with eigenvalues smaller than $|C|^2$. To each other measurement outcome, we associate the remaining conditional state to a branch of the computation, with the state in branch i denoted ρ_i. We then coherently apply the above EPR-based protocol on each branch using an entanglement embezzling state between Charlie and Alice, and another between Alice and Bob, to provide the necessary EPR pairs, as well as to absorb any EPR pair created, up to small error. Different amounts of EPR pairs are generated and consumed on each branch, hence the need for entanglement

embezzling states. We also transmit the register containing the coherent measurement outcomes, to make it possible to undo these. This procedure then flattens the eigenvalue spectra on the C system, hence the min- and max-entropies $H^{\varepsilon}_{\min}(C)_{\rho_i}$, $H^{\varepsilon}_{\max}(C)_{\rho_i}$ are both equal to the rank of ρ_i^C, up to a small error. This allows us to replace the max-entropy term by a min-entropy term when the B register is trivial, and similarly when A is trivial, and in such a case we can use the lemmas given in [22] to relate this to smooth max-information, and obtain a provably optimal rate. See [22] for a formal definition of the ρ_i's. In general, the communication grows as

$$\frac{1}{2}\max_i \left(H^{\varepsilon}_{\max}(C|B)_{\rho_i} - H^{\varepsilon}_{\min}(C|BR)_{\rho_i} \right) \tag{3.4.60}$$

up to small additive terms. This is however not optimal in general, and it is still unclear whether this can be of any help for obtaining tight bounds for state redistribution (cf. Section 3.5). An approach that might hold some promise could be to allow for interaction in the state redistribution protocol. For example, in a two-message protocol in which Bob speaks first, this would then allow Bob to also do some preprocessing similar to what Alice does here, and possibly obtain improved flattening in the general case. [2]

3.4.3 Conditional Mutual Information Bounds

The fully quantum asymptotic equipartition property [126, 127] enables us to recover the previously known optimal asymptotic bounds in terms of conditional mutual information.

Theorem 3.4.3 (Fully Quantum Asymptotic Equipartition Property [126, 127])**.** *For any ε, there exists n_0 such that for any $n \geq n_0$ and any state ρ^{AB} with purifying register R, the following holds:*

$$\frac{1}{n} H^{\varepsilon}_{min}(A^{\otimes n}|B^{\otimes n}) \geq H(A|B) - \frac{\delta(\varepsilon, v)}{\sqrt{n}},$$

2. Using the pre-processing from [22] would only amount to a sub-linear communication cost from Bob to Alice, and thus vanishing back communication cost in the iid asymptotic setting.

in which $\delta(\varepsilon, v) = 4 \log v \sqrt{\log(2/\varepsilon^2)}$ *and* $v = \sqrt{2^{H_{max}(A|B)}} + \sqrt{2^{H_{max}(A|R)}} + 1$.

To be able to compress a protocol proportionally to its quantum information cost in Chapter 4, we show how to compress a single message down to a communication cost proportional to its conditional mutual information, as in asymptotic state redistribution. Entanglement is deemed free for the compression.

The idea is to apply the above achievability bound along with the substate theorem of Jain, Radhakrishnan and Sen [78, 80, 83] to obtain a bound on one-shot state redistribution in terms of von Neumann conditional mutual information. This can then be applied iteratively in order to get bounded-round protocol compression proportional to the information cost. Let us first restate the substate theorem in the form that we will use.

Theorem 3.4.4 (Substate theorem [78, 80, 83])**.** *For* $\rho \in \mathscr{D}(A), \sigma \in \mathscr{P}(A)$ *and* $\varepsilon \in (0,1)$,

$$D^{\varepsilon}_{max}(\rho\|\sigma) \leq \frac{1}{\varepsilon^2}\left(D(\rho\|\sigma) + 1\right) + \log\left(1/(1-\varepsilon^2)\right).$$

Note that the square factor here is due to a difference in the distance function used compared to the one of [78] (we use the purified distance, they use its square). The smoothing parameter then changes accordingly. Smoothing is also done over a larger set here, since we allow for subnormalized states, and as a consequence it is possible that smooth max-relative entropy is slightly smaller according to our definition than to the one of Ref. [78]. This is however not an issue, since the smoothing is in the correct direction for the inequality in the substate theorem.

This leads to a lower bound on the conditional min-entropy, or equivalently, by the duality relations, to an upper bound on the conditional max-entropy, in terms of the conditional von Neumann entropy for a normalized state ρ:

Lemma 3.4.1. *For* $\rho \in \mathscr{D}(AB)$ *and* $\varepsilon \in (0,1)$,

$$H^{\varepsilon}_{min}(A|B)_\rho \geq \frac{1}{\varepsilon^2}(H(A|B)_\rho - 1) - \log(1/(1-\varepsilon^2)).$$

Proof. This is a direct consequence of the substate theorem:

$$\begin{aligned}
H_{min}^{\varepsilon}(A|B)_{\rho} &= \sup_{\bar{\rho}^{AB}\in B_{\varepsilon}(\rho^{AB})} H_{min}(A|B)_{\bar{\rho}} \\
&= \sup_{\bar{\rho}^{AB}\in B_{\varepsilon}(\rho^{AB})} (-\inf_{\sigma^{B}\in\mathscr{D}(B)} D_{max}(\bar{\rho}^{AB}\|I^{A}\otimes\sigma_{B}) \\
&\geq \sup_{\bar{\rho}^{AB}\in B_{\varepsilon}(\rho^{AB})} (-D_{max}(\bar{\rho}^{AB}\|I^{A}\otimes\rho^{B})) \\
&= -D_{max}^{\varepsilon}(\rho^{AB}\|I^{A}\otimes\rho^{B}) \\
&\geq -(D(\rho^{AB}\|I^{A}\otimes\rho^{B})+1)/\varepsilon^{2}-\log(1/(1-\varepsilon^{2})) \\
&= (H(A|B)_{\rho}-1)/\varepsilon^{2}-\log(1/(1-\varepsilon^{2})).
\end{aligned}$$

□

We get the following bound in terms of von Neumann conditional mutual information for one-shot state redistribution.

Lemma 3.4.2. *For all $\varepsilon\in(0,1/2),\rho\in\mathscr{D}(ABC)$ with purifying register R, there exists a one-message protocol $\Pi\in\mathscr{T}(\mathscr{R},\rho,\varepsilon)$ with quantum communication satisfying*

$$QCC(\Pi)\leq\frac{61}{\varepsilon^{2}}I(C;R|B)_{\rho}+\frac{242}{\varepsilon^{2}}+16.$$

Proof. We take $\varepsilon_1=\varepsilon_2=\varepsilon/11>0$ and $\varepsilon_3=\varepsilon_4=(\varepsilon_1)^2/75$ in Theorem 3.4.2, so $\varepsilon'=8\varepsilon_1+2\varepsilon_2+4\sqrt{3\varepsilon_3}+\sqrt{3\varepsilon_4}=\varepsilon$. Note that for $x\geq 4$, $2\log x\leq x$, and for $y\in(1,2)$, $\log y\leq$

1. Then

$$\begin{aligned}
QCC(\Pi) \leq & \frac{1}{2} H_{max}^{\varepsilon_1}(C|B)_\rho - \frac{1}{2} H_{min}^{\varepsilon_1}(C|BR)_\rho + 2\log\frac{1}{\varepsilon_3} + 2 \\
= & -\frac{1}{2} H_{min}^{\varepsilon_1}(C|AR)_\rho - \frac{1}{2} H_{min}^{\varepsilon_1}(C|BR)_\rho + 2\log\frac{1}{\varepsilon_1^2} + 2\log 75 + 2 \\
\leq & \frac{1}{2}(-H(C|AR)_\rho + 1)/\varepsilon_1^2 + \frac{1}{2}\log(1/(1-\varepsilon_1^2) \\
& + \frac{1}{2}(-H(C|BR)_\rho + 1)/\varepsilon_1^2 + \frac{1}{2}\log(1/(1-\varepsilon_1^2) + 2\log\frac{1}{\varepsilon_1^2} + 13 + 2 \\
= & \frac{1}{2\varepsilon_1^2} I(C;R|B)_\rho + 1/\varepsilon_1^2 + 2\log\frac{1}{\varepsilon_1^2} + \log(1/(1-\varepsilon_1^2) + 15 \\
\leq & \frac{1}{2\varepsilon_1^2} I(C;R|B)_\rho + 2/\varepsilon_1^2 + 16 \\
= & \frac{61}{\varepsilon^2} I(C;R|B)_\rho + 242/\varepsilon^2 + 16.
\end{aligned}$$

□

3.5 Conclusion

Let us conclude with a discussion of our results and some further research directions.

3.5.1 Discussion

We have proved that one-shot quantum state redistribution of ρ^{ABCR} up to error ε can be achieved at communication cost at most

$$\frac{1}{2}\left(H_{\max}^{\varepsilon}(C|B)_\rho - H_{\min}^{\varepsilon}(C|BR)_\rho\right) + O\left(\log(1/\varepsilon)\right), \tag{3.5.1}$$

when free entanglement assistance is available (independently, this bound has also been derived by Datta, Hsieh and Oppenheim [52]). The structure of the protocol achieving this performs a decomposition of state redistribution into two state merging protocols. Such a decomposition was proposed in [107] in order to achieve asymptotically tight rates. Note that we could alternatively use a decomposition into a state merging and a

state splitting protocol, as proposed in [141], to achieve similar bounds. An important technical ingredient for our proof is the bi-decoupling lemma that we prove as an extension of the well-known decoupling theorem [22]. A similar lemma was derived in [141], with bounds in terms of dimensions rather than conditional min-entropies. This lemma states that for two states on the same system C, there exists at least one unitary on C that acts as a decoupling unitary for both states simultaneously, when parameters are appropriately chosen. Perhaps surprisingly, this idea allows us to smooth both the conditional min- and max- entropy terms appearing in our bounds, notwithstanding the fact that it is in general unknown how to simultaneously smooth marginals of overlapping quantum systems (see, e.g., [57] and references therein).

3.5.2 Open Questions

However, it is known from the work on one-shot state merging and splitting [22] that, for arbitrary shared entanglement, the bound (3.4.8) can in general not be optimal, and in fact for some states the achievable communication can be substantially lower. An interesting open problem is to obtain a tight characterization of the minimal quantum communication cost. Recent works on the Rényi generalizations of conditional mutual information in the quantum regime [24] might shed some light on this question. In particular, it would be of interest to link some version of our improved bound (3.4.60) to a smooth version of the conditional max-information from Ref. [24],

$$I_{\max}(C;R|B)_\rho = D_{\max}\left(\rho^{CBR}\middle\|\left(\rho^{BR}\right)^{1/2}\left(\rho^{B}\right)^{-1/2}\rho^{BC}\left(\rho^{B}\right)^{-1/2}\left(\rho^{BR}\right)^{1/2}\right). \qquad (3.5.2)$$

In turn this would also shine some light on the Rényi generalizations of the conditional mutual information in Ref. [24].

In a recent work, Anshu, Devabathini and Jain [6] obtain tight bounds on quantum state redistribution, up to a small additive term. Another interesting open question, stated in [6], is whether it is possible to achieve quantum state redistribution in a different setting, in which we allow variable length protocols, possibly interactive, and we are interested in the average classical cost of communication when allowing for arbitrary

entanglement assistance. Note that for the classical analogue of this task, Braverman and Rao [36] give such a protocol that has average cost exactly equal to the conditional mutual information, up to second order. However, the protocol in Ref. [36] is highly interactive. Could we hope for something similar in the quantum setting? Note that to obtain such a result even in the much simpler case of source coding would already be quite surprising, due to the fact that this variable-length code would have to leak at most a negligible amount of information about the underlying quantum state.

CHAPTER 4

QUANTUM INFORMATION COMPLEXITY

4.1 Introduction

The area of classical information complexity has been thriving recently. We briefly introduce it, discuss some previous attempts at generalizing it to the quantum setting, and then present the new notion we study in this chapter.

4.1.1 Classical Information Complexity

The paradigm of information complexity has been quite successful recently in classical communication complexity. What started out as a useful tool for proving communication complexity lower bounds has recently developed into an important subfield of its own. The definition of information cost, the sum of the mutual information between the protocol transcript and each party's input conditional on the other party's input, makes it possible to bring powerful tools from information theory to study interactive communication. Many recent results show that this paradigm has enabled researchers to tackle questions that seemed out of reach not so long ago, like an exact characterization by Braverman, Garg, Pankratov, and Weinstein [38] of the communication complexity of the disjointness function, as well as direct sum and direct product results [10, 36, 39, 84]. See Ref. [34] for a recent survey.

The classical notion of information cost was introduced by Chakrabarti, Shi, Wirth and Yao [46], who used it to derive a direct sum result for the simultaneous message passing model. The notion they introduced is similar to what is known today as the external information cost. A notion similar to what is now known as the internal information cost was later introduced by Bar-Yossef, Jayram, Kumar and Sivakumar [9] to take advantage of a direct sum property of information for composite problems that decompose into simpler ones, like the disjointness function in term of the AND function. In Ref. [68], Harsha, Jain, McAllester and Radhakrishnan obtain a direct sum result for

multi-round protocols run on product distributions using a notion of external information cost. The modern notions of external and internal information costs were formally introduced by Barak, Braverman, Chen and Rao [10], in which they prove general direct sum theorems for randomized communication complexity.

For input random variables X and Y of Alice and Bob, respectively, shared randomness R, private randomness S_A, S_B available to Alice and Bob, respectively, and protocol transcript $\Pi(X,Y,R,S_A,S_B)$, the internal and external information costs are defined, respectively, as

$$IC_{int}(\Pi,\mu) = I(X;\Pi|YR) + I(Y;\Pi|XR), \tag{4.1.1}$$

$$IC_{ext}(\Pi,\mu) = I(XY;\Pi|R). \tag{4.1.2}$$

Note that we have used Π to represent both the protocol and the protocol transcript, while μ is the prior distribution on the inputs X,Y. An interpretation of internal information cost is as the amount of information about Alice's input leaked to Bob plus the amount of information about Bob's input leaked to Alice, while for the external information cost it is as the amount of information about the joint input of Alice and Bob leaked to an external observer.

Subsequent work by Braverman and Rao [36] provided an operational interpretation of internal information complexity as the amortized distributional communication complexity, i.e. the communication complexity per copy for computing n copies of a task in parallel, in the asymptotic limit of large n. They also provide a general direct sum theorem for bounded round communication complexity (that does not maintain round complexity however). Braverman [33] provides a similar operational interpretation of a prior-free version of information complexity as the amortized randomized communication complexity. He also lists several interesting open questions related to information complexity, one of which is to develop a quantum analog of information complexity. He asks whether the inherent reversibility of quantum computing, among other properties of quantum information, will impose a limit on the potential applications of such a quantity. Note that our results finally settle this: a notion of quantum information complexity with

a similar operational interpretation and similar potential for applications as the classical one can indeed be defined.

4.1.2 Previous Notions of Quantum Information Cost

In the quantum setting, many difficulties are immediately apparent in trying to generalize the classical definition. Firstly, by the no-cloning theorem [55, 136], there is no direct analogue for quantum communication of the notion of a transcript, available to all parties and containing all previous messages. In the entanglement-assisted model, we can replace quantum communication by twice as much classical communication, by using teleportation [15]. However, if we consider the transcript obtained by replacing quantum communication by classical communication in this way, this transcript will be completely uncorrelated to the corresponding quantum messages and to the inputs. Indeed, the classical messages sent in the teleportation protocol are uniformly random.

A possible way around this might be to try to adapt the classical definition by measuring the correlations between the inputs and the whole state, after reception of each message, of the receiving party, i.e. the Holevo information in each round. We can then even sum over the information contained in all messages. This yield a sensible notion of quantum information cost which is partly classical, and a similar quantity was used by Jain, Radhakrishnan and Sen to obtain a beautiful proof of a lower bound on the bounded round quantum communication complexity of the disjointness function [81]. A further variation on this was used by Jain and Nayak to obtain a lower bound for a variant of the Index function [79]. Works on direct sum results for a single round of communication also consider related notions [7, 77, 82].

However, these partly classical notions of quantum information cost all suffer from the drawback that they are only a lower bound on the communication cost once they have been divided by the number of messages, and already in this sense they do not provide the right quantum generalization of information complexity. Even if these definitions can be successful for obtaining interesting results in a bounded round scenario, it is quite plausible that they are also limited to such applications. Indeed, given these previous definitions of quantum information cost, it is quite easy to find particular in-

puts and protocols with M messages and quantum communication cost C such that the quantum information cost is as large as $M \cdot C$. In hindsight, comparing these with the notion that we introduce, we can say that this is partly due to the fact that they do not take into consideration the available quantum side information. Then, the corresponding notion of quantum information complexity does not have the clear operational interpretation of classical information complexity as the amortized communication complexity, and is probably restricted to applications in bounded round scenarios. However, there is no straightforward way to take into consideration quantum side information; quantum information quantities, and quantum correlations more specifically, often behave counterintuitively, and we will need to make a substantial detour in order to find the appropriate way to account for quantum side information while maintaining quantum correlations in protocols.

4.1.3 Overview of Results

We propose a new, fully quantum notion of quantum information cost, and a corresponding notion of complexity. These are the first fully quantum definitions for such quantities. In particular, the notion of cost applies to arbitrary bipartite quantum protocols that are run on arbitrary bipartite quantum inputs, and the notion of complexity applies to arbitrary quantum tasks on arbitrary quantum input. Of particular interest in the setting of quantum communication complexity that we focus on in this work is the case of quantum protocols implementing classical tasks, e.g. evaluating arbitrary bipartite classical functions or relations on arbitrary bipartite input distributions below a specified error bound. However, the notion could also find applications for fully quantum tasks.

4.1.3.1 Quantum Information Complexity and Amortized Communication

To arrive at the new definition of quantum information complexity, we propose a new interpretation of the classical internal information cost. Indeed, if we view each message generation in a protocol as a channel, then the information cost can be seen to be equal to

the sum of the asymptotic costs of simulating many copies of each such channel with side information at the receiver and feedback to the sender [100], a task related to the reverse Shannon theorem [3, 19, 20, 22, 68, 135], with application in particular in the setting of rate distortion theory with side-information at the receiver [137]. Using known bounds for this task [100], this yields a strengthening of the classical amortized communication result for bounded round complexity [36]. That is, we prove a bounded round variant of the result of Braverman and Rao relating amortized communication and information complexity.

In the fully quantum setting, channel simulation, with side information at the receiver and with environment given as feedback to the sender, is equivalent to the state redistribution task [54, 100, 134, 140]. This insight leads to the new, fully quantum definitions of information cost and complexity, and the link between state redistribution and protocol compression is then apparent. These new definitions are the firsts to simultaneously satisfy all of the properties that we state as desirable for these quantum notions. We prove many important structural properties for them, and in particular we prove the following. Please refer to Section 4.2 for formal definitions of the M-message quantum information complexity QIC^M of a task and of a product task.

Theorem 4.1.1. *For any classical task (T,μ,ε), product task $\otimes_i(T_i,\mu_i,\varepsilon_i)$, protocol Π and number of message M,*

$$\begin{aligned} QIC(\Pi,\mu) &\leq QCC(\Pi), \\ QIC^M(T,\mu,\varepsilon) &= AQCC^M(T,\mu,\varepsilon), \\ QIC^M\big(\bigotimes_{i=1}^{n}(T_i,\mu_i,\varepsilon_i)\big) &= \sum_{i=1}^{n} QIC^M(T_i,\mu_i,\varepsilon_i). \end{aligned}$$

4.1.3.2 Alternative Characterization for Classical Inputs

Our definition of quantum information cost is stated, for an arbitrary bipartite state $\rho^{A_{in}B_{in}}$, as a sum of terms measuring the correlations between the messages C_i and a purifying system R for ρ. While this may seem natural for a quantum input, when the input state is a classical distribution, we could argue that it might be more natural to

measure correlations between the messages and the classical input of the sender of this message, as in classical information cost. We can take R to be a quantum copy of the joint input of Alice and Bob and then expand the corresponding cost for each message to indeed get a term measuring correlations with the classical input of the sender, plus an additional remainder term measuring correlations with the classical input of the receiver. This additional term would always be zero for classical protocols, whereas in general for quantum protocols it can be non-zero. We show that it has an operational interpretation as the amount of information about the receiver's input the sender of the message is forgetting. These results are from a collaboration with Mathieu Laurière [94].

4.1.3.3 Protocol Compression and Direct Sum

We present the first *general* direct sum theorem for quantum communication complexity that holds for more than a single round of communication. A direct sum theorem states that to compute n tasks simultaneously requires as much resources as the amount of the given resource required for computing them separately. By a general direct sum theorem, we mean one that holds for arbitrary relations on arbitrary inputs. The direct sum question, and the related direct product question, are of central importance in the different models of communication complexity, and in computational complexity in general. They have been the subject of a lot of attention in recent years. Many results have been obtained for different models of classical communication complexity (see e.g. Refs [10, 36, 39, 84] and references therein). Progress for quantum communication complexity has been slower, with most results focusing on a single round of communication [7, 12, 77, 82]. Some notable exceptions for the multi-round case are the work of Klauck, Špalek and de Wolf [89] in which they derive a direct product theorem for disjointness, and the works of Shaltiel [117], Lee, Shraibman and Špalek [95], and Sherstov [119] deriving direct product theorems for functions for which particular lower bound methods known as the discrepancy or generalized discrepancy method are tight. Even for a single round of communication, a general direct sum theorem was only proved earlier this year, using techniques much different from ours [7]. Previous to that work, techniques were restricted to proving results for the restricted case of product in-

puts [82]. As a corollary of our results, we also obtain slightly improved parameters for the direct sum theorem of Ref. [7], for the single round case.

To obtain our direct sum theorem, we first prove a protocol compression result stating that we can compress a single copy of a bounded round protocol proportionally to its information cost. An important ingredient in this proof is a single-message one-shot state redistribution protocol. A state redistribution protocol on input state ρ^{ABC}, with the A and C registers initially held by Alice, and the B register held by Bob, is a protocol that effectively transmits the C register to Bob while keeping the overall correlation with a purifying register R, up to some small error ε. We use the new achievability bound in Lemma 3.4.2 for a communication cost proportional to the conditional mutual information, as in asymptotic state redistribution. Our compression protocol applies this single message compression iteratively, and satisfies the following. Recall that we also denote by Π the channel in $\mathscr{C}(A_{in}B_{in}, A_{out}B_{out})$ implemented by protocol Π.

Lemma 4.1.1. *For each M-message protocol Π, input $\rho^{A_{in}B_{in}}$ with purifying register R, and error parameter ε, there exists an M-message compression protocol $\Pi' \in \mathscr{T}(\Pi, \rho, \varepsilon)$ satisfying*

$$QCC(\Pi') \in O\Big(\frac{M^2}{\varepsilon^2}\big(QIC(\Pi,\rho)+M\big)\Big).$$

By combining this protocol compression result with many properties of quantum information complexity in Theorem 4.1.1 above, we can obtain a direct sum theorem for bounded round quantum communication complexity that holds for all quantum tasks. Note that the theorem holds in the model in which we allow for arbitrary pre-shared entanglement. For concreteness, we state the result for classical relations.

Theorem 4.1.2. *For any product task $\otimes(T_i, \varepsilon_i)$, error parameters $\varepsilon' \in (0, 1/2)$ and any number of message M,*

$$QCC^M\big(\bigotimes_{i=1}^{n}(T_i,\varepsilon_i)\big) \in \Omega\Big(\sum_{i=1}^{n}\big((\frac{\varepsilon'}{M})^2 QCC^M(T_i,\varepsilon_i+\varepsilon')-M\big)\Big).$$

4.1.3.4 Bounded Round Disjointness

A further application of the quantum information complexity paradigm is in obtaining tight lower bound for quantum communication complexity of specific functions. We provide such an example by studying the bounded round quantum communication complexity of disjointness. The quantum communication complexity of disjointness, like its classical analogue, has a rich history. While classically a linear amount of communication is required to solve this problem, it was shown by Buhrman, Cleve and Widgerson [42] that a distributed variant of Grover's search algorithm [26, 66] could be applied to obtain a quadratic saving, up to a logarithmic term. This was further improved by Høyer and de Wolf [76], and finally the optimal bound of $O(\sqrt{n})$ was shown to be achievable by Aaronson and Ambainis [1]. Meanwhile, Razborov [110] had proved such a tight lower bound of $\Omega(\sqrt{n})$. But it should be noted that all known protocols were highly interactive, and it was proved by Buhrman and de Wolf [41], based on Nayak's bound on random access codes [102], that any one-message protocol must have linear communication. A natural question is then: what is the quantum communication complexity of disjointness when restricting to protocols with only M messages?

The previously best known lower bound for this problem was from the work of Jain, Radhakrishnan and Sen [81], who derived a bound of $\Omega(n/M^2 + M)$ on the quantum communication complexity of M-message protocols solving disjointness on n bits. They also made the remark that the optimal protocol of Aaronson and Ambainis implies a $O(n/M + M)$ upper bound. Their approach can be seen as using a different notion of quantum information complexity, and reducing the quantum information complexity of disjointness to that of the AND function. They then obtain a $\Omega(1/M)$ lower bound for AND with their notion of quantum information complexity. However, their notion of quantum information cost can be as high as $M \cdot C$ for some protocols communicating a total of C qubits, so they lose a factor of $1/M$ when going from information back to communication.

Using our notion of quantum information complexity, it is possible to obtain a similar reduction from disjointness to AND. Moreover, since our notion is directly a lower

bound on communication, independently of the number of rounds, it seems at first sight that we can get a $1/M$ improvement for a lower bound on the communication complexity for disjointness. However, like the classical information cost, our notion of quantum information cost is defined in terms of a conditional quantum mutual information while theirs was in term of an Holevo information. The conditional quantum mutual information is notoriously hard to lower bound [98], though a recent breakthrough result by Fawzi and Renner [60] might find applications in the context of quantum information complexity. However, such a direct approach has not yet led to an improvement over the bound of Ref. [81]. With Mark Braverman, Ankit Garg, Young Kun Ko and Jieming Mao [40], we have been able to obtain a tight lower bound, up to polylogarithmic terms, of $\tilde{\Omega}(n/M+M)$ with an alternative approach. The idea we use is to reduce back the quantum information complexity of AND to that of disjointness. By showing that the generalized discrepancy method yields a lower bound on the quantum information complexity of binary functions, we can then complete the argument. We give a high level overview of the proof in Section 4.7. Along the way to proving this result, we need to prove many more important properties of quantum information complexity, which appear throughout this chapter.

Organization: This Chapter is structured as follows: in section 4.2, we propose a different perspective on classical information complexity and its link to amortized communication, leading to our definition of quantum information cost and complexity. Section 4.3 states many properties of the newly defined quantities that prove useful in order to obtain the operational interpretation as amortized quantum communication complexity as well as the direct sum theorem, and some more that were developed in Ref. [40] in order to prove the lower bound for disjointness. In Section 4.4, we argue that studying quantum information complexity in the hybrid model is sufficient. In Section 4.5, we present an alternative characterization of quantum information cost when inputs are classical. In Section 4.6, we prove the operational interpretation of quantum information complexity as the amortized quantum communication complexity, and show an additional one-shot compression result leading to our bounded round direct sum theorem. In Section 4.7, we present a high-level overview of the optimal lower bound for bounded

round quantum communication complexity of disjointness. We conclude with a discussion of our results, potential applications and further research directions.

4.2 Definition of Quantum Information Complexity

In this section, we present our new notion of quantum information cost, and the corresponding notion of quantum information complexity. These are the first quantum notions for such quantities to simultaneously possess an additivity property while being direct lower bounds on quantum communication. Quantum information complexity of a task possesses an operational interpretation as the amortized quantum communication complexity of the task, giving a quantum analogue to the result of Braverman and Rao for classical information complexity [36].

4.2.1 A Different Perspective on Classical Information Cost

As we argued in the introduction, the usual definition of classical information cost does not seem to yield itself to a straightforward quantum generalization that would maintain most of its desirable properties. We give an alternate, but equivalent, characterization of the classical information cost for which we can give a quantum generalization that will share a lot of the properties of its classical analogue.

The main difference from the standard definition is not so much in the formal rewriting of this definition, which to some extent was already implicitly used in previous proofs of some properties of the classical information cost [33, 36] and is simply an application of the chain rule and basic properties of conditional mutual information. It is rather in the interpretation of every message transmission as the simulation of a channel (the generation of the message from the input, previous messages, and randomness) with feedback to the sender and side information at the receiver, a variant of the setting of the classical reverse Shannon theorem studied in the information theory literature [19, 20, 68, 100].

Using the same notation as in (4.1.1) for IC_{int} but now distinguishing between the

messages M_i of the transcript Π, we can see that

$$\begin{aligned} IC_{int}(\Pi,\mu) = I(M_1^B;XR^A|YR^B) \\ + I(M_2^A;YM_1^BR^B|XM_1^AR^A) \\ + I(M_3^B;XM_2^AM_1^AR^A|YM_2^BM_1^BR^B)+\cdots, \end{aligned} \tag{4.2.1}$$

in which we distinguish between Alice's and Bob's copy of the public randomness R and messages M_i. Note that this is easily seen to be equivalent to the standard definition for IC_{int} given in the introduction by using the fact that M_i^A, M_i^B and R^A, R^B are just copies of one another. However, the above rewriting has a clear operational interpretation. Indeed, the above characterization comes from viewing each conditional message $C|X$ in the protocol as a channel in which the output C is sent over a noiseless channel to a receiver who has side information S about the input X to the channel, but for which also a copy C_F of the output is given as feedback to the sender. As noted above, the problem of simulating the sending of the output of a channel with feedback has been studied in the literature under the name of classical reverse Shannon theorem [19, 20, 68]. When there is side information S at the receiver, $I(C;X|S)$ characterizes the amount of information that needs to be sent over the noiseless channel from sender to receiver, and it is shown in Ref. [100] that asymptotically, this quantity characterizes the optimal (unidirectional) classical communication rate for this task when sufficient shared randomness is available. In Ref. [36], a rejection sampling protocol is used to perform a similar task in a one-shot setting, but such that the same communication efficiency is achieved on average, up to lower order terms. A caveat is that their protocol to do so is interactive, while the one in Ref. [100] is not. Thus, we can combine the above characterization and the result of Ref. [100] to obtain a simulation protocol for amortized communication that asymptotically achieves communication at the information cost of the protocol, while keeping the same round complexity, and error parameter arbitrarily close to the original one. Details on how to extend this to a direct coding theorem for multi-round protocols that maintains round complexity follow along similar lines as in the quantum setting. Combining this with the direct sum property for information complexity from Ref. [36],

which also maintains round complexity, we get the following theorem, in which IC_{int}^{M} and ACC^{M} are the bounded round internal information complexity and amortized communication complexity for classical protocols, respectively, defined analogously to the quantum quantities.

Theorem 4.2.1. *For a classical task (T, μ, ε) and any bound $M \in \mathbb{N}$ on the number of messages,*

$$ACC^{M}(T, \mu, \varepsilon) = IC_{int}^{M}(T, \mu, \varepsilon).$$

4.2.2 Quantum Information Cost of a Protocol

Given the alternate characterization of the classical information cost in the preceding section, it is possible to define an analogous notion of quantum information cost for quantum protocols. A potential problem is that we cannot keep a copy of a channel input and output at the sender. As already noted, the fully quantum task analogous to channel simulation with feedback to the sender and side information at the receiver is equivalent to quantum state redistribution, and avoids this problem as follows.

4.2.2.1 Quantum State Redistribution

In quantum state redistribution, there are 4 systems of interest. At the outset of the protocol, the A, C systems are in the possession of Alice, and would be for us the coherent feedback of the channel and the output to be transmitted, respectively, while Bob holds the side information B, and the ABC joint system is purified by a reference register R that no party has access to. Thus, the only system changing hands is the C subsystem that is to be transmitted from Alice to Bob. Recall the definition for quantum state redistribution stated in Chapter 3.

Definition 4.2.1. *We say that the bipartite channel $\mathscr{R} \in \mathscr{C}(A_{in}B_{in}, A_{out}B_{out})$ implements* state redistribution *on input $\rho^{A_{in}B_{in}}$, with $A_{in} = AC$, $B_{in} = B$, $A_{out} = A$, $B_{out} = BC$: it implements the identity channel on such a state and such a partition of the input-output*

registers, i.e. it transfers the C part of ρ from Alice to Bob. We say that a protocol Π is an ε-error state redistribution protocol for ρ^{ABC} if $\Pi \in \mathscr{T}(\mathscr{R}, \rho, \varepsilon)$.

It is proved in [54, 140] that this can be accomplished, in the limit of asymptotically many copies of this task and with free entanglement assistance, at a communication cost of $\frac{1}{2}I(R;C|B)$ qubits per copy. The following variant follows from developments in Chapter 3.

Theorem 4.2.2 ([54, 140])**.** *For all $\rho \in \mathscr{D}(ABC)$ and $\delta > 0$, there exists $c > 0$, $n_0 \in \mathbb{N}$ such that for all $n \geq n_0$, there exists a one-message protocol $\Pi_n \in \mathscr{T}(\mathscr{R}^{\otimes n}, \rho^{\otimes n}, 2^{-cn})$ satisfying*

$$QCC(\Pi_n) \leq n\big(I(C;R|B) + \delta\big).$$

4.2.2.2 Quantum Information Cost in the Hybrid Model

Now, in analogy with our rewriting of the classical information cost in (4.2.1), we define the quantum information cost of a protocol, and the corresponding notion of quantum information complexity of a relation, in the following way, by considering the sum of the asymptotic communication costs. Note that throughout this chapter, except when explicitly mentioned, all protocols are in the hybrid model. This is justified in Section 4.4. Please refer to Section 2.4.1 for the definition of a protocol Π and its corresponding registers. Recall that register R is the purifying register for input state $\rho \in \mathscr{D}(A_{in}B_{in})$.

Definition 4.2.2. *For a protocol Π and an input state ρ, we define the* quantum information cost *of Π on input ρ as*

$$QIC(\Pi, \rho) = \sum_{i \geq 1,\ odd} \frac{1}{2} I(C_i;R|B_i) + \sum_{i \geq 1,\ even} \frac{1}{2} I(C_i;R|A_i).$$

Note that even for protocols with communication, the quantum information cost on pure state input is zero, since the purifying register R is trivial in such a case.

Due to the the fact that on pure four-partite states ρ^{ABCR}, it holds that $I(C;R|B) = I(C;R|A)$, we get the following alternate characterizations.

Lemma 4.2.1. *For a protocol* Π *and an input state* ρ*,*

$$\begin{aligned} QIC(\Pi,\rho) &= \sum_{i\geq 1} \frac{1}{2} I(C_i;R|B_i) \\ &= \sum_{i\geq 1} \frac{1}{2} I(C_i;R|A_i). \end{aligned}$$

4.2.2.3 Quantum Information Cost in the Randomized Model

For worst-case classical tasks, $\mathscr{T}(T,\varepsilon)$ and $\mathscr{T}^M(T,\varepsilon)$ are sets of protocols in the randomized model. In fact, Corollary 4.3.1 tells us that it is sufficient to consider protocols in the hybrid model, which are equivalently described as protocols Π_ν in the randomized model with ν having support of size 1. Mainly for convenience in notation, we nevertheless provide a definition valid for all protocols in the randomized model.

Definition 4.2.3. *For a protocol* Π_ν *in the randomized model and an input state* ρ*, we define the* quantum information cost *of* Π_ν *on input* ρ *as*

$$QIC_R(\Pi_\nu,\rho) = \sum_r \nu(r) \cdot QIC(\Pi_r,\rho).$$

4.2.3 Quantum Information Complexity of Classical Tasks

We consider the quantum information complexity of two different kinds of classical tasks: distributional tasks and worst-case tasks.

4.2.3.1 Distributional Quantum Information Complexity

When considering classical inputs in the distributional setting, we noted in Section 2.3.1 that the purification register can be thought of as containing a (quantum) copy of the classical input. The definition of quantum information cost is however invariant under the choice of R and corresponding purification. We define the quantum information complexity for classical tasks in the following way.

Definition 4.2.4. *For a classical task (T,μ,ε) and a number of message M, we define the ε-error* quantum information complexity *of T on input μ as*

$$QIC(T,\mu,\varepsilon) = \inf_{\Pi\in\mathscr{T}(T,\mu,\varepsilon)} QIC(\Pi,\mu),$$

and the M-message, ε-error quantum information complexity *of T on input μ as*

$$QIC^M(T,\mu,\varepsilon) = \inf_{\Pi\in\mathscr{T}^M(T,\mu,\varepsilon)} QIC(\Pi,\mu).$$

For a product task $\bigotimes_i(T_i,\mu_i,\varepsilon_i)$, we have

$$QIC\big(\bigotimes_{i=1}^{n}(T_i,\mu_i,\varepsilon_i)\big) = \inf_{\Pi\in\mathscr{T}(\otimes_i(T_i,\mu_i,\varepsilon_i))} QIC(\Pi,\otimes_i\mu_i),$$
$$QIC^M\big(\bigotimes_{i=1}^{n}(T_i,\mu_i,\varepsilon_i)\big) = \inf_{\Pi\in\mathscr{T}^M(\otimes_i(T_i,\mu_i,\varepsilon_i))} QIC(\Pi,\otimes_i\mu_i).$$

Using properties that are satisfied by these definitions and that are stated in the next section, we get the operational interpretation for quantum information complexity as the amortized quantum communication complexity, i.e. the second statement in Theorem 4.1.1. An unbounded round variant of this result also holds. See Section 4.6.1 for a proof of these statements.

Note that taking an infimum has already been proven to be necessary for the unbounded round definition in the analogous classical context [38]. The reason is that an infinite sequence of protocols, using more and more rounds, might indeed be necessary to asymptotically approach the quantum information complexity, with each message containing an infinitesimal amount of information.

Taking an infimum might also be necessary in the bounded round setting, but for a different reason: an infinite sequence of protocol might also be necessary, with larger and larger entanglement registers. This is somewhat related to the fact that no good bounds are known on the amount of entanglement required for the best protocols; see Ref. [101] and references therein for related discussions.

4.2.3.2 Prior-free Quantum Information Complexity

We want to define a sensible notion of quantum information complexity for prior-free classical tasks. As in the classical setting [33], there are two sensible orderings for the optimization over inputs and protocols. We provide the two corresponding definitions and later investigate the link between them in Section 4.3.2.4: it turns out that we can *almost* reverse the quantifiers.

Definition 4.2.5. *For a classical task (T,ε) and a number of message M, the ε-error* max-distributional quantum information complexity *of T is*

$$QIC_D(T,\varepsilon) = \max_{\mu\in\mathscr{D}_{XY}} QIC(T,\mu,\varepsilon),$$

and the M-message, ε-error max-distributional quantum information complexity *of T is*

$$QIC_D^M(T,\varepsilon) = \max_{\mu\in\mathscr{D}_{XY}} QIC^M(T,\mu,\varepsilon).$$

Definition 4.2.6. *For a classical task (T,ε) and a number of message M, the ε-error* quantum information complexity *of T is*

$$QIC(T,\varepsilon) = \inf_{\Pi\in\mathscr{T}(T,\varepsilon)} \max_{\mu\in\mathscr{D}_{XY}} QIC(\Pi,\mu),$$

and the M-message, ε-error quantum information complexity *of T is*

$$QIC^M(T,\varepsilon) = \inf_{\Pi\in\mathscr{T}^M(T,\varepsilon)} \max_{\mu\in\mathscr{D}_{XY}} QIC(\Pi,\mu).$$

It holds that $QIC_D(T,\varepsilon) \leq QIC(T,\varepsilon)$ and $QIC_D^M(T,\varepsilon) \leq QIC^M(T,\varepsilon)$.

For product tasks, there are two possible sets of distributions over which we can optimize. Given relations T_i with input sets X_i and Y_i, denote $\mathscr{D}^{\times}_{X^nY^n}$ the set of all product distributions over $X^n\times Y^n = (X_1\times\cdots\times X_n)\times(Y_1\times\cdots\times Y_n)$, while $\mathscr{D}_{X^nY^n}$ is the set of all distributions, possibly non-product, over $X^n\times Y^n$. Taking Definition 4.2.6 as our basis, we get the following definitions.

Definition 4.2.7. *For a product classical task $\otimes_i(T_i, \varepsilon_i)$ and a number of message M, the product* quantum information complexity *of $\otimes(T_i, \varepsilon_i)$ is*

$$QIC_{\times}(\otimes_i(T_i, \varepsilon_i)) = \inf_{\Pi_n \in \mathscr{T}(\otimes_i(T_i,\varepsilon_i))} \max_{\mu^n \in \mathscr{D}^{\times}_{X^nY^n}} QIC(\Pi_n, \mu^n),$$

and the M-message product quantum information complexity *of $\otimes_i(T_i, \varepsilon_i)$ is*

$$QIC^M_{\times}(\otimes_i(T_i, \varepsilon_i)) = \inf_{\Pi_n \in \mathscr{T}^M(\otimes_i(T_i,\varepsilon_i))} \max_{\mu^n \in \mathscr{D}^{\times}_{X^nY^n}} QIC(\Pi_n, \mu^n).$$

Definition 4.2.8. *For a product classical task $\otimes_i(T_i, \varepsilon_i)$ and a number of message M, the* quantum information complexity *of $\otimes_i(T_i, \varepsilon_i)$ is*

$$QIC(\otimes_i(T_i, \varepsilon_i)) = \inf_{\Pi_n \in \mathscr{T}(\otimes_i(T_i,\varepsilon_i))} \max_{\mu^n \in \mathscr{D}_{X^nY^n}} QIC(\Pi_n, \mu^n),$$

and the M-message quantum information complexity *of $\otimes_i(T_i, \varepsilon_i)$ is*

$$QIC^M(\otimes_i(T_i, \varepsilon_i)) = \inf_{\Pi_n \in \mathscr{T}^M(\otimes_i(T_i,\varepsilon_i))} \max_{\mu^n \in \mathscr{D}_{X^nY^n}} QIC(\Pi_n, \mu^n).$$

We later study the link between these definitions: optimizing over product or arbitrary distributions is in fact equivalent.

4.2.4 Quantum Information Complexity of Quantum Tasks

For quantum tasks, the notion of quantum information complexity is defined as follows.

Definition 4.2.9. *For a quantum task $(\mathscr{N}, \rho, \varepsilon)$ and a number of message M, we define the ε-error* quantum information complexity *of $\mathscr{N}$ on input ρ as*

$$QIC(\mathscr{N}, \rho, \varepsilon) = \inf_{\Pi \in \mathscr{T}(\mathscr{N},\rho,\varepsilon)} QIC(\Pi, \rho),$$

and the M-message, ε-error quantum information complexity *of $\mathcal{N}$ on input ρ as*

$$QIC^{M}(\mathcal{N},\rho,\varepsilon) = \inf_{\Pi \in \mathcal{T}^{M}(\mathcal{N},\rho,\varepsilon)} QIC(\Pi,\rho).$$

The quantum information complexity of quantum tasks also has an operational interpretation as the amortized quantum communication complexity; see Section 4.6.1.

4.3 Properties of Interactive Quantum Information

We investigate in more detail the properties of the newly defined quantum information cost and complexity. In particular, we show that they are direct lower bounds on communication, that they are asymptotic upper bounds on communication, that they satisfy an additivity as well as a subadditivity property, are convex and continuous in the error parameter, and concave and continuous in the input. We also show that when composing protocols, side-information can only decrease quantum information cost, and provide a first round-independent lower bound on quantum information cost.

Since a variety of quantum tasks can be defined, with error parameter ranging from a classical probability of failure to fully quantum distance notions, we first derive most results for protocols. Such results are usually more versatile; in particular, they can be applied in a similar fashion to prove results about quantum protocols implementing classical and fully quantum tasks. These results about protocols usually maintain the round complexity. For properties of quantum information complexity, we can then obtain bounded-round variants. For worst-case classical tasks, we also prove that we may almost reverse the quantifiers in the definition of quantum information complexity, and that the error can be decreased without increasing the information cost too much.

These lemmata share some similarities with those in foundational works on classical information complexity [33, 36], but there are some particular difficulties associated to the fact that we are handling quantum registers. In particular, we must be careful when conditioning on quantum registers and evaluating quantum information quantities. In general, quantum conditional mutual information does not have the interpretation as an average over possible values of the side information. Moreover, the no-cloning theorem

forbids copying of quantum states, so we can only evaluate information quantities on registers that can be defined at the same moment in time. It then becomes important to keep track of all registers in a purification of a protocol, to split purifications of pre-shared entangled states in the appropriate way, and to properly set them such that we can take classical average when appropriate.

4.3.1 Interactive Protocols

We first study properties of the quantum information cost of protocols. Except when explicitly mentioned, all protocols are in the hybrid model.

4.3.1.1 Quantum Information Lower Bounds Communication

In this section, we make the important remark that in any protocol, the quantum information cost is non-negative and, more importantly, is a lower bound on the quantum communication cost. This holds when considering the quantum information cost with respect to any input state. This follows from the fact that for any quantum state ρ^{BCR}, $0 \leq \frac{1}{2} I(C;R|B) \leq \log|C|$. Applying this to all terms in the quantum information cost versus all terms in the quantum communication cost, we get the result.

Lemma 4.3.1. *For any protocol Π and input state ρ, the following holds:*

$$0 \leq QIC(\Pi, \rho) \leq QCC(\Pi).$$

For protocols in the randomized model, information also lower bounds communication, which can be shown by using a convexity argument and the above result.

Lemma 4.3.2. *For any protocol Π_v in the randomized model and any input state ρ, the following holds:*

$$0 \leq QIC_R(\Pi_v, \rho) \leq QCC(\Pi_v).$$

4.3.1.2 Quantum Information Upper Bounds Amortized Communication

We also state a weak converse result to the above one: the quantum information cost is an upper bound on the amortized quantum communication cost. This is a consequence of the link between our notion of information cost and asymptotic bounds on quantum state redistribution [54, 140].

Lemma 4.3.3. *For any M-message protocol Π, any input state $\rho^{A_{in}B_{in}}$ and any $\delta > 0$, there exists $c > 0$, $n_0 \in \mathbb{N}$ such that for all $n \geq n_0$, there exists a protocol $\Pi_n \in \mathscr{T}^M(\Pi^{\otimes n}, \rho^{\otimes n}, 2^{-cn})$ satisfying*

$$\frac{1}{n} QCC(\Pi_n) \leq QIC(\Pi, \rho) + \delta.$$

Proof. Given any M-message protocol Π and any state $\rho^{A_{in}B_{in}R}$ with purification register R, let $\rho_1^{A_1C_1B_1R} = U_1(\rho \otimes \psi)$, $\rho_2^{A_2C_2B_2R} = U_2(\rho_1), \cdots, \rho_M^{A_MC_MB_MR} = U_M(\rho_{M-1})$. Then, for any $\delta > 0$, take $\varepsilon = 2^{-cn}$ and $Q_i = \frac{1}{2}I(C_i;R|B_i) + \frac{\delta}{2M} = \frac{1}{2}I(C_i;R|A_i) + \frac{\delta}{2M}$, and for each i let n_0^i be the corresponding n_0 for error $\frac{\varepsilon}{M}$ in Theorem 4.2.2, and take $n_0' = \max\{n_0^i\}$ and $n_0 = \max(n_0', \frac{2M}{\delta})$. Then for any $n \geq n_0$, we have a one-message quantum state redistribution protocol $\Pi_i \in \mathscr{T}(\mathscr{R}_i^{\otimes n}, \rho_i^{\otimes n}, \varepsilon/M)$ satisfying $QCC(\Pi_i) = \lceil Q_i n \rceil$. We define the following protocol Π_n starting from the protocol Π. The state ψ is the shared entanglement used in Π, and its isometries are $U_1, U_2, \cdots, U_M, U_{M+1}$. For each i, the state ϕ_i is the shared entanglement used in the quantum state redistribution protocol Π_i, and its isometries are V_1^i, V_2^i. Note that for even i, we will act V_1^i on Bob's side and V_2^i on Alice's side.

Protocol Π_n on input σ in registers $A_{in}^{\otimes n}$, $B_{in}^{\otimes n}$ of $\Pi^{\otimes n}$

- Take entangled state $\hat{\psi} = \psi^{\otimes n} \otimes \phi_1 \otimes \cdots \otimes \phi_M$.
- Take unitaries $\hat{U}_1 = V_1^1 \circ U_1^{\otimes n}$, $\hat{U}_2 = V_1^2 \circ U_2^{\otimes n} \circ V_2^1, \cdots, \hat{U}_M = V_1^M \circ U_M^{\otimes n} \circ V_2^{M-1}$, $\hat{U}_{M+1} = U_{M+1}^{\otimes n} \circ V_2^M$.
- Take as output the $A_{out}^{\otimes n}$, $B_{out}^{\otimes n}$ registers of $\Pi^{\otimes n}$.

Note that the communication cost of Π_n satisfies

$$\begin{aligned}
QCC(\Pi_n) &= \sum_i \log|\hat{C}^i| \\
&= \sum_i \lceil Q_i n \rceil \\
&\leq n\Big(\sum_{i\geq 1} \frac{1}{2} I(C_i;R|B_i) + \frac{M\delta}{2M} + \frac{M}{n}\Big) \\
&\leq n(QIC(\Pi,\rho)+\delta).
\end{aligned}$$

This is also a M-message protocol, so is left to bound the error on input $\sigma = \rho^{\otimes n}$ to make sure that $\Pi_n \in \mathscr{T}(\Pi^{\otimes n}, \rho^{\otimes n}, \varepsilon)$. We have

$$\begin{aligned}
\|\Pi_n(\rho^{\otimes n}) - \Pi^{\otimes n}(\rho^{\otimes n})\| &= \| \operatorname{Tr}_{\neg A_{out}^{\otimes n} B_{out}^{\otimes n}} U_{M+1}^{\otimes n} V_2^M V_1^M U_M^{\otimes n} V_2^{M-1} \cdots V_1^1 U_1^{\otimes n} (\rho^{\otimes n} \otimes \hat{\psi}) \\
&\quad - \operatorname{Tr}_{\neg A_{out}^{\otimes n} B_{out}^{\otimes n}} U_{M+1}^{\otimes n} U_M^{\otimes n} \cdots U_1^{\otimes n} (\rho^{\otimes n} \otimes \psi^{\otimes n}) \| \\
&= \| \operatorname{Tr}_{\neg A_{out}^{\otimes n} B_{out}^{\otimes n}} U_{M+1}^{\otimes n} \Pi_M U_M^{\otimes n} \Pi_{M-1} \cdots \Pi_1 U_1^{\otimes n} (\rho^{\otimes n} \otimes \psi^{\otimes n}) \\
&\quad - \operatorname{Tr}_{\neg A_{out}^{\otimes n} B_{out}^{\otimes n}} U_{M+1}^{\otimes n} U_M^{\otimes n} \cdots U_1^{\otimes n} (\rho^{\otimes n} \otimes \psi^{\otimes n}) \| \\
&\leq \| \operatorname{Tr}_{(A')^{\otimes n}(B')^{\otimes n}} U_{M+1}^{\otimes n} \Pi_M \cdots \Pi_2 U_2^{\otimes n} \Pi_1 (\rho_1^{\otimes n}) \\
&\quad - \operatorname{Tr}_{(A')^{\otimes n}(B')^{\otimes n}} U_{M+1}^{\otimes n} \Pi_M \cdots \Pi_2 U_2^{\otimes n} (\rho_1^{\otimes n}) \| \\
&\quad + \| \operatorname{Tr}_{(A')^{\otimes n}(B')^{\otimes n}} U_{M+1}^{\otimes n} \Pi_M U_M^{\otimes n} \Pi_{M-1} \cdots \Pi_3 U_3^{\otimes n} \Pi_2 (\rho_2^{\otimes n}) \\
&\quad - \operatorname{Tr}_{(A')^{\otimes n}(B')^{\otimes n}} U_{M+1}^{\otimes n} \Pi_M U_M^{\otimes n} \Pi_{M-1} \cdots \Pi_3 U_3^{\otimes n} (\rho_2^{\otimes n}) \| \\
&\quad + \cdots \\
&\quad + \| \operatorname{Tr}_{(A')^{\otimes n}(B')^{\otimes n}} U_{M+1}^{\otimes n} \Pi_M U_M^{\otimes n} \Pi_{M-1} (\rho_{M-1}^{\otimes n}) \\
&\quad - \operatorname{Tr}_{(A')^{\otimes n}(B')^{\otimes n}} U_{M+1}^{\otimes n} \Pi_M U_M^{\otimes n} (\rho_{M-1}^{\otimes n}) \| \\
&\quad + \| \operatorname{Tr}_{(A')^{\otimes n}(B')^{\otimes n}} U_{M+1}^{\otimes n} \Pi_M (\rho_M^{\otimes n}) \\
&\quad - \operatorname{Tr}_{(A')^{\otimes n}(B')^{\otimes n}} U_{M+1}^{\otimes n} (\rho_M^{\otimes n}) \| \\
&\leq \|\Pi_1(\rho_1^{\otimes n}) - (\rho_1^{\otimes n})\| \\
&\quad + \|\Pi_2(\rho_2^{\otimes n}) - (\rho_2^{\otimes n})\| \\
&\quad + \cdots
\end{aligned}$$

$$
\begin{aligned}
&+ \|\Pi_{M-1}(\rho_{M-1}^{\otimes n}) - (\rho_{M-1}^{\otimes n})\| \\
&+ \|\Pi_{M}(\rho_{M}^{\otimes n}) - (\rho_{M}^{\otimes n})\| \\
&\leq M\frac{\varepsilon}{M} \\
&= \varepsilon.
\end{aligned}
$$

The first equality is by definition, the second one by taking the channel view for the protocols Π_i, since the corresponding A'_i, B'_i left at the end of these are traced out, and only U_1^i, U_2^i act on ϕ^i, the first inequality is by the triangle inequality and by definition of the ρ_i's, the second inequality is due to the monotonicity of trace distance under noisy channels, and the next is because $\Pi_i \in \mathscr{T}(\mathscr{R}_i^{\otimes n}, \rho_i^{\otimes n}, \varepsilon/M)$. □

4.3.1.3 Additivity

The quantum information complexity of tasks satisfies an exact direct sum property on product inputs. This follows from two technical lemmata about the additivity of quantum information cost of protocols. Remember that for a protocol Π, we also denote by Π the channel implemented by the protocol.

Lemma 4.3.4. *For any M-message protocol Π and any input states $\rho_1 \in \mathscr{D}(A_{in}^1 B_{in}^1)$ and $\rho_2 \in \mathscr{D}(A_{in}^2 B_{in}^2)$, there exist M-message protocols Π^1, Π^2 satisfying for all σ_1 and σ_2*

$$
\begin{aligned}
\Pi^1(\sigma_1) &= \mathrm{Tr}_{A_{out}^2 B_{out}^2} \circ \Pi(\sigma_1 \otimes \rho_2), \\
\Pi^2(\sigma_2) &= \mathrm{Tr}_{A_{out}^1 B_{out}^1} \circ \Pi(\rho_1 \otimes \sigma_2), \\
QIC(\Pi^1, \rho_1) + QIC(\Pi^2, \rho_2) &= QIC(\Pi, \rho_1 \otimes \rho_2).
\end{aligned}
$$

Proof. Given Π, we define the protocols Π^1, Π^2 in the following way. Let $\psi^{T_A T_B}$ be the entangled state used in the Π protocol.

Protocol Π^1 on input σ^1

1. Let $\rho_2^{A_{in}^2 B_{in}^2 R^2}$ be a purification of ρ_2. The entangled state for the protocol will

be $\rho_2 \otimes \psi$, with the A^2_{in}, R^2, T_A registers given to Alice, and the B^2_{in}, T_B registers given to Bob.

2. Using the ρ_2 state given as pre-shared entanglement to simulate the other input, run protocol Π on input $\sigma_1 \otimes \rho_2$.

3. Take as output the $A^1_{out}B^1_{out}$ output registers.

Protocol Π^2 on input σ^2

1. Let $\rho_1^{A^1_{in}B^1_{in}R^1}$ be a purification of ρ_1. The entangled state for the protocol will be $\rho_1 \otimes \psi$, with the A^1_{in}, T_A registers given to Alice, and the B^1_{in}, R^1, T_B registers given to Bob.

2. Using the ρ_1 state given as pre-shared entanglement to simulate the other input, run protocol Π on input $\rho_1 \otimes \sigma_2$.

3. Take as output the $A^2_{out}B^2_{out}$ output registers.

By the definitions of protocols Π^1 and Π^2, the channels they implement are $\Pi^1(\sigma_1) = \mathrm{Tr}_{A^2_{out}B^2_{out}} \circ \Pi(\sigma_1 \otimes \rho_2)$ and $\Pi^2(\sigma_2) = \mathrm{Tr}_{A^1_{out}B^1_{out}} \circ \Pi(\rho_1 \otimes \sigma_2)$, respectively. Also, Π^1 and Π^2 are M-message protocols, so is left to analyse their quantum information costs on input ρ_1, ρ_2, respectively. Note that R_1R_2 is a purifying register for $\rho_1 \otimes \rho_2$. By definition and the structure of the protocols,

$$\begin{aligned} 2QIC(\Pi^1, \rho_1) &= I(C_1; R^1|B_0) + I(C_2; R^1|A_1R^2) + \cdots, \\ 2QIC(\Pi^2, \rho_2) &= I(C_1; R^2|B_0R^1) + I(C_2; R^2|A_1) + \cdots, \end{aligned}$$

and we get by rearranging terms and using the chain rule

$$\begin{aligned} 2QIC(\Pi^1, \rho_1) + 2QIC(\Pi^2, \rho_2) &= I(C_1; R^1R^2|B_0) + I(C_2; R^1R^2|A_1) + \cdots \\ &= 2QIC(\Pi, \rho_1 \otimes \rho_2). \end{aligned}$$

□

Lemma 4.3.5. *For any two protocols Π^1 and Π^2 with M_1 and M_2 messages, respectively, there exists an M-message protocol Π, satisfying $\Pi = \Pi^1 \otimes \Pi^2$, $M = \max(M_1, M_2)$, such that the following holds for any input states ρ^1, ρ^2:*

$$QIC(\Pi, \rho^1 \otimes \rho^2) = QIC(\Pi^1, \rho^1) + QIC(\Pi^2, \rho^2).$$

Proof. Given protocols Π^1 and Π^2, we assume without loss of generality that $M_1 \geq M_2$, and we define the protocol Π in the following way.

Protocol Π on input σ

1. Take as input the registers A^1_{in}, B^1_{in}, A^2_{in}, B^2_{in} of both Π^1 and Π^2
2. Run protocols Π^1, Π^2 in parallel for M_2 messages on corresponding input registers until Π^2 has finished.
3. Finish running protocol Π^1.
4. Take as output the registers A^1_{out}, B^1_{out}, A^2_{out}, B^2_{out} of both Π^1 and Π^2.

By the definition of protocol Π, the channel it implements is $\Pi = \Pi^1 \otimes \Pi^2$, and the number of messages satisfies $M = \max(M_1, M_2)$, so is left to analyse its quantum information cost on input $\sigma = \rho_1 \otimes \rho_2$. The first thing to notice is that we can find a purification of $\rho_1 \otimes \rho_2$ that is also in product form, i.e. there exists a purification with the purifying system $R = R^1 \otimes R^2$ and such that $(\rho_1 \otimes \rho_2)^{A^1_{in}B^1_{in}A^2_{in}B^2_{in}R} = \rho_1^{A^1_{in}B^1_{in}R^1} \otimes \rho_2^{A^2_{in}B^2_{in}R^2}$. Also note that throughout the protocol, due to the structure of Π_2 and the fact that the input $\rho_1 \otimes \rho_2$ is in product form, any registers corresponding to Π^1 stays in product form

with any register corresponding to Π^2. Then

$$\begin{aligned}
2QIC(\Pi, \rho_1 \otimes \rho_2) &= I(C_1^1 C_1^2; R^1 R^2 | B_1^1 B_1^2) + I(C_2^1 C_2^2; R^1 R^2 | A_2^1 A_2^2) \\
&\quad + \cdots + I(C_{M_2}^1 C_{M_2}^2; R^1 R^2 | A_{M_2}^1 A_{M_2}^2) \\
&\quad + I(C_{M_2+1}^1; R^1 R^2 | B_{M_2}^1 B_{M_2}^2) + \cdots + I(C_{M_1}^1; R^1 R^2 | A_{M_1}^1 A_{M_2}^2) \\
&= I(C_1^1; R^1 | B_1^1) + I(C_2^1; R^1 | A_2^1) + \cdots + I(C_{M_1}^1; R^1 | A_{M_1}^1) \\
&\quad + I(C_1^2; R^2 | B_1^2) + I(C_2^2; R^2 | A_2^2) + \cdots \\
&= 2QIC(\Pi^1, \rho_1) + 2QIC(\Pi^2, \rho_2).
\end{aligned}$$

The first equality is by definition of quantum information cost of Π, and due to its parallel structure, the second equality is because registers of Π^1, Π^2 are in product form, and then the last equality follows from definition and the structure of Π. □

In general, inputs are not necessarily in product form. Then, the following subadditivity result holds.

Lemma 4.3.6 (Subadditivity)**.** *For any two protocols Π^1, Π^2 with M_1, M_2 messages, respectively, there exists a M-message protocol Π, satisfying $\Pi = \Pi^1 \otimes \Pi^2$, $M = \max(M_1, M_2)$, such that the following holds for any joint input state $\rho_{12} \in \mathscr{D}(A_{in}^1 B_{in}^1 A_{in}^2 B_{in}^2)$:*

$$QIC(\Pi, \rho_{12}) \leq QIC(\Pi^1, \rho_1) + QIC(\Pi^2, \rho_2),$$

with $\rho_1 = \mathrm{Tr}_{A_{in}^2 B_{in}^2}(\rho_{12})$ and $\rho_2 = \mathrm{Tr}_{A_{in}^1 B_{in}^1}(\rho_{12})$.

Proof. The protocol Π is as in Lemma 4.3.5: simply run Π^1, Π^2 in parallel.

Protocol Π on input σ

1. Take as input the registers A_{in}^1, B_{in}^1, A_{in}^2, B_{in}^2 of both Π^1 and Π^2
2. Run protocols Π^1, Π^2 in parallel for M_2 messages on corresponding input registers until Π^2 has finished.
3. Finish running protocol Π^1.

4. Take as output the registers A^1_{out}, B^1_{out}, A^2_{out}, B^2_{out} of both Π^1 and Π^2.

By the definition of Π, the channel that it implements is $\Pi = \Pi^1 \otimes \Pi^2$, and the number of messages satisfies $M = \max(M_1, M_2)$, so is left to analyse its quantum information cost on input ρ_{12}. Let R_{12} be a purifying register such that $\rho_{12}^{A^1_{in}B^1_{in}A^2_{in}B^2_{in}R_{12}}$ is a pure state. Also, denote the purified joint state in round i as $(\rho^i_{12})^{A^1_iB^1_iC^1_iA^2_iB^2_iC^2_iR_{12}}$, and the local state for protocol Π^1 as

$$(\rho^i_1)^{A^1_iB^1_iC^1_i} = \mathrm{Tr}_{A^2_iB^2_iC^2_iR_{12}}((\rho^i_{12})^{A^1_iB^1_iC^1_iA^2_iB^2_iC^2_iR_{12}}), \tag{4.3.1}$$

and similarly for that of protocol Π^2. Notice that for all i, $(\rho^i_1)^{A^1_iB^1_iC^1_i}$ is purified by $(\rho^i_1)^{A^1_iB^1_iC^1_iA^2_{in}B^2_{in}R_{12}} \otimes \phi_2^{T^2_AT^2_B}$, with $A^2_{in}B^2_{in}R_{12}$ the registers of state ρ_{12} before application of the unitaries corresponding to Π^1, and ϕ_2 the pure entangled state used in Π^2. If we denote, for $i \geq M_2+1$, $A^2_i = A^2_{M_2}$, $B^2_i = B^2_{M_2}$, then by the definition of QIC and application of chain rule,

$$\begin{aligned}
2\cdot QIC(\Pi,\rho_{12}) = &\sum_{i=1,\ i\ odd}^{M_2} I(C^1_iC^2_i;R_{12}|B^1_iB^2_i)_{\rho_{12}} + \sum_{i=1,\ i\ even}^{M_2} I(C^1_iC^2_i;R_{12}|A^1_iA^2_i)_{\rho_{12}} \\
&+ \sum_{i=M_2+1,\ i\ odd}^{M_1} I(C^1_i;R_{12}|B^1_iB^2_i)_{\rho_{12}} + \sum_{i=M_2+1,\ i\ even}^{M_1} I(C^1_i;R_{12}|A^1_iA^2_i)_{\rho_{12}} \\
= &\sum_{i=1,\ i\ odd}^{M_2} I(C^2_i;R_{12}|B^1_iB^2_iC^1_i)_{\rho_{12}} + \sum_{i=1,\ i\ even}^{M_2} I(C^2_i;R_{12}|A^1_iA^2_iC^1_i)_{\rho_{12}} \\
&+ \sum_{i=1,\ i\ odd}^{M_1} I(C^1_i;R_{12}|B^1_iB^2_i)_{\rho_{12}} + \sum_{i=1,\ i\ even}^{M_1} I(C^1_i;R_{12}|A^1_iA^2_i)_{\rho_{12}}.
\end{aligned}$$

Now for protocol Π^1, as noted above, the registers $A^2_{in}B^2_{in}R_{12}T^2_AT^2_B$ purify $(\rho^i_1)^{A^1_iB^1_iC^1_i}$ for all i, so

$$
\begin{aligned}
2\cdot QIC(\Pi^1,\rho_1) &= \sum_{i=1,\ i\ odd}^{M_1} I(C_i^1;A_{in}^2B_{in}^2R_{12}T_A^2T_B^2|B_i^1)_{\rho_1} + \sum_{i=1,\ i\ even}^{M_1} I(C_i^1;A_{in}^2B_{in}^2R_{12}T_A^2T_B^2|A_i^1)_{\rho_1} \\
&= \sum_{i=1,\ i\ odd}^{M_1} I(C_i^1;A_i^2B_i^2C_i^2R_{12}|B_i^1)_{\rho_{12}} + \sum_{i=1,\ i\ even}^{M_1} I(C_i^1;A_i^2B_i^2C_i^2R_{12}|A_i^1)_{\rho_{12}} \\
&= \sum_{i=1,\ i\ odd}^{M_1} I(C_i^1;B_i^2|B_i^1)_{\rho_{12}} + \sum_{i=1,\ i\ even}^{M_1} I(C_i^1;A_i^2|A_i^1)_{\rho_{12}} \\
&\quad + \sum_{i=1,\ i\ odd}^{M_1} I(C_i^1;R_{12}|B_i^1B_i^2)_{\rho_{12}} + \sum_{i=1,\ i\ even}^{M_1} I(C_i^1;R_{12}|A_i^1A_i^2)_{\rho_{12}} \\
&\quad + \sum_{i=1,\ i\ odd}^{M_1} I(C_i^1;A_i^2C_i^2|B_i^1B_i^2R_{12})_{\rho_{12}} + \sum_{i=1,\ i\ even}^{M_1} I(C_i^1;B_i^2C_i^2|A_i^1A_i^2R_{12})_{\rho_{12}} \\
&\geq \sum_{i=1,\ i\ odd}^{M_1} I(C_i^1;R_{12}|B_i^1B_i^2)_{\rho_{12}} + \sum_{i=1,\ i\ even}^{M_1} I(C_i^1;R_{12}|A_i^1A_i^2)_{\rho_{12}},
\end{aligned}
$$

in which the first equality is by definition, the second is by isometric invariance of the conditional quantum mutual information (CQMI), the third by the chain rule for CQMI, and the inequality is by non-negativity of CQMI. Similarly for protocol Π^2, with a slightly different application of the chain rule, we get

$$
\begin{aligned}
2\cdot QIC(\Pi^2,\rho_2) &= \sum_{i=1,\ i\ odd}^{M_2} I(C_i^2;A_{in}^1B_{in}^1R_{12}T_A^1T_B^1|B_i^2)_{\rho_2} + \sum_{i=1,\ i\ even}^{M_2} I(C_i^2;A_{in}^1B_{in}^1R_{12}T_A^1T_B^1|A_i^2)_{\rho_2} \\
&= \sum_{i=1,\ i\ odd}^{M_2} I(C_i^2;A_i^1B_i^1C_i^1R_{12}|B_i^2)_{\rho_{12}} + \sum_{i=1,\ i\ even}^{M_2} I(C_i^2;A_i^1B_i^1C_i^1R_{12}|A_i^2)_{\rho_{12}} \\
&= \sum_{i=1,\ i\ odd}^{M_2} I(C_i^2;B_i^1C_i^1|B_i^2)_{\rho_{12}} + \sum_{i=1,\ i\ even}^{M_2} I(C_i^2;A_i^1C_i^1|A_i^2)_{\rho_{12}} \\
&\quad + \sum_{i=1,\ i\ odd}^{M_2} I(C_i^2;R_{12}|B_i^1B_i^2C_i^1)_{\rho_{12}} + \sum_{i=1,\ i\ even}^{M_2} I(C_i^2;R_{12}|A_i^1A_i^2C_i^1)_{\rho_{12}} \\
&\quad + \sum_{i=1,\ i\ odd}^{M_2} I(C_i^2;A_i^1|B_i^1B_i^2C_i^1R_{12})_{\rho_{12}} + \sum_{i=1,\ i\ even}^{M_2} I(C_i^2;B_i^2|A_i^1A_i^2C_i^1R_{12})_{\rho_{12}} \\
&\geq \sum_{i=1,\ i\ odd}^{M_2} I(C_i^2;R_{12}|B_i^1B_i^2C_i^1)_{\rho_{12}} + \sum_{i=1,\ i\ even}^{M_2} I(C_i^2;R_{12}|A_i^1A_i^2C_i^1)_{\rho_{12}}.
\end{aligned}
$$

The result then follows by comparing terms. □

4.3.1.4 Convexity, Concavity and Continuity

We now show that quantum information cost is convex in the protocol parameter, concave in the input state parameter, and also continuous in the input state.

Lemma 4.3.7 (Convexity in Protocol)**.** *For any $p \in [0,1]$, any two protocols Π^1, Π^2 with M_1, M_2 messages, respectively, there exists an M-message protocol Π satisfying $\Pi = p\Pi^1 + (1-p)\Pi^2$, $M = \max(M_1, M_2)$, such that the following holds for any state ρ:*

$$QIC(\Pi, \rho) = pQIC(\Pi^1, \rho) + (1-p)QIC(\Pi^2, \rho).$$

Proof. Given Π^1, Π^2, we assume without loss of generality that $M_1 \geq M_2$, and we define Π in the following way:

Protocol Π on input ρ

1. The entangled state ψ contains many parts: it contains both entangled states ψ_1, ψ_2 for the corresponding protocols Π^1, Π^2, it contains selector registers in state $\left|\sigma_p\right\rangle = \sqrt{p}\left|1\right\rangle^{S_A}\left|1\right\rangle^{S_B} + \sqrt{1-p}\left|2\right\rangle^{S_A}\left|2\right\rangle^{S_B}$, and it contains padding pure states to feed as input to the protocol that is not selected, held in registers D_A, D_B.
2. Coherently control, according to the selector registers, what to input into the two protocols: on control set to 1, input state ρ into protocol Π^1 and the padding pure state into protocol Π^2 , and vice-versa on control set to 2.
3. Run protocols Π^1, Π^2 in parallel for M_2 messages on given input until Π^2 has finished.
4. Finish running protocol Π^1.
5. Coherently control what to output: on control set to 1, take as output the A_{out}, B_{out} registers of protocol Π^1, and on control set to 2, take those of protocol Π^2.

Note that by the structure of the above protocol and because the selector registers are traced out at the end, the output is of the form $\Pi(\rho) = p\Pi^1(\rho) + (1-p)\Pi^2(\rho)$, and Π is an M-message protocol. We must now verify that the quantum information cost satisfies the stated property. First note that if Alice's registers are traced out, then Bob's selector register is effectively a classical register, and similarly for Alice's selector register if Bob's registers are traced out. Also note that throughout the protocol, due to the structure of Π, conditional on some classical state of the selector register, the reference register R can only be correlated with registers in the corresponding protocol Π^1 or Π^2. Also, still conditional on some classical state of the selector register, any register corresponding to Π^1 is in product form with any register corresponding to Π^2. Then

$$\begin{aligned}
2QIC(\Pi,\rho) &= I(C_1^1C_1^2;R|B_{in}D_BT_B^1T_B^2S_B) + I(C_2^1C_2^2;R|A_2^1A_2^2S_A) \\
&\quad + \cdots + I(C_{M_2}^1C_{M_2}^2;R|A_{M_2}^1A_{M_2}^2S_A) \\
&\quad + I(C_{M_2+1}^1;R|B_{M_2+1}^1B_{M_2}^2S_B) + \cdots + I(C_{M_1}^1;R|A_{M_1}^1A_{M_2}^2S_A) \\
&= p(I(C_1^1C_1^2;R|B_{in}D_BT_B^1T_B^2(S_B=1)) + I(C_2^1C_2^2;R|A_2^1A_2^2(S_A=1)) \\
&\quad + \cdots + I(C_{M_2}^1C_{M_2}^2;R|A_{M_2}^1A_{M_2}^2(S_A=1)) \\
&\quad + I(C_{M_2+1}^1;R|B_{M_2+1}^1B_{M_2}^2(S_B=1)) + \cdots + I(C_{M_1}^1;R|A_{M_1}^1A_{M_2}^2(S_A=1))) \\
&\quad + (1-p)(I(C_1^1C_1^2;R|B_{in}D_BT_B^1T_B^2(S_B=2)) + I(C_2^1C_2^2;R|A_2^1A_2^2(S_A=2)) \\
&\quad + \cdots + I(C_{M_2}^1C_{M_2}^2;R|A_{M_2}^1A_{M_2}^2(S_A=2)) \\
&\quad + I(C_{M_2+1}^1;R|B_{M_2+1}^1B_{M_2}^2(S_B=2)) + \cdots + I(C_{M_1}^1;R|A_{M_1}^1A_{M_2}^2(S_A=2))) \\
&= p(I(C_1^1;R|B_{in}^1T_B^1(S_B=1)) + I(C_2^1;R|A_2^1(S_A=1)) + \cdots + I(C_{M_1}^1;R|A_{M_1}^1(S_A=1))) \\
&\quad + (1-p)(I(C_1^2;R|B_{in}^2T_B^2(S_B=2)) + I(C_2^2;R|A_2^2(S_A=2)) + \cdots) \\
&= p\cdot 2QIC(\Pi^1,\rho) + (1-p)\cdot 2QIC(\Pi^2,\rho).
\end{aligned}$$

The first equality is by definition of quantum information cost of Π, and due to its parallel structure, the second equality uses the above remark about the selector register of one party being classical when the registers of the other party are traced out, along with

a convex rewriting of conditional mutual information, the third equality uses the above remark about the product structure of R and the registers corresponding to Π^1, Π^2, respectively, depending on the classical state of the selector register, and the last equality is due to the fact that conditional on some classical state of the selector register, the state in the registers considered is the same as the one in the corresponding protocol. □

In light of the above lemma, it holds that in the context of quantum information complexity, considering protocols in the hybrid model instead of protocols in the randomized model is sufficient. We have the following corollary.

Corollary 4.3.1. *For any M-message protocol Π_ν in the randomized model, there is an M-message protocol Π in the hybrid model such that for any ρ it holds that*

$$\Pi = \Pi_\nu,$$
$$QIC(\Pi, \rho) = QIC_R(\Pi_\nu, \rho).$$

Proof. If ν has support of size 1, then this is immediate. If it has support of size 2, the protocol from Lemma 4.3.7 will do. Otherwise, use the lemma recursively. □

We now show that quantum information is concave in its input state parameter, while a subsequent lemma gives a bound on how far it is from being convex.

Lemma 4.3.8 (Concavity in input)**.** *For any $p \in [0,1]$, define $\rho = p \cdot \rho_1 + (1-p) \cdot \rho_2$ for any two input states ρ_1, ρ_2. Then the following holds for any protocol Π:*

$$QIC(\Pi, \rho) \geq p\, QIC(\Pi, \rho_1) + (1-p)\, QIC(\Pi, \rho_2).$$

Proof. Let R be a register holding a purification of ρ_1 and ρ_2, then we can purify ρ with two copies S_1, S_2 of a selector reference register, such that

$$|\rho\rangle^{A_{in}B_{in}RS_1S_2} = \sqrt{p}\,|\rho_1\rangle^{A_{in}B_{in}R}\,|1\rangle^{S_1}\,|1\rangle^{S_2} + \sqrt{1-p}\,|\rho_2\rangle^{A_{in}B_{in}R}\,|2\rangle^{S_1}\,|2\rangle^{S_2}.$$

We can then expand each term in the quantum information cost as

$$I(C_i;RS_1S_2|B_i)_\rho = I(C_i;S_1|B_i)_\rho + I(C_i;R|B_iS_1)_\rho + I(C_i;S_2|B_iRS_1)_\rho.$$

The result follows by summing over all rounds since

$$I(C_i;R|B_iS_1)_\rho = pI(C_i;R|B_i)_{\rho_1} + (1-p)I(C_i;R|B_i)_{\rho_2},$$

and the remainder terms are non-negative. □

Lemma 4.3.9 (Quasi-convexity in input)**.** *For any $p \in [0,1]$, define $\rho = p\cdot\rho_1 + (1-p)\cdot\rho_2$ for any two input states ρ_1, ρ_2. Then the following holds for any M-message protocol Π:*

$$QIC(\Pi,\rho) \leq pQIC(\Pi,\rho_1) + (1-p)QIC(\Pi,\rho_2) + MH(p).$$

Proof. The result follows from the proof of the above lemma and by noting that $H(S_1) = H(S_2) = H(p)$ upper bounds the two remainder terms in each of the M messages. Indeed, S_1 is classical if S_2 is traced out, and vice versa, so $I(C_i;S_1|B_i)_\rho \leq H(S_1)$ and $I(C_i;S_2|B_iRS_1)_\rho = I(C_i;S_2|A_i)_\rho \leq H(S_2)$.

□

Note that since entropy is concave, i.e. $H(B|X) \leq H(B)$, the above lemma can also be extended to arbitrary convex combinations of states. Combining the two previous lemmata, we can prove continuity in the input.

Lemma 4.3.10 (Continuity in input)**.** *The quantum information cost for M-message protocols is uniformly continuous in the input state. This holds uniformly over all M-message protocols with input $\mathscr{D}(A_{in}B_{in})$. That is, for all M, A_{in}, B_{in}, and $\varepsilon > 0$, there exists $\delta \in (0,1/2)$ such that for all ρ_1 and ρ_2 that are δ-close and all M-message protocols Π,*

$$|QIC(\Pi,\rho_1) - QIC(\Pi,\rho_2)| \leq \varepsilon.$$

Proof. Since we are using the trace distance as a measure of distance, this follows by finding the closest purification between ρ_1 and ρ_2. This can be done up to error $O(\sqrt{\delta})$ by using the Fuchs-van de Graaf inequalities (see Eq. 3.2.17) and the characterization of fidelity as the maximum overlap between purifications. Remember that quantum information cost is invariant under the choice of purification. Then continuity of the conditional quantum mutual information M-times yields $\varepsilon \in O(M \cdot \sqrt{\delta} \cdot (\log|A_{in}| + \log|B_{in}|) + MH(\sqrt{\delta}))$. This bound is independent of μ_1, μ_2, depends on Π only through M and $\log|A_{in}|$, $\log|B_{in}|$, and goes to zero as δ does, so the result follows.

However, this is not sufficient for an application of continuity needed in the result about disjointness in Section 4.7, due to the $\sqrt{\delta}$ terms appearing instead of δ: see Lemma 4.7.6. We prove a tighter result, with $\sqrt{\delta}$ replaced by δ.

Let's start with the case $\rho_1 = (1-\delta')\rho_2 + \delta'\sigma$, for $\delta' \in (0, 1/2)$ and some state σ. Note that this case is sufficient for classical inputs: for δ-close distributions we can always find such a rewriting. Then by the lemma about quasi-convexity in the input, we get $QIC(\Pi, \rho_1) \leq (1-\delta')QIC(\Pi, \rho_2) + \delta' QIC(\Pi, \sigma) + MH(\delta')$. By concavity in the input and non-negativity, we also get $(1-\delta')QIC(\Pi, \rho_2) \leq QIC(\Pi, \rho_1)$. It follows that $|QIC(\Pi, \rho_1) - QIC(\Pi, \rho_2)| \leq \delta' \cdot M \cdot (\log|A_{in}| + \log|B_{in}|) + MH(\delta')$, since $H(R) \leq \log|A_{in}| + \log|B_{in}|$.

Now, for general ρ_1 and ρ_2 that are δ-close, we use a trick from Alicki and Fannes [5]. Write $\theta = (1-\delta)\rho_1 + |\rho_1 - \rho_2|$, $\sigma_1 = \frac{1}{\delta}|\rho_1 - \rho_2|$, $\sigma_2 = \frac{1-\delta}{\delta}(\rho_1 - \rho_2) + \frac{1}{\delta}|\rho_1 - \rho_2|$, with $|O| = (O^\dagger O)^{1/2}$ for an operator O. It then holds that $\theta = (1-\delta)\rho_1 + \delta\sigma_1 = (1-\delta)\rho_2 + \delta\sigma_2$. The desired result then follows by the triangle inequality and the special case above: $|QIC(\Pi, \rho_1) - QIC(\Pi, \rho_2)| \leq |QIC(\Pi, \rho_1) - QIC(\Pi, \theta)| + |QIC(\Pi, \theta) - QIC(\Pi, \rho_2)| \leq 2\delta \cdot M \cdot (\log|A_{in}| + \log|B_{in}|) + 2MH(\delta)$. □

4.3.1.5 Composition of Protocols

The following lemmata show that when running a protocol as a subroutine, classical side information can be conditioned on, and more generally quantum side information can be safely discarded without increasing quantum information cost.

Definition 4.3.1. *Let $\rho^{A_{in}B_{in}O_AO_B}$ be a state with purification register R and of the form*

$$|\rho\rangle^{A_{in}B_{in}O_AO_BR} = \sum_o p_O(o)\,|o\rangle^{O_A}\,|o\rangle^{O_B}\,|\rho_o\rangle^{A_{in}B_{in}R},$$

for some distribution p_O and states ρ_o. Also let Π be an M-message protocol acting on input registers A_{in}, B_{in}. Then we define the quantum information cost *of Π on ρ* conditional *on O as*

$$QIC(\Pi,\rho|O) = \sum_{i>0,\ odd} I(C_i;R|B_iO_B) + \sum_{i>0,\ even} I(C_i;R|A_iO_A).$$

The next lemma follows directly from the definition and Eq. (2.2.29).

Lemma 4.3.11. *In the setting of Definition 4.3.1,*

$$QIC(\Pi,\rho|O) = \sum_o p_O(o) QIC(\Pi,\rho_o).$$

The next lemma follows from definitions and the fact that we can implement an identity channel with the trivial protocol that does not communicate at all.

Lemma 4.3.12. *In the setting of Definition 4.3.1, define $\sigma^{A'_{in}B'_{in}} = \rho^{A_{in}B_{in}O_AO_B}$, with $A'_{in} = A_{in}O_A$ and $B'_{in} = B_{in}O_B$. Then there exists an M-round protocol Π' satisfying*

$$\Pi' = \Pi \otimes I^{O_AO_B},$$
$$QIC(\Pi',\sigma) = QIC(\Pi,\rho|O).$$

In the case of arbitrary quantum side information, we get the lemma below. It follows from subadditivity, the fact that quantum information cost is invariant under the choice of purification and the fact that we can implement an identity channel with the trivial protocol that does not communicate at all.

Lemma 4.3.13. *Let Π be an M-round protocol acting on input registers A_{in}, B_{in}, let $\rho^{A_{in}B_{in}}$ and $\sigma^{A'_{in}B'_{in}}$ be states with $A'_{in} = A_{in}\tilde{A}$ and $B'_{in} = B_{in}\tilde{B}$ for some arbitrary finite-dimensional registers $\tilde{A}$, $\tilde{B}$, and such that $\mathrm{Tr}_{\tilde{A}\tilde{B}}(\sigma) = \rho$. Then there exists an M-message*

protocol Π' *satisfying*

$$\Pi' = \Pi \otimes I^{\tilde{A}\tilde{B}},$$
$$QIC(\Pi', \sigma) \leq QIC(\Pi, \rho).$$

4.3.1.6 Round-Independent Lower Bound

The link between each term in the quantum information cost and the task of quantum state redistribution suggests that the following relation should hold for the information Bob has about R:

$$\sum_{i\geq 0} I(R;C_{2i+1}|B_{2i+1}) - \sum_{i\geq 1} I(R;C_{2i}|B_{2i}) = I(R;B_{out}B') - I(R;B_{in}).$$

We prove the following stronger result.

Lemma 4.3.14. *Given a protocol* Π*, a state* ρ *with purifying register* R *having an arbitrary decomposition* $R = R_a R_b R_c$*, the following holds:*

$$\sum_{i\geq 0} I(R_a;C_{2i+1}|R_b B_{2i+1}) - \sum_{i\geq 1} I(R_a;C_{2i}|R_b B_{2i}) = I(R_a;B_{out}B'|R_b) - I(R_a;B_{in}|R_b).$$

Proof. We will prove that

$$I(R_a;B_{2k+1}C_{2k+1}|R_b) = \sum_{0\leq i\leq k} I(R_a;C_{2i+1}|R_b B_{2i+1}) - \sum_{1\leq i\leq k} I(R_a;C_{2i}|R_b B_{2i}) + I(R_a;B_{in}|R_b)$$

by induction on k. If M is odd, the result follows since Bob receives the last message and $I(R_a;B_{out}B'|R_b) = I(R_A;B_M C_M|R_b)$. If M is even and Bob sends the last message, the result follows since $B_M = B_{out}B'$ and $I(R_a;B_M|R_b) = I(R_a;B_{M-1}C_{M-1}|R_b) - I(R_a;C_M|R_b B_M)$ by using the chain rule and isometric invariance for $B_{M-1}C_{M-1} \to B_M C_M$. Now for the induction, the base case follows from

$$\begin{aligned} I(R_a;B_1C_1|R_b) &= I(R_a;B_1|R_b) + I(R_a;C_1|R_b B_1) \\ &= I(R_a;B_{in}|R_b) + I(R_a;C_1|R_b B_1), \end{aligned}$$

in which the first equality is by the chain rule and the second is because $B_1 = B_0 = B_{in}T_B$ along with the fact that T_B is product to R_a, R_b and B_{in}.

Now for the induction step, we have

$$\begin{aligned} I(R_a;B_{2k+3}C_{2k+3}|Y) &= I(R_a;B_{2k+3}|R_b) + I(R_a;C_{2k+2}|R_bB_{2k+2}) \\ &\quad + I(R_a;C_{2k+3}|R_bB_{2k+3}) - I(R_a;C_{2k+2}|R_bB_{2k+2}) \\ &= I(R_a : B_{2k+2}C_{2k+2}|R_b) + I(R_a;C_{2k+3}|R_bB_{2k+3}) \\ &\quad - I(R_a;C_{2k+2}|R_bB_{2k+2}) \\ &= I(R_a : B_{2k+1}C_{2k+1}|R_b) + I(R_a;C_{2k+3}|R_bB_{2k+3}) \\ &\quad - I(R_a;C_{2k+2}|R_bB_{2k+2}), \end{aligned}$$

in which the first equality is by the chain rule and also by adding and subtracting the same term, the second is also by the chain rule and because $B_{2k+3} = B_{2k+2}$, and the third is by isometric invariance on $B_{2k+1}C_{2k+1} \to B_{2k+2}C_{2k+2}$. The induction step follows by comparing terms. □

A similar result holds for Alice.

Corollary 4.3.2. *Given a protocol Π, a state ρ with purifying register R having an arbitrary decomposition $R = R_aR_bR_c$, the following holds:*

$$\sum_{i\geq 0} I(R_a;C_{2i+2}|R_bA_{2i+2}) - \sum_{i\geq 0} I(R_a;C_{2i+1}|R_bA_{2i+1}) = I(R_a;A_{out}A'|R_b) - I(R_a;A_{in}|R_b).$$

Combining the above two results, we get the following lower bound on quantum information cost, stated as a difference between the amount of correlations of the reference with the output and the input.

Corollary 4.3.3. *Given a protocol Π, a state ρ with purifying register R having an*

arbitrary decomposition $R = R_a R_b R_c$, the following holds:

$$\begin{aligned}
QIC(\Pi,\rho) &\geq I(R_a;A_{out}A'|R_b) - I(R_a;A_{in}|R_b),\\
QIC(\Pi,\rho) &\geq I(R_a;B_{out}B'|R_b) - I(R_a;B_{in}|R_b),\\
QIC(\Pi,\rho) &\geq I(R_a;A_{out}A'|R_b) - I(R_a;A_{in}|R_b)\\
&\quad + I(R_a;B_{out}B'|R_b) - I(R_a;B_{in}|R_b).
\end{aligned}$$

Note that the first two inequalities are not subsumed by the third: the difference on the right hand side could be negative in these.

4.3.2 Classical Tasks

From the properties about the quantum information cost of protocols we proved in the previous section, we can derive many corresponding properties for classical tasks.

4.3.2.1 Quantum Information Lower Bounds Communication

The result that quantum information complexity lower bounds quantum communication complexity follows by taking infimum on both sides in Lemma 4.3.1.

Lemma 4.3.15. *For any product task $\otimes_i(T_i,\mu_i,\varepsilon_i)$ and number of message M,*

$$\begin{aligned}
0 \leq QIC(\otimes_i(T_i,\mu_i,\varepsilon_i)) \leq QCC(\otimes_i(T_i,\mu_i,\varepsilon_i)),\\
0 \leq QIC^M(\otimes_i(T_i,\mu_i,\varepsilon_i)) \leq QCC^M(\otimes_i(T_i,\mu_i,\varepsilon_i)).
\end{aligned}$$

4.3.2.2 Additivity

Quantum information complexity satisfies an exact direct sum property, a consequence of the following additivity property. For two product tasks $S_1 = \bigotimes_{i=1}^{s_1}(T_i,\mu_i,\varepsilon_i)$, $S_2 = \bigotimes_{j=s_1+1}^{s_1+s_2}(T_j,\mu_j,\varepsilon_j)$, define the product task $S_1 \otimes S_2 = \bigotimes_{k=1}^{s_1+s_2}(T_k,\mu_k,\varepsilon_k)$. We get the following result.

Lemma 4.3.16. *For any two product classical tasks* $S_1 = \bigotimes_{i=1}^{s_1}(T_i, \mu_i, \varepsilon_i), S_2 = \bigotimes_{j=s_1+1}^{s_1+s_2}(T_j, \mu_j, \varepsilon_j)$ *and any bound* $M \in \mathbb{N}$ *on the number of messages,*

$$QIC(S_1 \otimes S_2) = QIC(S_1) + QIC(S_2),$$
$$QIC^M(S_1 \otimes S_2) = QIC^M(S_1) + QIC^M(S_2).$$

Proof. We first prove the $\leq$ direction. Let Π^1 and Π^2 be protocols succeeding at the corresponding tasks S_1, S_2, and achieving, for an arbitrary small $\delta > 0$, $QIC(\Pi^1, \mu_1 \otimes \cdots \otimes \mu_{s_1}) \leq QIC(S_1) + \delta, QIC(\Pi^2, \mu_{s_1+1} \otimes \cdots \otimes \mu_{s_1+s_2}) \leq QIC(S_2) + \delta$, respectively. Taking the corresponding protocol Π from Lemma 4.3.5, its induced channel succeeds at the product task $S_1 \otimes S_2$, and we get

$$\begin{aligned} QIC(S_1 \otimes S_2) &\leq QIC(\Pi, \mu_1 \otimes \cdots \otimes \mu_{s_1} \otimes \mu_{s_1+1} \otimes \cdots \otimes \mu_{s_1+s_2}) \\ &= QIC(\Pi^1, \mu_1 \otimes \cdots \otimes \mu_{s_1}) + QIC(\Pi^2, \mu_{s_1+1} \otimes \cdots \otimes \mu_{s_1+s_2}) \\ &\leq QIC(S_1) + QIC(S_2) + 2\delta. \end{aligned}$$

Now for the $\geq$ direction, let Π be a protocol succeeding at the product task and achieving $QIC(\Pi, \mu_1 \otimes \cdots \otimes \mu_{s_1} \otimes \mu_{s_1+1} \otimes \cdots \otimes \mu_{s_1+s_2}) \leq QIC(S_1 \otimes S_2) + \delta$ for an arbitrary small $\delta > 0$. Taking the corresponding protocols Π^1, Π^2 from Lemma 4.3.4 for tasks S_1, S_2, their induced channels succeed at their respective task, and we get

$$\begin{aligned} &QIC(S_1) + QIC(S_2) \\ &\quad \leq QIC(\Pi^1, \mu_1 \otimes \cdots \otimes \mu_{s_1}) + QIC(\Pi^2, \mu_{s_1+1} \otimes \cdots \otimes \mu_{s_1+s_2}) \\ &\quad = QIC(\Pi, \mu_1 \otimes \cdots \otimes \mu_{s_1} \otimes \mu_{s_1+1} \otimes \cdots \otimes \mu_{s_1+s_2}) \\ &\quad \leq QIC(S_1 \otimes S_2) + \delta. \end{aligned}$$

Keeping tracks of rounds, we also get the bounded round result. □

An induction then yields the following corollary.

Corollary 4.3.4. *For a product task* $\otimes_i(T_i, \mu_i, \varepsilon_i)$ *and any bound* $M \in \mathbb{N}$ *on the number*

of messages, the following holds:

$$QIC(\otimes_i(T_i,\mu_i,\varepsilon_i)) = \sum_i QIC(T_i,\mu_i,\varepsilon_i),$$
$$QIC^M(\otimes_i(T_i,\mu_i,\varepsilon_i)) = \sum_i QIC^M(T_i,\mu_i,\varepsilon_i).$$

Combining the above corollary with Lemma 4.3.15, we get the following.

Corollary 4.3.5. *For a product task $\otimes_i(T_i,\mu_i,\varepsilon_i)$ and any bound $M \in \mathbb{N}$ on the number of messages, the following holds:*

$$\sum_i QIC(T_i,\mu_i,\varepsilon_i) \leq QCC(\otimes_i(T_i,\mu_i,\varepsilon_i)),$$
$$\sum_i QIC^M(T_i,\mu_i,\varepsilon_i) \leq QCC^M(\otimes_i(T_i,\mu_i,\varepsilon_i)).$$

4.3.2.3 Convexity and Continuity

We now show that quantum information complexity is convex in the error parameter. A corollary is that it is continuous in the error parameter. Recall that quantum information complexity is non-increasing in the error parameter.

Lemma 4.3.17. *For any T, μ, ε_1, ε_2, M_1, M_2 and $p \in [0,1]$, if we define $\varepsilon = p\varepsilon_1 + (1-p)\varepsilon_2$, $M = \max(M_1,M_2)$, then the following holds:*

$$QIC(T,\mu,\varepsilon) \leq pQIC(T,\mu,\varepsilon_1) + (1-p)QIC(T,\mu,\varepsilon_2),$$
$$QIC^M(T,\mu,\varepsilon) \leq pQIC^{M_1}(T,\mu,\varepsilon_1) + (1-p)QIC^{M_2}(T,\mu,\varepsilon_2).$$

Proof. Let Π^1 and Π^2 be protocols satisfying, for $i \in \{1,2\}$, $\Pi^i \in \mathscr{T}(T,\mu,\varepsilon_i), QIC(\Pi^i,\mu) \leq QIC(T,\mu,\varepsilon_i) + \delta$ for an arbitrary small $\delta > 0$, and take the corresponding protocol Π of Lemma 4.3.7. By linearity of expectation, protocol Π successfully accomplishes its task of achieving average error ε. We must now verify that the quantum information cost

satisfies the convexity property:

$$\begin{aligned} QIC(T,\mu,\varepsilon) &\leq QIC(\Pi,\mu) \\ &= pQIC(\Pi^1,\mu) + (1-p)QIC(\Pi^2,\mu) \\ &\leq pQIC(T,\mu,\varepsilon_1) + (1-p)QIC(T,\mu,\varepsilon_2) + \delta. \end{aligned}$$

Keeping track of messages, we get the bounded round result. □

Lemma 4.3.18 (Continuity in average error)**.** *Quantum information complexity is continuous in the error. This holds uniformly in the input. That is, for all T, M and ε, $\delta > 0$, there exists $\varepsilon' \in (0,\varepsilon)$ such that for all $\varepsilon'' \in (\varepsilon',\varepsilon)$ and all μ,*

$$\begin{aligned} |QIC(T,\mu,\varepsilon'') - QIC(T,\mu,\varepsilon)| &\leq \delta, \\ |QIC^M(T,\mu,\varepsilon'') - QIC^M(T,\mu,\varepsilon)| &\leq \delta. \end{aligned}$$

Proof. Note that we can drop the absolute values since quantum information complexity is non-increasing in the error, i.e. $QIC(T,\mu,\varepsilon) \leq QIC(T,\mu,\varepsilon'') \leq QIC(T,\mu,\varepsilon')$. Using Lemma 4.3.17 with $\varepsilon_1 = 0, \varepsilon_2 = \varepsilon, \varepsilon' = p\varepsilon$ for the current ε, we get

$$\begin{aligned} QIC(T,\mu,\varepsilon'') &\leq QIC(T,\mu,\varepsilon') \\ &\leq pQIC(T,\mu,0) + (1-p)QIC(T,\mu,\varepsilon) \\ &\leq pQCC(T,0) + QIC(T,\mu,\varepsilon). \end{aligned}$$

Rearranging terms, we get

$$|QIC(T,\mu,\varepsilon'') - QIC(T,\mu,\varepsilon)| \leq \frac{\varepsilon'}{\varepsilon} QCC(T,0).$$

This bound is independent of μ, and goes to zero as p and ε' do, so the result follows. The bounded round result is proved in the same way, obtaining $QCC^M(T,0)$ in the final bound instead. □

4.3.2.4 Prior-free Quantum Information Complexity

Many similar properties hold for worst-case tasks: quantum information complexity lower bounds quantum communication complexity, it is convex in the error parameter, and it is also continuous in the error parameter. In Theorem 4.3.1, we also prove that we may almost reverse the quantifiers on protocols and inputs. First, we prove that information lower bounds communication.

Lemma 4.3.19 (Information lower bounds communication)**.** *For any worst-case product task $\otimes_i(T_i, \varepsilon_i)$ and number of message M,*

$$0 \leq QIC(\otimes_i(T_i, \varepsilon_i)) \leq QCC(\otimes_i(T_i, \varepsilon_i)),$$
$$0 \leq QIC^M(\otimes_i(T_i, \varepsilon_i)) \leq QCC^M(\otimes_i(T_i, \varepsilon_i)).$$

Proof. Let Π be a protocol implementing product task $\otimes_i(T_i, \varepsilon_i)$ and satisfying $QCC(\Pi) = QCC(\otimes_i(T_i, \varepsilon_i))$. We get the result by noting that $QIC(\otimes_i(T_i, \varepsilon_i)) \leq \max_{\mu^n \in \mathscr{D}_{X^n Y^n}} QIC(\Pi, \mu^n) \leq QCC(\Pi)$. The bounded round result follows in the same way, by keeping tracks of messages. □

For two product tasks $S_1 = \bigotimes_{i=1}^{s_1}(T_i, \varepsilon_i), S_2 = \bigotimes_{j=s_1+1}^{s_1+s_2}(T_j, \varepsilon_j)$, define the product task $S_1 \otimes S_2 = \bigotimes_{k=1}^{s_1+s_2}(T_k, \varepsilon_k)$. We get the following additivity result, similar to Theorem 4.2 in Ref. [33].

Lemma 4.3.20. *For any two product classical tasks $S_1 = \bigotimes_{i=1}^{s_1}(T_i, \varepsilon_i)$, $S_2 = \bigotimes_{j=s_1+1}^{s_1+s_2}(T_j, \varepsilon_j)$ and any bound $M \in \mathbb{N}$ on the number of messages,*

$$\begin{aligned} QIC(S_1 \otimes S_2) &= QIC_{\times}(S_1 \otimes S_2) \\ &= QIC(S_1) + QIC(S_2), \\ QIC^M(S_1 \otimes S_2) &= QIC^M_{\times}(S_1 \otimes S_2) \\ &= QIC^M(S_1) + QIC^M(S_2). \end{aligned}$$

Proof. We first prove the statement of the lemma for $QIC_{\times\times}(S_1 \otimes S_2)$ instead of $QIC_{\times}(S_1 \otimes$

S_2), with $QIC_{\times\times}(S_1 \otimes S_2)$ allowing for product distributions $\mu_1 \otimes \mu_2$ in which μ_1 is an arbitrary distribution for the task S_1 and similarly for μ_2 and S_2.

The inequality $QIC(S_1 \otimes S_2) \geq QIC_{\times\times}(S_1 \otimes S_2)$ follows from definition since the maximization is over a larger set in QIC than in $QIC_{\times\times}$.

Then we prove that $QIC_{\times\times}(S_1 \otimes S_2) \geq QIC(S_1) + QIC(S_2)$. Let Π be a protocol succeeding at the product task $S_1 \otimes S_2$ and achieving, for all $\mu_1 \otimes \mu_2$, $QIC(\Pi, \mu_1 \otimes \mu_2) \leq QIC_{\times}(S_1 \otimes S_2) + \delta$ for an arbitrary small $\delta > 0$. Fix any μ_1, μ_2, and take the corresponding protocols Π^1, Π^2 from Lemma 4.3.4 for tasks S_1, S_2. They succeed at their respective task, and we get for all μ_1 and μ_2

$$\begin{aligned} QIC(\Pi^1, \mu_1) + QIC(\Pi^2, \mu_2) &= QIC(\Pi, \mu_1 \otimes \mu_2) \\ &\leq QIC_{\times\times}(S_1 \otimes S_2) + \delta. \end{aligned}$$

Optimizing over μ_1 and μ_2, we get $QIC(S_1) + QIC(S_2) \leq QIC_{\times\times}(S_1 \otimes S_2) + \delta$, as desired.

Finally, to prove that $QIC(S_1) + QIC(S_2) \geq QIC(S_1 \otimes S_2)$, let Π^1 and Π^2 be protocols succeeding at the corresponding tasks S_1, S_2, and achieving, for an arbitrary small $\delta > 0$ and all μ_1, μ_2, $QIC(\Pi^1, \mu_1) \leq QIC(S_1) + \delta$, $QIC(\Pi^2, \mu_2) \leq QIC(S_2) + \delta$, respectively. Taking the corresponding protocol Π from Lemma 4.3.6, it succeeds at the product task $S_1 \otimes S_2$, and we get, for any $\mu_{12} \in \mathscr{D}_{X^2Y^2}$ and corresponding marginals $\mu_1 \in \mathscr{D}_{X_1Y_1}$, $\mu_2 \in \mathscr{D}_{X_2Y_2}$,

$$QIC(\Pi, \mu_{12}) \leq QIC(\Pi^1, \mu_1) + QIC(\Pi^2, \mu_2).$$

Maximizing over all μ_{12}, we get the result

$$QIC(S_1 \otimes S_2) \leq QIC(S_1) + QIC(S_2) + 2\delta.$$

The result for $QIC_{\times\times}$ follows by taking δ to zero. Keeping tracks of rounds, we also get the bounded round result.

To complete the proof of the lemma, we show that for any $S = \otimes_{i=1}^{s}(T_i, \varepsilon_i)$, $QIC(S) =$

$QIC_\times(S) = \sum_{i=1}^{s} QIC(T_i, \varepsilon_i)$. The inequality $QIC(S) \geq QIC_\times(S)$ follows from definition since the maximization is over a larger set in QIC than in $QIC_\times$. Then, the result $QIC(S) = \sum_{i=1}^{s} QIC(T_i, \varepsilon_i)$ follows by induction on what we already proved.

Finally, we prove that $QIC_\times(S) \geq \sum_{i=1}^{s} QIC(T_i, \varepsilon_i)$ by induction on s. The base case follows by definition. For the induction step, assume the result for $s-1$ and let Π be a protocol for the task S and achieving, for all $\otimes_{i=1}^{s} \mu_i$, $QIC(\Pi, \otimes_{i=1}^{s} \mu_i) \leq QIC_\times(S) + \delta$ for an arbitrary small $\delta > 0$. Fix any $\otimes_{i=1}^{s-1} \mu_i$ and μ_s, and take the corresponding protocols Π^1, Π^2 from Lemma 4.3.4 for tasks $\otimes_{i=1}^{s-1}(T_i, \varepsilon_i)$ and (T_s, ε_s). They succeed at their respective task, and we get for all $\otimes_{i=1}^{s} \mu_i$

$$\begin{aligned} QIC(\Pi^1, \otimes_{i=1}^{s-1} \mu_i) + QIC(\Pi^2, \mu_s) &= QIC(\Pi, \otimes_{i=1}^{s} \mu_i) \\ &\leq QIC_\times(S) + \delta. \end{aligned}$$

Optimizing over $\otimes_{i=1}^{s} \mu_i$, we get $QIC_\times(\otimes_{i=1}^{s-1}(T_i, \varepsilon_i)) + QIC(T_s, \varepsilon_s) \leq QIC_\times(S) + \delta$. Using the induction hypothesis on $QIC_\times(\otimes_{i=1}^{s-1}(T_i, \varepsilon_i))$ yields the desired result. Keeping tracks of rounds, we also get the bounded round result. □

Let us state the following corollary.

Corollary 4.3.6. *For a product task $\otimes_i(T_i, \varepsilon_i)$ and any bound $M \in \mathbb{N}$ on the number of messages, the following holds:*

$$\begin{aligned} QIC(\otimes_i(T_i, \varepsilon_i)) &= \sum_i QIC(T_i, \varepsilon_i), \\ QIC^M(\otimes_i(T_i, \varepsilon_i)) &= \sum_i QIC^M(T_i, \varepsilon_i). \end{aligned}$$

Combining the above corollary with Lemma 4.3.19, we get the following.

Corollary 4.3.7. *For a product task (T_i, ε_i) and any bound $M \in \mathbb{N}$ on the number of*

messages, the following holds:

$$\sum_i QIC(T_i, \varepsilon_i) \leq QCC(\otimes_i (T_i, \varepsilon_i)),$$
$$\sum_i QIC^M(T_i, \varepsilon_i) \leq QCC^M(\otimes_i (T_i, \varepsilon_i)).$$

We now show that quantum information complexity is convex in the error.

Lemma 4.3.21 (Convexity in error)**.** *For any $p \in [0,1]$, T and ε, ε_1, $\varepsilon_2 \in [0,1]$ satisfying $\varepsilon = p \cdot \varepsilon_1 + (1-p) \cdot \varepsilon_2$ and for any bound $M = \max(M_1, M_2)$, M_1, $M_2 \in \mathbb{N}$ on the number of messages, the following holds:*

$$QIC(T, \varepsilon) \leq pQIC(T, \varepsilon_1) + (1-p)QIC(T, \varepsilon_2),$$
$$QIC^M(T, \varepsilon) \leq pQIC^{M_1}(T, \varepsilon_1) + (1-p)QIC^{M_2}(T, \varepsilon_2).$$

Proof. The proof is similar to the one for the analogous result with fixed input. Given $\delta > 0$, let Π^1 and Π^2 be protocols satisfying, for all μ, for $i \in \{1,2\}, \Pi^i \in \mathscr{T}(T, \varepsilon_i)$ and $QIC(\Pi^i, \mu) \leq QIC(T, \varepsilon_i) + \delta$, and take the corresponding protocol Π of Lemma 4.3.7. First, it holds that protocol Π successfully accomplishes its task, i.e. it implements task T on all inputs with error bounded by $\varepsilon = p\varepsilon_1 + (1-p)\varepsilon_2$. We must now verify that the quantum information cost satisfies the convexity property:

$$\begin{aligned} QIC(T, \varepsilon) &\leq \max_\mu QIC(\Pi, \mu) \\ &= \max_\mu \left(pQIC(\Pi^1, \mu) + (1-p)QIC(\Pi^2, \mu) \right) \\ &\leq p \max_\mu QIC(\Pi^1, \mu) + (1-p) \max_\mu QIC(\Pi^2, \mu) \\ &\leq pQIC(T, \varepsilon_1) + (1-p)QIC(T, \varepsilon_2) + 2\delta. \end{aligned}$$

The result follows by taking δ to zero. Keeping track of rounds, we get the bounded round result. □

As for distributional tasks, we get as a corollary that quantum information complexity

is continuous in the error.

Corollary 4.3.8 (Continuity in error)**.** *Quantum information complexity is continuous in the error. That is, for all T, r and ε, $\delta > 0$, there exists $\varepsilon' \in (0,\varepsilon)$ such that for all $\varepsilon'' \in (\varepsilon',\varepsilon)$,*

$$|QIC(T,\varepsilon'') - QIC(T,\varepsilon)| \leq \delta,$$
$$|QIC^r(T,\varepsilon'') - QIC^r(T,\varepsilon)| \leq \delta.$$

Similar results hold for the max-distributional quantum information complexity. However, we prove that we can almost reverse the quantifiers for the two possible notions of prior-free quantum information complexity. The proof idea follows along the lines of the proof of Theorem 3.5 in Ref. [33], but special care must be taken for quantum protocols. The idea we use is to take an ε-*net* over $\mathscr{D}_{XY}$, and then take a δ-optimal protocol for each distribution in the net. An ε-net over $\mathscr{D}_{XY}$ is a set $N_{XY} \subset \mathscr{D}_{XY}$ such that for all $\mu \in \mathscr{D}_{XY}$, there exists $\nu \in N_{XY}$ such that $\|\mu - \nu\|_1 \leq \varepsilon$. Since $\mathscr{D}_{XY}$ is a compact set, there exists a finite such N_{XY}.

Theorem 4.3.1. *For a classical task (T,ε) with $\varepsilon > 0$, a number of messages M and each value $\alpha \in (0,1)$,*

$$QIC^M\left(T,\frac{\varepsilon}{\alpha}\right) \leq \frac{QIC_D^M(T,\varepsilon)}{1-\alpha}.$$

Proof. Fix T, M, ε, α and denote $I = QIC_D^M(T,\varepsilon)$. For any $\delta_1 \in (0,1)$, we want to prove the existence of a protocol $\Pi \in \mathscr{T}^M(T, \frac{\varepsilon}{\alpha}\cdot(1+2\delta_1))$ satisfying $QIC(\Pi,\mu) \leq \frac{I\cdot(1+2\delta_1)}{1-\alpha}$ for all $\mu \in \mathscr{D}_{XY}$. This shows that $QIC^M(T, \frac{\varepsilon}{\alpha}\cdot(1+2\delta_1)) \leq \frac{I}{1-\alpha}\cdot(1+2\delta_1)$, and then by continuity of quantum information complexity in the error, we get the result by taking δ_1 to 0. The proof follows along the lines of the one for the analogous result for classical information complexity [33], using a minimax argument. We take extra care to account for the continuum of quantum protocols, the round-by-round definition of quantum information cost, and the fact that we do not have a bound on the size of the

entanglement. Let $\delta_2 \in (0, \varepsilon\delta_1)$ satisfy the following two properties uniformly for all μ_1, μ_2 that are δ_2-close, and also uniformly for all M-message protocols Π:

$$|QIC(\Pi, \mu_1) - QIC(\Pi, \mu_2)| \leq I \cdot \frac{\delta_1}{10}, \tag{4.3.2}$$

$$|QIC^M(T, \mu_1, \varepsilon - \delta_2) - QIC^M(T, \mu_1, \varepsilon)| \leq I \cdot \frac{\delta_1}{10}. \tag{4.3.3}$$

The first inequality is possible by Lemma 4.3.10, i.e. by the uniform continuity of quantum information cost in the input, uniformly over all M-message protocols, and the second is possible by Lemma 4.3.25, i.e. the continuity of quantum information complexity in the error, uniformly over all inputs. Fix a finite δ_2-net for $\mathscr{D}_{XY}$, that we denote N_{XY}. For each $\mu \in N_{XY}$, fix a protocol $\Pi_\mu \in \mathscr{T}^M(T, \mu, \varepsilon - \delta_2)$ such that $QIC(\Pi_\mu, \mu) \leq QIC^M(T, \mu, \varepsilon - \delta_2) \cdot (1 + \frac{\delta_1}{10})$ and denote the set of all such protocols P_N. We then have $|P_N| = |N_{XY}| < \infty$, and we get using (4.3.3) that

$$\begin{aligned} QIC(\Pi_\mu, \mu) &\leq QIC^M(T, \mu, \varepsilon - \delta_2) \cdot (1 + \frac{\delta_1}{10}) \\ &\leq \left(QIC^M(T, \mu, \varepsilon) + I \cdot \frac{\delta_1}{10}\right)(1 + \frac{\delta_1}{10}) \\ &\leq I \cdot (1 + \frac{\delta_1}{10})^2 \\ &\leq I \cdot (1 + \frac{\delta_1}{2}). \end{aligned} \tag{4.3.4}$$

We define the following two-player zero-sum game over these two sets. Player A comes up with a quantum protocol $\Pi \in P_N$. Player B comes up with a distribution $\mu \in N_{XY}$. Player B's payoff is given by

$$P_B(\Pi, \mu) = (1 - \alpha) \cdot \frac{QIC(\Pi, \mu)}{I} + \alpha \cdot \frac{P_e(\Pi, \mu)}{\varepsilon},$$

and then player A's is given by $P_A(\Pi, \mu) = -P_B(\Pi, \mu)$. Recall that $P_e(\Pi, \mu)$ is the average error of protocol Π for implementing T on μ. We first show the following.

Claim 4.3.1. *The value of the game for player B is bounded by* $1 + \delta_1$.

Proof. Let ν_B be a probability distribution over N_{XY} representing a mixed strategy for player B. To prove the claim, it suffices to show that there is a protocol $\Pi \in P_N$ such that $\mathbb{E}_{\nu_B}[P_B(\Pi,\mu)] < 1+\delta_1$. Let $\bar{\mu}$ be the distribution corresponding to averaging over ν_B, that is

$$\bar{\mu}(x,y) = \mathbb{E}_{\nu_B}\mu(x,y).$$

Let $\mu' \in N_{XY}$ be a distribution that is δ_2-close to $\bar{\mu}$, and $\Pi' \in P_N$ the corresponding protocol. We first show that Π' is also good for $\bar{\mu}$. We have

$$\begin{aligned} P_e(\Pi',\bar{\mu}) &\leq P_e(\Pi',\mu') + \delta_2 \\ &\leq \varepsilon - \delta_2 + \delta_2 \\ &= \varepsilon, \end{aligned}$$

in which the first inequality follows from the fact that $\bar{\mu}$ and μ' are δ_2-close and the second inequality from the fact that $\Pi' \in P_N$ is the protocol corresponding to $\mu' \in N_{XY}$, i.e. $\Pi' \in \mathscr{T}^r(T,\mu',\varepsilon-\delta_2)$. We also have

$$\begin{aligned} QIC(\Pi',\bar{\mu}) &\leq QIC(\Pi',\mu') + I \cdot \frac{\delta_1}{2} \\ &\leq I \cdot (1+\delta_1), \end{aligned}$$

in which the first inequality follows from (4.3.2) and the second from the fact that $\Pi' \in$

P_N is the protocol corresponding to $\mu' \in N_{XY}$ along with (4.3.4). We obtain

$$\begin{aligned}
\mathbb{E}_{\nu_B}[P_B(\Pi', \mu)] &= \mathbb{E}_{\nu_B}\left[(1-\alpha)\cdot\frac{QIC(\Pi', \mu)}{I} + \alpha\cdot\frac{P_e(\Pi', \mu)}{\varepsilon}\right] \\
&= (1-\alpha)\cdot\mathbb{E}_{\nu_B}\left[\frac{QIC(\Pi', \mu)}{I}\right] + \alpha\cdot\frac{P_e(\Pi', \bar{\mu})}{\varepsilon} \\
&\leq (1-\alpha)\cdot\left[\frac{QIC(\Pi', \bar{\mu})}{I}\right] + \alpha\cdot\frac{P_e(\Pi', \bar{\mu})}{\varepsilon} \\
&< (1-\alpha)\cdot(1+\delta_1) + \alpha \\
&< 1+\delta_1,
\end{aligned}$$

in which the first equality is by definition, the second by linearity of expectation, the first inequality is by Lemma 4.3.8, i.e. concavity of quantum information cost in the input state, and the second inequality is by the above results about Π'. This concludes the proof of the claim. □

By the minimax theorem for zero-sum games, the above claim implies that there exists a probability distribution ν_A over P_N representing a mixed strategy for player A and such that the value of the game for player B is at most $1+\delta_1$. That is, for all $\mu \in N_{XY}$,

$$\mathbb{E}_{\nu_A}(P_B(\Pi, \mu)) < 1+\delta_1.$$

Let $\bar{\Pi} = \mathbb{E}_{\nu_A}(\Pi)$ be the M-message protocol obtained by publicly averaging over ν_A, as per Lemma 4.3.7. This is the protocol we are looking for. The following claim holds.

Claim 4.3.2. *For all* $\mu \in \mathscr{D}_{XY}$, $(1-\alpha)\cdot\frac{QIC(\bar{\Pi}, \mu)}{I} + \alpha\cdot\frac{P_e(\bar{\Pi}, \mu)}{\varepsilon} < 1+2\delta_1$.

Proof. Fix any $\mu \in \mathscr{D}_{XY}$, and let $\mu' \in N_{XY}$ be a distribution that is δ_2-close to μ. Then

we obtain

$$
\begin{aligned}
(1-\alpha)\cdot\frac{QIC(\bar{\Pi},\mu)}{I}+\alpha\cdot\frac{P_e(\bar{\Pi},\mu)}{\varepsilon} & \\
&\leq (1-\alpha)\cdot\frac{QIC(\bar{\Pi},\mu')+I\delta_1}{I}+\alpha\cdot\frac{P_e(\bar{\Pi},\mu')+\delta_2}{\varepsilon} \\
&= (1-\alpha)\cdot\frac{QIC(\bar{\Pi},\mu')}{I}+\alpha\cdot\mathbb{E}_{\nu_A}\frac{P_e(\Pi,\mu')}{\varepsilon} \\
&\qquad +(1-\alpha)\cdot\delta_1+\alpha\cdot\frac{\delta_2}{\varepsilon} \\
&\leq (1-\alpha)\cdot\mathbb{E}_{\nu_A}\left[\frac{QIC(\Pi,\mu')}{I}\right]+\alpha\cdot\mathbb{E}_{\nu_A}\left[\frac{P_e(\Pi,\mu')}{\varepsilon}\right]+\delta_1 \\
&= \mathbb{E}_{\nu_A}[P_B(\Pi,\mu')]+\delta_1 \\
&< 1+2\delta_1,
\end{aligned}
$$

in which the first inequality follows from (4.3.2) and the fact that μ,μ' are δ_2-close, the first equality is because we take expectation over a probability, the second inequality is because $\delta_2 \leq \varepsilon\cdot\delta_1$ and also by Lemma 4.3.7, i.e. by the convexity of quantum information cost in the protocol, the second equality is by linearity of expectation and the definition of $P_B(\Pi,\mu')$, and the last inequality is because ν_A represents the mixed strategy obtained by the minimax theorem. Since this holds for all $\mu\in\mathscr{D}_{XY}$, this conclude the proof of the claim. □

To conclude the proof of the theorem, we first note that the above claim implies that for all $\mu\in\mathscr{D}_{XY}$,

$$QIC(\bar{\Pi},\mu)\leq\frac{I}{1-\alpha}(1+2\delta_1),$$

so $\bar{\Pi}$ satisfies the quantum information cost property we are looking for. It is left to verify that it also has low error on all inputs. The above claim also implies that for all μ,

$$P_e(\bar{\Pi},\mu)\leq\frac{\varepsilon}{\alpha}\cdot(1+2\delta_1).$$

Letting μ run over all distributions with support of size one, we get the desired error

property, i.e. $\bar{\Pi} \in \mathscr{T}^M(T, \frac{\varepsilon}{\alpha}(1+2\delta_1))$, and so

$$QIC^M\left(T, \frac{\varepsilon}{\alpha}\cdot(1+2\delta_1)\right) \leq \frac{I}{1-\alpha}(1+2\delta_1),$$

as desired. □

To extend this result to the unbounded round quantum setting, we adapt a compactness argument from Ref. [38], itself adapted from Ref. [125].

Theorem 4.3.2. *For a classical task (T,ε) with $\varepsilon > 0$ and each value $\alpha \in (0,1)$,*

$$QIC\left(T, \frac{\varepsilon}{\alpha}\right) \leq \frac{QIC_D(T,\varepsilon)}{1-\alpha}.$$

Proof. Let $I = QIC_D(T,\varepsilon)$, and denote by P_T the set of all protocols over the same input and output spaces as T. For any Π, $P_e(\Pi,\mu)$ is continuous in μ by properties of the statistical distance. Given $\delta > 0$, define

$$A(\Pi) = \{\mu \in \mathscr{D}_{XY} : QIC(\Pi,\mu) \geq I + 2\cdot\delta \text{ or } P_e(\Pi,\mu) \geq \varepsilon + \delta\}.$$

By continuity of both $QIC(\Pi,\mu)$ and $P_e(\Pi,\mu)$ in μ, these sets are closed for all $\Pi \in P_T$. Then, by definition of I, for all μ there exists $\Pi_\mu \in \mathscr{T}(T,\mu,\varepsilon)$ such that $QIC(\Pi_\mu,\mu) \leq I+\delta$, and so $\cap_{\Pi\in P_T} A(\Pi) = \emptyset$. Since $\mathscr{D}_{XY}$ is compact and the sets $A(\Pi)$ are closed, we get that there exists a finite set $Q \subset P_T$ such that $\cap_{\Pi\in Q} A(\Pi) = \emptyset$. We get that for all μ, there exists $\Pi_\mu \in Q$ such that $QIC(\Pi_\mu,\mu) < I+2\delta$ and $P_e(\Pi_\mu,\mu) < \varepsilon+\delta$. Let $M^* = \max\{M : \text{there is } \Pi \in Q \text{ with } M \text{ messages }\}$, then

$$\begin{aligned}
I + 2\delta &\geq \max_{\mu} \min_{\Pi\in Q\cap\mathscr{T}(T,\mu,\varepsilon+\delta)} QIC(\Pi,\mu) \\
&\geq QIC_D^{M^*}(T,\varepsilon+\delta) \\
&\geq (1-\alpha)\cdot QIC^{M^*}\left(T, \frac{\varepsilon}{\alpha}+\frac{\delta}{\alpha}\right) \\
&\geq (1-\alpha)\cdot QIC\left(T, \frac{\varepsilon}{\alpha}+\frac{\delta}{\alpha}\right).
\end{aligned}$$

The result follows by continuity of QIC and by taking δ to zero. □

4.3.2.5 Reducing the Error for Functions

Similarly to communication, it is possible to reduce the error when computing functions without increasing too much the information.

Lemma 4.3.22. *For any function f and error parameter $\varepsilon > 0$, the following holds:*

$$QIC(f,\varepsilon) \in O\left(\log 1/\varepsilon \cdot QIC(f,1/3)\right).$$

Proof. Given $\delta > 0$, let Π be a protocol computing f with worst-case error at most $1/3$ on every input and satisfying $QIC(\Pi,\mu) \leq QIC(f,1/3)+\delta$ for all μ. Let $n \in O(\log 1/\varepsilon)$ be given by the Chernoff bound such that protocol Π_n running Π n times in parallel as per Lemma 4.3.6, with each input being a copy of the instance to f, and taking a majority vote (with arbitrary tie-breaking) computes f correctly except with probability ε on every input. This n can be chosen independently of δ. We now argue on the quantum information cost of Π_n. Consider an arbitrary distribution μ for f, and let μ_n be the distribution once the n copies have been made. If we denote the marginal for the i-th copy by μ^i, then $\mu^i = \mu$ for all i. By Lemma 4.3.6 and an induction, we then get that

$$\begin{aligned} QIC(f,\varepsilon) &\leq QIC(\Pi_n,\mu_n) \\ &\leq nQIC(\Pi,\mu) \\ &\leq n(QIC(f,1/3)+\delta). \end{aligned}$$

The result follows by taking δ to 0. □

4.3.3 Quantum Tasks

The following properties hold for quantum tasks. Most proofs follow exactly as for distributional classical tasks; except when a different proof is required, we only state the result without proof.

4.3.3.1 Quantum Information Lower Bounds Communication

For quantum tasks also, information lower bounds communication.

Corollary 4.3.9. *For a product quantum task $\otimes_i(\mathcal{N}_i, \rho_i, \varepsilon_i)$ and any bound $M \in \mathbb{N}$ on the number of messages, the following holds*

$$0 \leq QIC(\otimes_i(\mathcal{N}_i, \rho_i, \varepsilon_i)) \leq QCC(\otimes_i(\mathcal{N}_i, \rho_i, \varepsilon_i)),$$
$$0 \leq QIC^M(\otimes_i(\mathcal{N}_i, \rho_i, \varepsilon_i)) \leq QCC^M(\otimes_i(\mathcal{N}_i, \rho_i, \varepsilon_i)).$$

4.3.3.2 Additivity

Quantum information complexity of quantum tasks also satisfy an exact direct sum property, a corollary of the following additivity result. For two product tasks $S_1 = \bigotimes_{i=1}^{s_1}(\mathcal{N}_i, \rho_i, \varepsilon_i), S_2 = \bigotimes_{j=s_1+1}^{s_1+s_2}(\mathcal{N}_j, \rho_j, \varepsilon_j)$, define the product task $S_1 \otimes S_2 = \bigotimes_{k=1}^{s_1+s_2}(\mathcal{N}_k, \rho_k, \varepsilon_k)$. We get the following results.

Lemma 4.3.23. *For any two product classical tasks $S_1 = \bigotimes_{i=1}^{s_1}(\mathcal{N}_i, \rho_i, \varepsilon_i)$, $S_2 = \bigotimes_{j=s_1+1}^{s_1+s_2}(\mathcal{N}_j, \rho_j, \varepsilon_j)$ and any bound $M \in \mathbb{N}$ on the number of messages,*

$$QIC(S_1 \otimes S_2) = QIC(S_1) + QIC(S_2),$$
$$QIC^M(S_1 \otimes S_2) = QIC^M(S_1) + QIC^M(S_2).$$

Corollary 4.3.10. *For a product task $(\mathcal{N}_i, \rho_i, \varepsilon_i)$ and any bound $M \in \mathbb{N}$ on the number of messages, the following holds:*

$$QIC(\otimes_i(\mathcal{N}_i, \rho_i, \varepsilon_i)) = \sum_i QIC(\mathcal{N}_i, \rho_i, \varepsilon_i),$$
$$QIC^M(\otimes_i(\mathcal{N}_i, \rho_i, \varepsilon_i)) = \sum_i QIC^M(\mathcal{N}_i, \rho_i, \varepsilon_i).$$

Corollary 4.3.11. *For a product task $(\mathcal{N}_i, \rho_i, \varepsilon_i)$ and any bound $M \in \mathbb{N}$ on the number*

of messages, the following holds:

$$\sum_i QIC(\mathcal{N}_i, \rho_i, \varepsilon_i) \leq QCC(\otimes_i(\mathcal{N}_i, \rho_i, \varepsilon_i)),$$
$$\sum_i QIC^M(\mathcal{N}_i, \rho_i, \varepsilon_i) \leq QCC^M(\otimes_i(\mathcal{N}_i, \rho_i, \varepsilon_i)).$$

4.3.3.3 Convexity and Continuity

We also prove that quantum information complexity is convex and continuous in the error parameter. For quantum tasks, it also make sense to consider linear combination of the channels to be implemented in the task. We get the following result.

Lemma 4.3.24. *For any $\mathcal{N}_1$, $\mathcal{N}_2$, ρ, ε_1, ε_2, M_1, M_2 and $p \in [0,1]$, if we define $\varepsilon = p\varepsilon_1 + (1-p)\varepsilon_2$, $\mathcal{N} = p\mathcal{N}_1 + (1-p)\mathcal{N}_2$, $M = \max(M_1, M_2)$, then the following holds:*

$$QIC(\mathcal{N}, \rho, \varepsilon) \leq pQIC(\mathcal{N}_1, \rho, \varepsilon_1) + (1-p)QIC(\mathcal{N}_2, \rho, \varepsilon_2),$$
$$QIC^M(\mathcal{N}, \rho, \varepsilon) \leq pQIC^{M_1}(\mathcal{N}_1, \rho, \varepsilon_1) + (1-p)QIC^{M_2}(\mathcal{N}_2, \rho, \varepsilon_2).$$

Proof. Let Π^1 and Π^2 be protocols satisfying, for $i \in \{1,2\}$, $\Pi^i \in \mathcal{T}(\mathcal{N}_i, \rho, \varepsilon_i)$ and $QIC(\Pi^i, \rho) \leq QIC(\mathcal{N}_i, \rho, \varepsilon_i) + \delta$ for an arbitrary small $\delta > 0$, and take the corresponding protocol Π of Lemma 4.3.7. We first verify that protocol Π successfully accomplishes its task. The result follows from the triangle inequality for the trace distance:

$$\begin{aligned}
\|\Pi(\rho^{A_{out}B_{out}R}) - \mathcal{N}(\rho^{A_{out}B_{out}R})\|_1 &= \|p\Pi^1(\rho^{A_{out}B_{out}R}) + (1-p)\Pi^2(\rho^{A_{out}B_{out}R}) \\
&\quad - (p\mathcal{N}_1 + (1-p)\mathcal{N}_2)(\rho^{A_{out}B_{out}R})\|_1 \\
&\leq \|p\Pi^1(\rho^{A_{out}B_{out}R}) - p\mathcal{N}_1(\rho^{A_{out}B_{out}R})\|_1 \\
&\quad + \|(1-p)\Pi^2(\rho) - (1-p)\mathcal{N}_2(\rho)\|_{A_{out}B_{out}R} \\
&\leq p\varepsilon_1 + (1-p)\varepsilon_2 \\
&= \varepsilon.
\end{aligned}$$

We must now verify that the quantum information cost satisfies the convexity property:

$$\begin{aligned} QIC(\mathcal{N},\rho,\varepsilon) &\leq QIC(\Pi,\rho) \\ &= pQIC(\Pi^1,\rho)+(1-p)QIC(\Pi^2,\rho) \\ &\leq pQIC(\mathcal{N}_1,\rho,\varepsilon_1)+(1-p)QIC(\mathcal{N}_2,\rho,\varepsilon_2)+2\delta. \end{aligned}$$

Keeping track of rounds, we get the bounded round result. □

We get as a corollary that quantum information complexity is continuous in its error parameter. Recall that quantum information complexity is non-increasing in the error parameter.

Lemma 4.3.25. *Quantum information complexity is continuous in the error. This holds uniformly in the input. That is, for all $\mathcal{N}$, M and ε, $\delta > 0$, there exists $\varepsilon' \in (0,\varepsilon)$ such that for all $\varepsilon'' \in (\varepsilon',\varepsilon)$ and all ρ,*

$$\begin{aligned} |QIC(\mathcal{N},\rho,\varepsilon'') - QIC(\mathcal{N},\rho,\varepsilon)| &\leq \delta, \\ |QIC^M(\mathcal{N},\rho,\varepsilon'') - QIC^M(\mathcal{N},\rho,\varepsilon)| &\leq \delta. \end{aligned}$$

4.4 Quantum Information Cost in the Cleve-Buhrman Model

We can also define a sensible notion of quantum information complexity for protocols in the Cleve-Buhrman model.

Keeping with our purified view of quantum protocols in the context of quantum information cost, a protocol Π_{CB} in the Cleve-Buhrman model would be mapped into an equivalent protocol Π' in the hybrid model, with each quantum instrument replaced by its isometric extension. Moreover, to account for the fact that messages are classical, the quantum copy of each message C_i generated by the isometric extension would be put into the reference register R'_i for subsequent rounds, inaccessible to both Alice and Bob. Remember that both Alice and Bob would also get a copy of C_i in M_i^A, M_i^B as defined in

Section 2.4.1.2. The quantum information cost of Π_{CB} would be defined accordingly in terms of Π' and the state $\rho_i^{A_i'B_i'C_i'R_1'\cdots R_i'R}$ after message i has been sent.

Definition 4.4.1. *For a protocol Π_{CB} in the Cleve-Buhrman model and its purification Π' in the hybrid model as defined above, and an input state ρ, we define the* quantum information cost *of Π_{CB} on input ρ as*

$$QIC_{CB}(\Pi_{CB},\rho) = \sum_{i\geq 1,\ odd} \frac{1}{2} I(C_i; RR_1'\cdots R_{i-1}'|B_iM_{i-1}^B) + \sum_{i\geq 1,\ even} \frac{1}{2} I(C_i; RR_1'\cdots R_{i-1}'|A_iM_{i-1}^A).$$

Note that Π' implements the same channel, has the same communication and the same number of messages of Π_{CB}. Moreover, once the reference register R_i' is traced out, the message C_i is classical. We now show that this quantum information cost for Π_{CB} evaluates to the standard quantum information cost for its purification Π', and that the factor of $1/2$ appearing in the standard definition in order to lower bound the quantum communication cost is in fact unnecessary here, since all communication is classical.

Lemma 4.4.1. *For a protocol Π_{CB} in the Cleve-Buhrman model and its purification Π' in the hybrid model as defined above, and an input state ρ, it holds that*

$$QIC_{CB}(\Pi_{CB},\rho) = QIC(\Pi',\rho),$$
$$2QIC_{CB}(\Pi_{CB},\rho) \leq QCC(\Pi_{CB}).$$

Proof. The results follow since the sender's copy in M_i^A or M_i^B of the message $C_1,\cdots,C_i$ is traced out in each term of the quantum information cost. Hence, C_i is classical in each term and the bound $2QIC_{CB}(\Pi_{CB},\rho) \leq QCC(\Pi_{CB})$ follows since for odd i, $I(C_i;RR_1'\cdots R_{i-1}'|B_iM_{i-1}^B) \leq \log|C_i|$ and for even i, $I(C_i;RR_1'\cdots R_{i-1}'|A_iM_{i-1}^A) \leq \log|C_i|$. Also, for odd i, M_{i-1}^B holds a classical copy of $R_1',\cdots,R_{i-1}'$, so that $I(C_i;R|B_iM_{i-1}^B) = I(C_i;R|B_iM_{i-1}^BR_2'R_4'\cdots R_{i-1}') = I(C_i;RR_1'\cdots R_{i-1}'|B_iM_{i-1}^B)$, and similarly on Alice's side, so that $QIC_{CB}(\Pi_{CB},\rho) = QIC(\Pi',\rho)$. □

Note that to avoid cumbersome notation, we have avoided defining the Cleve-Buhrman model with a pattern of communication that might depend on the partial transcript, and

maybe also on shared randomness as in the randomized model. If we do not care about the number of message, then this can affect communication by at most a factor of two if we assume bit communication, by having Alice and Bob send bits in alternation. If we want to maintain the round complexity, then the communication can be at most affected by a multiplicative factor of the number of messages. For information however, this would be of no effect on both round complexity and information complexity. Using ideas developed here and the fact that sending padding messages in a pure state, as in the proof of Lemma 4.3.7, contributes zero quantum information cost, we can see that it is also possible to obtain the result of Lemma 4.4.1 for such an extension of the Cleve-Buhrman in which the order of communication might depend on the partial transcript and additional shared randomness.

Also note that with similar considerations, we could have alternatively defined the quantum information cost in the randomized model by considering a purification R_ν of the shared randomness used in protocol Π_ν, inaccessible to both Alice and Bob, and having each term of the form $I(C_i;RR_\nu|B_iS_B)$, with S_B being Bob's copy of the shared randomness. This is also easily seen to yield the same value as in Definition 4.2.3 that we adopted, as an average over ν.

4.5 The Cost of Forgetting Classical Information

In this section, we show that even though quantum protocols are reversible and thus each party can somehow forget information that he learned about the input, there is a quantum information cost associated in particular with forgetting classical information. This holds either for information about a party's own input, or for information about the other party's input.

4.5.1 Safe Copies Do Not Increase Quantum Information Cost

We first show that making *safe copies* of classical inputs do not increase the quantum information cost. A safe copy is a copy that is made by each party at the outset of the protocol and not acted upon during the remainder of the protocol. Consider any protocol

Π, and let a protocol Π' be a protocol in which Alice and Bob make a coherent copy of their respective inputs X,Y at the outset of the protocol into safe registers X',Y', and then run Π. Recall that there are also coherent copies held in purification registers R_X,R_Y. That is, on input distribution μ on X,Y, we denote as ρ_μ^{XY} the state

$$\rho_\mu^{XY} = \sum_{x,y} \mu(x,y)|x\rangle\langle x|^X \otimes |y\rangle\langle y|^Y, \tag{4.5.1}$$

and then a purification is

$$\left|\rho_\mu\right\rangle^{XYR_XR_Y} = \sum_{x,y} \sqrt{\mu(x,y)}\,|x\rangle^X\,|y\rangle^Y\,|x\rangle^{R_X}\,|y\rangle^{R_Y}\,. \tag{4.5.2}$$

In the protocol Π', these registers X',Y' are then left untouched for the remainder of the protocol, which is then identical to protocol Π after such copies are made. We want to show that the quantum information cost of Π' is never greater than that of Π. More formally, define the isometries

$$U_X^{X\to XX'} = \sum_{x\in X} |x\rangle^X\,|x\rangle^{X'}\,\langle x|^X, \tag{4.5.3}$$

$$U_Y^{Y\to YY'} = \sum_{y\in Y} |y\rangle^Y\,|y\rangle^{Y'}\,\langle y|^Y\,. \tag{4.5.4}$$

Then the protocol Π' is defined by applying U_X on Alice's side before applying U_1 in the first round, and by applying U_Y on Bob's side before applying U_2 in the second round, and then running U_i in round i for $i \geq 3$. We then get the following result.

Lemma 4.5.1. *For any protocol Π and any input distribution μ for X,Y, the protocol Π' as defined above satisfies*

$$QIC(\Pi',\mu) \leq QIC(\Pi,\mu). \tag{4.5.5}$$

Proof. This follows from Lemma 4.3.13, since U_Y commutes with U_1 on Alice's side, and we can implement a protocol applying U_X,U_Y without any communication. □

Let us illustrate the difference between $QIC(\Pi, \mu)$ and $QIC(\Pi', \mu)$ with a simple example. Consider an input distribution μ such that X is uniformly distributed, and $Y = X$. Consider a protocol in which Alice directly sends her input to Bob. Then the costs are

$$\begin{aligned} QIC(\Pi, \mu) &= I(X : R_X R_Y | Y)_{\rho_\mu} && (4.5.6)\\ &= H(X|Y)_{\rho_\mu} - H(X|R_X R_Y Y)_{\rho_\mu} && (4.5.7)\\ &= H(X)_{\rho_\mu} && (4.5.8)\\ &= |X|, && (4.5.9)\\ QIC(\Pi', \mu) &= I(X : R_X R_Y | Y'Y)_{\rho'_\mu} && (4.5.10)\\ &= 0, && (4.5.11) \end{aligned}$$

in which we used for $QIC(\Pi', \mu)$ the fact that all registers are classical once X' is traced out along with the fact that $X = Y$, similarly for $H(X|Y)_{\rho_\mu}$, and finally, since $\rho_\mu^{XYR_XR_Y}$ is pure, $H(X|R_X R_Y Y)_{\rho_\mu} = -H(X)_{\rho_\mu}$.

Notice that for protocols with classical inputs, making such a local copy of the input does not change the quantum communication cost. Hence, whenever we are interested in minimizing the quantum information cost, we may always consider such protocols that start by making a local copy of their inputs.

4.5.2 Alternate Characterization for Classical Inputs

For protocols with classical inputs, we provide an alternative characterization of their quantum information cost that does not require introducing a purification register. Before proving this, we first introduce some new quantities. In a recent work, Kerenidis, Laurière, Le Gall and Rennela define a notion of classical input information cost [87]. They also define an asymmetric version of quantum information cost. They have the following definitions, in which we consider protocols making safe copies of the input into registers $X'Y'$, as in the previous section.

Definition 4.5.1. *For a protocol* Π *and an input distribution* μ*, the* classical input infor-

mation cost from Alice to Bob *is defined as*

$$CIC_{A\to B}(\Pi,\mu)=\frac{1}{2}\sum_{i\geq 1,\ odd} I(C_i;X'|Y'B_i),$$

the classical input information cost from Bob to Alice *as*

$$CIC_{B\to A}(\Pi,\mu)=\frac{1}{2}\sum_{i\geq 1,\ even} I(C_i;Y'|X'A_i),$$

the quantum information cost from Alice to Bob *as*

$$QIC_{A\to B}(\Pi,\mu)=\frac{1}{2}\sum_{i\geq 1,\ odd} I(C_i;R_XR_Y|Y'B_i),$$

and the quantum information cost from Bob to Alice *as*

$$QIC_{B\to A}(\Pi,\mu)=\frac{1}{2}\sum_{i\geq 1,\ even} I(C_i;R_XR_Y|X'A_i).$$

It follows from the data processing inequality that $CIC_{A\to B}\leq QIC_{A\to B}$, a fact noted in Ref. [87]. They also consider another notion of information cost for classical inputs, termed the superposed information cost, and prove that it lies in between these two.

Note that $QIC(\Pi,\mu)=QIC_{A\to B}(\Pi,\mu)+QIC_{B\to A}(\Pi,\mu)$, so similarly we define a symmetric version of classical input information cost $CIC(\Pi,\mu)=CIC_{A\to B}(\Pi,\mu)+CIC_{B\to A}(\Pi,\mu)$. We want to compare these two quantities, and in particular we find that they are related with a further notion of information cost, which we call the Holevo information cost. This quantity evaluates the Holevo information each party possesses at the end of the protocol about the other party's input, conditional on his own input. Note that similar considerations can be made in each round i by considering the protocol Π_i that runs Π up to round i and then stops (with an appropriate partition of the registers in round i, depending on whether i is even or odd, and who holds C_i).

Definition 4.5.2. *For a protocol* Π *and an input distribution* μ*, the* Holevo information

cost from Alice to Bob *is defined as*

$$HIC_{A\to B}(\Pi,\mu) = I(X';B_{out}B'|Y'),$$

and the Holevo information cost from Bob to Alice *as*

$$HIC_{B\to A}(\Pi,\mu) = I(Y';A_{out}A'|X').$$

The introduction of the reference register R in the definition of quantum information cost, which can be decomposed into R_X, R_Y for classical inputs, is natural when discussing compression while keeping quantum correlations, and for general quantum inputs, but when discussing protocols implementing classical tasks it might appear somewhat artificial. We now present an alternative characterization of quantum information cost on classical inputs that does not involve such purification registers and only mention the classical input registers, similar to the notion of classical input information cost (CIC) of Ref. [87]. We start by expanding the ith term in the quantum information cost. For odd i,

$$I(C_i;R_XR_Y|Y'B_i)_{\rho'_i} = I(C_i;R_X|Y'B_i)_{\rho'_i} + I(C_i;R_Y|R_XY'B_i)_{\rho'_i}, \tag{4.5.12}$$

and similarly for even i with the conditioning instead on $X'A_i$. The first term on the right end side is the classical input information cost term $I(C_i;R_X|Y'B_i) = I(C_i;X'|Y'B_i)$ in round i and somehow quantifies the amount of information that message C_i contains about X for someone who already knows Y and possesses B_i as side-information, while the second one does not immediately have such an intuitive interpretation. However, we can rewrite it as $I(C_i : R_Y|X'A_i) = I(C_i : Y'|X'A_i)$ since $X'A_i$ contain a purification of $\rho_i^{B_iC_iR_XR_YY'}$. Notice that X', Y' are both classical in this term, which can now be informally interpreted as the amount of information that message C_i contains about Y for someone who already knows X and possess A_i. But remember that it is Alice who generated message C_i, so that $I(C_i : Y'|X'A_i)$ would always evaluate to 0 in a classical protocol. However, quantum protocols are reversible, so it is somehow possible to forget informa-

tion along the way. The term $I(C_i : Y'|X'A_i)$ then somehow corresponds to the amount of information Alice is forgetting about Y when transmitting C_i, and so we call the sum of these terms the classical input *reverse* information cost (CRIC).

Definition 4.5.3. *For a protocol* Π *and an input distribution* μ*, the* classical input reverse information cost from Bob back to Alice *is defined as*

$$CRIC_{A\leftarrow B}(\Pi,\mu) = \frac{1}{2}\sum_{i\geq 1,\ even} I(C_i;X'|Y'B_i),$$

and the classical input reverse information cost from Alice back to Bob *as*

$$CRIC_{B\leftarrow A}(\Pi,\mu) = \frac{1}{2}\sum_{i\geq 1,\ odd} I(C_i;Y'|X'A_i).$$

We soon make the above intuition more precise by providing an operational interpretation, but let's first consider a simple example. Let μ be an input distribution with X,Y distributed independently and uniformly on n bits, and consider a protocol in which after the second round, Alice has received a copy of Bob's input Y. At this point, Alice then copies the first m out of the n bits of Y, and then sends back Y to Bob. Then the term with $i = 3$ in $CRIC_{B\leftarrow A}$ will amount to the $n-m$ bits of information about Y that Alice is forgetting.

Now, to make this more precise, we can consider the following scenario. Consider an input state XY purified by R_XR_Y. Alice is given her input X as usual, but also the purification R_Y of Bob's input. Bob is only given his input Y, and so only the register R_X is held by the referee. Alice is given the register R_Y in order for her to be able to generate any message C_i in the protocol, for i odd as well as i even, and then transmit this message to Bob, after giving him his side information B_iY'. We are interested in how much new information about X this message C_i contains, hence we are only putting R_X in the referee's hand. More formally, suppose that we are interested in this information for round i. We then ask what is the asymptotic quantum communication cost for redistributing the C_i register of this state from Alice to Bob if, apart from C_i, Alice holds the A_i, X', R_Y registers and Bob holds the B_i, Y' registers. This is $I(C_i : R_X|B_iY') = I(C_i : X'|Y'B_i)$,

for classical registers X', Y'. Depending on whether i is odd or even, this is the ith term in $CIC_{A\to B}$ or in $CRIC_{A\leftarrow B}$. Remember that quantum communication in state redistribution is symmetric under time-reversal, so that the cost is the same if Bob decides to send back this message to Alice. Hence, not only does this scenario gives an operational interpretation to CIC as the amount of information about X Alice is sending to Bob in odd rounds, but also to CRIC as the amount of information about X Bob is forgetting by sending it back to Alice in even rounds.

From this operational interpretation, it is then intuitive that in any odd round i, after reception by Bob of message C_i from Alice, the conditional Holevo information $I(X' : B_iC_i|Y')$ Bob has about Alice's input can be written as follows:

$$I(X';B_iC_i|Y') = \sum_{j\leq i,\, j\ odd} I(C_j;X'|Y'B_j) - \sum_{j\leq i,\, j\ even} I(C_j;X'|Y'B_j), \tag{4.5.13}$$

in which on the right hand side the first sum corresponds to terms in $CIC_{A\to B}$ and the second one to terms in $CRIC_{A\leftarrow B}$. Note that this result follows from Lemma 4.3.14 with classical registers $R_a = X$, $R_b = Y$, along with the fact that for two classical copies Y_1, Y_2 of Y, $I(C_i;X|Y_1Y_2B_i) = I(C_i;X|YB_i)$ and $I(X;Y_1|Y_2) = 0$. Similar statements hold for Alice, with the role of odd and even rounds interchanged, and the proof is almost symmetric once we realize that the first term in $CRIC_{B\leftarrow A}$ vanishes, i.e. $I(C_1;Y'|X'A_1) = 0$, which follows by the following chain of inequality, in which we use the data processing inequality along with isometric invariance:

$$\begin{aligned} 0 &\leq I(Y' : C_1|X') \\ &\leq I(Y' : A_1C_1|X') \\ &= I(Y' : A_{in}T_A|X') \\ &= 0. \end{aligned}$$

Also note that from this and Lemma 4.3.13, the developments in Ref. [48] imply a linear lower bound on the quantum information complexity of computing the inner product function with bounded error. Similarly, the developments in Ref. [104] imply a

linear lower bound on the quantum information complexity of computing the distributed quantum Fourier transform between Alice's and Bob's input qubits.

The above characterization of the Holevo information cost also implies a relationship between classical input information cost and quantum information cost.

Lemma 4.5.2. *For a protocol* Π *and an input distribution* μ*, it holds that*

$$QIC(\Pi,\mu) \leq 2CIC(\Pi,\mu).$$

Proof. This follows from the fact that $I(X' : B_iC_i|Y) \geq 0$ along with (4.5.13), and the following chain of inequality:

$$\begin{aligned}
2QIC(\Pi,\mu) &= \sum_{1\leq i,\ odd} \big(I(C_i : X'|Y'B_i) + I(C_i : Y'|X'A_i)\big) && (4.5.14)\\
&\quad + \sum_{1\leq i,\ even} \big(I(C_i : Y'|X'A_i) + I(C_i : X'|Y'B_i)\big) && (4.5.15)\\
&\leq \sum_{1\leq i,\ odd} \big(I(C_i : X'|Y'B_i) + I(C_i : X'|Y'B_i)\big) && (4.5.16)\\
&\quad + \sum_{1\leq i,\ even} \big(I(C_i : Y'|X'A_i) + I(C_i : Y'|X'A_i)\big) && (4.5.17)\\
&= 4CIC(\Pi,\mu). && (4.5.18)
\end{aligned}$$

□

Note that the following remark also makes it intuitively clear that $QIC(\Pi,\mu) \leq 2CIC(\Pi,\mu)$ should hold. Given an M-message protocol Π, let Π' be the protocol that runs Π forward and then, without making any copy of the output, runs Π backward. Then the $(M+k)$th message is identical to the $(M-k)$th message, except that the role of the sender and the receiver have been exchanged. Since the terms in QIC are symmetric under time-reversal, we have $QIC_{A\to B}(\Pi',\mu) = QIC_{B\to A}(\Pi',\mu) = QIC(\Pi,\mu)$. Also, $CIC_{A\to B}(\Pi',\mu) = CIC_{A\to B}(\Pi,\mu) + CRIC_{A\leftarrow B}(\Pi,\mu) = QIC(\Pi,\mu)$ and $CIC_{B\to A}(\Pi',\mu) = CIC_{B\to A}(\Pi,\mu) + CRIC_{B\leftarrow A}(\Pi,\mu) = QIC(\Pi,\mu)$ since the last M messages in Π' consist of the M message of Π run backward and thus contribute the $CRIC$ terms. Thus,

$QIC(\Pi', \mu) = 2QIC(\Pi, \mu)$ and $CIC(\Pi', \mu) = QIC(\Pi, \mu)$. Intuitively, it is clear that $CRIC_{A \leftarrow B}(\Pi, \mu)$ should be at most $CIC_{A \rightarrow B}(\Pi, \mu)$ and $CRIC_{B \leftarrow A}(\Pi, \mu)$ should be at most $CIC_{B \rightarrow A}(\Pi, \mu)$, since it should not be possible to send back more information about the other party's input then what was received. This intuition also leads to the inequality $QIC(\Pi, \mu) \leq 2CIC(\Pi, \mu)$.

Note that we also get the following result if we make a copy of the classical output before running Π backward, as in the case of clean protocols as defined in Ref. [48].

Lemma 4.5.3. *For any classical input μ and protocol Π with classical output, let Π' be the protocol that runs Π forward, makes a copy of the output into A'_{out}, B'_{out}, and then runs Π backward. Then Π', with output in $A'_{out}B'_{out}$, implements the same channel as Π, and its quantum information cost satisfies*

$$QIC(\Pi', \mu) \leq 2QIC(\Pi, \mu).$$

Notice that the above developments shed new light on why previous definitions of quantum information cost, which were more similar in spirit to classical input information cost than to quantum information cost, were restricted to compression results for a single round. In the first round, both quantities evaluate to the same value, since Alice does not yet possess any information on Bob's input (aside from what she can infer from her own input), and then the $CRIC_{A \leftarrow B}$ term evaluates to zero. For one-round protocols, it is then immaterial whether one uses classical input information cost or quantum information cost. But then in subsequent rounds, generally the CRIC term can be much larger than the CIC term in the quantum information cost. In fact, it is easy to construct from the example above a protocol in which in the second round the CRIC term is large while the CIC term is zero. Thus, trying to compress such a quantum message down to CIC, that is, almost at no cost, while keeping the overall state of the protocol almost equivalent to that in the original protocol is bound to fail: we know from our developments that to forget information we must invest communication! As a consequence, we see that for quantum protocols, it is important to take into account the cost of forgetting information.

4.6 Direct Sum Theorem

In this section, we prove the first general multi-round direct sum theorem for quantum communication complexity. In order to prove such a result, we present a protocol compression result in terms of the quantum information cost. But first, in order to get an intuition for why quantum information complexity should be an appropriate notion to study direct sum questions, we prove its link to amortized quantum communication complexity.

4.6.1 Amortized Quantum Communication

Before getting to the bounded round direct sum theorem, we first prove that quantum information equals amortized quantum communication. That is, we prove the following.

Theorem 4.6.1. *For any classical task (T,μ,ε) with $\varepsilon > 0$ and any number of message M, the following holds:*

$$QIC(T,\mu,\varepsilon) = AQCC(T,\mu,\varepsilon),$$
$$QIC^{M}(T,\mu,\varepsilon) = AQCC^{M}(T,\mu,\varepsilon).$$

Proof. The converse part, which states that quantum information complexity is a lower bound on the amortized quantum communication complexity, follows directly from Corollary 4.3.5, and taking the limit. That is, it holds for all n that

$$\begin{aligned} QIC(T,\mu,\varepsilon) &= \frac{1}{n} QIC((T,\mu,\varepsilon)^{\otimes n}) \\ &\leq \frac{1}{n} QCC((T,\mu,\varepsilon)^{\otimes n}). \end{aligned}$$

The direct coding theorem, which states that the quantum information complexity is an achievable rate for amortized quantum communication complexity, follows from Lemma 4.3.3 and continuity of quantum information complexity in the error. That is, take an arbitrarily small $\delta > 0$, and use Corollary 4.3.25 to find an $0 < \varepsilon' < \varepsilon$ such that $QIC(T,\mu,\varepsilon-\varepsilon') \leq QIC(T,\mu,\varepsilon) + \delta$. We then consider a protocol $\Pi \in \mathscr{T}(T,\mu,\varepsilon-\varepsilon')$

satisfying $QIC(\Pi,\mu) \leq QIC(T,\mu,\varepsilon-\varepsilon') + \delta$. We now use Lemma 4.3.3 to find, for a sufficiently large n_0 such that for all $n \geq n_0$, a protocol $\Pi_n \in \mathscr{T}(\Pi^{\otimes n}, \mu^{\otimes n}, \varepsilon')$ satisfying $\frac{1}{n}QCC(\Pi_n) \leq QIC(\Pi,\mu) + \delta$. We then have the following chain of inequality:

$$\begin{aligned}\frac{1}{n}QCC(\Pi_n) &\leq QIC(\Pi,\mu) + \delta \\ &\leq QIC(T,\mu,\varepsilon-\varepsilon') + 2\delta \\ &\leq QIC(T,\mu,\varepsilon) + 3\delta.\end{aligned}$$

Since $\delta > 0$ is arbitrarily small and this holds for all sufficiently large n, we only have to verify that $\Pi_n \in \mathscr{T}((T,\mu,\varepsilon)^{\otimes n})$ to complete the proof. Recall that $P_e^i(\Pi_n, \mu^{\otimes n})$ denotes the error Π_n makes for the i-th coordinate. We have for each $i \in [n]$,

$$\begin{aligned}P_e^i(\Pi_n, \mu^{\otimes n}) &\leq P_e^i(\Pi^{\otimes n}, \mu^{\otimes n}) + \varepsilon' \\ &\leq \varepsilon - \varepsilon' + \varepsilon' \\ &= \varepsilon,\end{aligned}$$

in which we first use properties of the trace distance, and then the fact that $\Pi \in \mathscr{T}(T,\mu,\varepsilon-\varepsilon')$. The result for bounded round follows similarly. □

A similar result holds for quantum tasks.

Theorem 4.6.2. *For any $\mathscr{N}$, ρ, M, $\varepsilon > 0$, the following holds:*

$$\begin{aligned}QIC(\mathscr{N},\rho,\varepsilon) &= AQCC(\mathscr{N},\rho,\varepsilon), \\ QIC^M(\mathscr{N},\rho,\varepsilon) &= AQCC^M(\mathscr{N},\rho,\varepsilon).\end{aligned}$$

Proof. The proof for the classical case also apply here, up to the verification that the corresponding protocol Π_n is in $\mathscr{T}((\mathscr{N},\rho,\varepsilon)^{\otimes n})$. The result in such a case follows since we have $\Pi \in \mathscr{T}(\mathscr{N},\rho,\varepsilon)$ and $\Pi_n \in \mathscr{T}(\Pi^{\otimes n}, \rho^{\otimes n}, \varepsilon')$. Then, for arbitrarily small

$\delta > 0$, we get for each $i \in [n]$,

$$\begin{aligned}\|\operatorname{Tr}_{\neg A^i_{out}B^i_{out}} \Pi_n(\rho^{\otimes n}) - \mathscr{N}(\rho)\| &\leq \|\operatorname{Tr}_{\neg A^i_{out}B^i_{out}} \Pi_n(\rho^{\otimes n}) - \Pi(\rho)\| \\ &\quad + \|\Pi(\rho) - \mathscr{N}(\rho)\| \\ &\leq \|\Pi_n(\rho^{\otimes n}) - \Pi^{\otimes n}(\rho^{\otimes n})\| + \varepsilon - \varepsilon' \\ &\leq \varepsilon,\end{aligned}$$

in which we first use the triangle inequality, and then monotonicity of the trace distance under partial trace. Hence, $\Pi_n \in \mathscr{T}((\mathscr{N}, \rho, \varepsilon)^{\otimes n})$ □

This asymptotic result makes it clear intuitively that quantum information complexity is the right notion to consider in the context of general direct sum. Indeed, by replacing asymptotic compression with a one-shot compression result, we obtain our direct sum result in Section 4.6.3.

4.6.2 Protocol Compression at Information Cost

To be able to compress a protocol proportionally to its quantum information cost, we first compress a single message down to a communication cost proportional to its conditional mutual information, as in asymptotic state redistribution. Entanglement is deemed free for the compression. We obtained the following result for single message compression in Chapter 3.

Lemma 4.6.1. *For all $\varepsilon \in (0, 1/2)$ and $\rho \in \mathscr{D}(ABC)$ with purifying register R, there exists a one-message protocol $\Pi \in \mathscr{T}(\mathscr{R}, \rho, \varepsilon)$ with quantum communication satisfying*

$$QCC(\Pi) \leq \frac{61}{\varepsilon^2} \cdot I(C;R|B)_\rho + \frac{242}{\varepsilon^2} + 16.$$

To compress a single protocol to a quantum communication cost close to its quantum information cost, we apply single message compression iteratively, once for each of the M messages in the protocol. We obtain the following result.

Lemma 4.6.2. *For any $\varepsilon \in (0,1/2)$, any M-message protocol Π, any input state ρ, there exists an M-message compression protocol $\Pi' \in \mathscr{T}(\Pi,\rho,\varepsilon)$ satisfying*

$$QCC(\Pi') \leq \frac{32M^2}{\varepsilon^2} QIC(\Pi,\rho) + M\left(\frac{242M^2}{\varepsilon^2} + 17\right).$$

Proof. Define $\varepsilon_1 = \varepsilon/M$ and $t = 242/\varepsilon_1^2 + 16$. Given any M-message protocol Π and any state $\rho^{A_{in}B_{in}R}$, let $\rho_1^{A_1C_1B_1R} = U_1(\rho \otimes \psi)$, $\rho_2^{A_2C_2B_2R} = U_2(\rho_1)$, $\cdots$, $\rho_M^{A_MC_MB_MR} = U_M(\rho_{M-1})$. Then, take $Q_i = \frac{1}{2\varepsilon_1^2} I(C_i;R|B_{i-1}) + t$. Then by Lemma 3.4.2, we have protocols Π_i, each with encoding and decoding maps V_1^i, V_2^i, along with corresponding entanglement $\phi_i \in D(T_A^i T_B^i)$ and communication register $\hat{C}^i$ of size $\dim \hat{C}^i = 2^{\lceil Q_i \rceil}$, with each satisfying

$$\|\Pi_i(\rho_i^{A_iC_iB_iR}) - \rho_i^{A_iC_iB_iR}\|_1 \leq \varepsilon_1. \tag{4.6.1}$$

We define the following protocol Π' starting from the protocol Π. The state ψ is the shared entanglement used in Π, and its isometries are $U_1, U_2, \cdots, U_M, U_{M+1}$. The state ϕ_i is the shared entanglement used in Π_i, and their isometries are V_1^i, V_2^i, respectively. Note that for even i, we will act V_1^i on Bob's side and V_2^i on Alice's side.

Protocol Π' on input σ in registers $A_{in}^{\otimes n}, B_{in}^{\otimes n}$ of $\Pi^{\otimes n}$

- Take entangled state $\hat{\psi} = \psi^{\otimes n} \otimes \phi_1 \otimes \cdots \otimes \phi_M$.
- Take unitaries $\hat{U}_1 = V_1^1 \circ U_1^{\otimes n}$, $\hat{U}_2 = V_1^2 \circ U_2^{\otimes n} \circ V_2^1$, $\cdots$, $\hat{U}_M = V_1^M \circ U_M^{\otimes n} \circ V_2^{M-1}$, $\hat{U}_{M+1} = U_{M+1}^{\otimes n} \circ V_2^M$.
- Take as output the $A_{out}^{\otimes n}, B_{out}^{\otimes n}$ registers of $\Pi^{\otimes n}$.

Note that the communication cost of Π' satisfies

$$\begin{aligned}
QCC(\Pi') &= \sum_i \log|\hat{C}^i| \\
&= \sum_i \lceil Q_i \rceil \\
&\leq \sum_{i>0,\ odd} \frac{61}{\varepsilon_1^2} I(C_i;R|B_{i-1}) + \sum_{i>0,\ even} \frac{61}{\varepsilon_1^2} I(C_i;R|A_{i-1}) + M(t+1) \\
&\leq \frac{32}{\varepsilon_1^2} QIC(\Pi,\rho) + M(t+1) \\
&\leq \frac{32M^2}{\varepsilon^2} QIC(\Pi,\rho) + M(\frac{242M^2}{\varepsilon^2} + 17).
\end{aligned}$$

This is also an M-message protocol. It remains to bound the error to make sure that Π' implements Π on ρ up to error ε. We have

$$\begin{aligned}
\|\Pi'(\rho) - \Pi(\rho)\| &= \| \operatorname{Tr}_{\neg A_{out}B_{out}} U_{M+1}V_2^M V_1^M U_M V_2^{M-1} \cdots V_1^1 U_1(\rho \otimes \hat{\psi}) \\
&\qquad - \operatorname{Tr}_{\neg A_{out}B_{out}} U_{M+1}U_M \cdots U_1(\rho \otimes \psi)\| \\
&= \| \operatorname{Tr}_{\neg A_{out}B_{out}} U_{M+1}\Pi_i U_M \Pi_{M-1} \cdots \Pi_1 U_1(\rho \otimes \psi) \\
&\qquad - \operatorname{Tr}_{\neg A_{out}B_{out}} U_{M+1}U_M \cdots U_1(\rho \otimes \psi)\| \\
&\leq \| \operatorname{Tr}_{(A')(B')} U_{M+1}\Pi_M \cdots \Pi_2 U_2 \Pi_1(\rho_1) \\
&\qquad - \operatorname{Tr}_{(A')(B')} U_{M+1}\Pi_M \cdots \Pi_2 U_2(\rho_1)\| \\
&\quad + \| \operatorname{Tr}_{(A')(B')} U_{M+1}\Pi_M U_M \Pi_{M-1} \cdots \Pi_3 U_3 \Pi_2(\rho_2) \\
&\qquad - \operatorname{Tr}_{(A')(B')} U_{M+1}\Pi_M U_M \Pi_{M-1} \cdots \Pi_3 U_3(\rho_2)\| \\
&\quad + \cdots \\
&\quad + \| \operatorname{Tr}_{(A')(B')} U_{M+1}\Pi_M U_M \Pi_{M-1}(\rho_{M-1}) \\
&\qquad - \operatorname{Tr}_{(A')(B')} U_{M+1}\Pi_M U_M(\rho_{M-1})\| \\
&\quad + \| \operatorname{Tr}_{(A')(B')} U_{M+1}\Pi_M(\rho_M) \\
&\qquad - \operatorname{Tr}_{(A')(B')} U_{M+1}(\rho_M)\| \\
&\leq \|\Pi_1(\rho_1\otimes) - (\rho_1)\| + \|\Pi_2(\rho_2) - (\rho_2)\|
\end{aligned}$$

$$
\begin{aligned}
&+\cdots \\
&+\|\Pi_{M-1}(\rho_{M-1})-(\rho_{M-1})\|+\|\Pi_M(\rho_M)-(\rho_M)\| \\
&\leq M\varepsilon_1 \\
&=\varepsilon.
\end{aligned}
$$

The first equality is by definition, the second one by taking the channel view for the protocols Π_i, since the corresponding A'_i, B'_i left at the end of these are traced out, and only U_1^i, U_2^i act on ϕ^i, the first inequality is by the triangle inequality and by definition of the ρ_i's, the second inequality is due to the monotonicity of trace distance under noisy channels, and the next is because $\Pi_i \in \mathscr{T}(\mathscr{R}_i, \rho_i, \varepsilon_1)$. □

4.6.3 Direct Sum for Bounded Round

By combining the above protocol compression result with many properties of quantum information complexity, we obtain the following direct sum theorem for distributional quantum communication complexity.

Theorem 4.6.3. *For a product classical task $\otimes_i(T_i, \mu_i, \varepsilon_i)$, any $\varepsilon \in (0, 1/2)$ and any number of message M,*

$$
QCC^M(\otimes_i(T_i,\mu_i,\varepsilon_i)) \geq \frac{\varepsilon^2}{32M^2}\sum_i \left(QCC^M(T_i,\mu_i,\varepsilon_i+\varepsilon) - (\frac{242M^2}{\varepsilon^2}+17)M \right).
$$

Proof. By Lemma 4.3.5, $QCC^M(\otimes_i(T_i,\mu_i,\varepsilon_i)) \geq \sum_i QIC^M(\otimes_i(T_i,\mu_i,\varepsilon_i))$. Given $\delta > 0$, let $\Pi_i \in \mathscr{T}^M(T_i,\mu_i,\varepsilon_i)$ be a protocol satisfying $QIC(\Pi_i,\mu_i) \leq QIC^M(T_i,\mu_i,\varepsilon_i)+\delta$ and let Π'_i be the compression protocol given by Lemma 4.6.2, satisfying $\Pi'_i \in \mathscr{T}^M(T_i,\mu_i,\varepsilon_i+\varepsilon)$ and $QCC(\Pi'_i) \leq \frac{32M^2}{\varepsilon^2} QIC(\Pi,\rho) + M(\frac{242M^2}{\varepsilon^2}+17)$. By rearranging terms, we get the result.

□

We can obtain a similar result for quantum tasks.

Theorem 4.6.4. *For a product quantum task $(\mathscr{N}_i, \rho_i, \varepsilon_i)$, any $\varepsilon \in (0, 1/2)$ and any number of message M,*

$$QCC^M(\otimes_i(\mathscr{N}_i, \rho_i, \varepsilon_i)) \geq \frac{\varepsilon^2}{32M^2} \sum_i \left(QCC^M(\mathscr{N}_i, \rho_i, \varepsilon_i + \varepsilon) - M(\frac{242M^2}{\varepsilon^2} + 17) \right).$$

4.6.3.1 Yao's Min-Max Theorem

In order to obtain a version of Theorem 4.6.3 for worst-case tasks, we also need a quantum version of Yao's Min-Max theorem.

Lemma 4.6.3. *For any ε and $\delta > 0$, any relation T and number of message M,*

$$QCC^M(T, \varepsilon + \delta) \leq \max_\mu QCC^M(T, \mu, \varepsilon),$$
$$QCC(T, \varepsilon + \delta) \leq \max_\mu QCC(T, \mu, \varepsilon).$$

Proof. Fix T, M, ε and denote $c = \max_\mu QCC^M(T, \mu, \varepsilon)$. For any $\delta > 0$, we want to prove the existence of a protocol $\Pi \in \mathscr{T}^M(T, \varepsilon + \delta)$ satisfying $QCC(\Pi) \leq c$. Recall that $P_e(\Pi, \mu)$ is the error of protocol Π on input μ. Fix a finite δ-net for $\mathscr{D}_{XY}$ containing all distributions with support of size one, that we denote N_{XY}. For each $\mu \in N_{XY}$, fix a protocol $\Pi_\mu \in \mathscr{T}^M(T, \mu, \varepsilon)$ such that $QCC(\Pi_\mu) \leq c$ and denote the set of all such protocols P_N. We then have $|P_N| = |N_{XY}| < \infty$. We define the following two-player zero-sum game over these two sets. Player A comes up with a quantum protocol $\Pi \in P_N$. Player B comes up with a distribution $\mu \in N_{XY}$. Player B's payoff is given by

$$P_B(\Pi, \mu) = P_e(\Pi, \mu),$$

and then player A's is given by $P_A(\Pi, \mu) = -P_B(\Pi, \mu)$. We first show the following.

Claim 4.6.1. *The value of the game for player B is bounded by $\varepsilon + \delta$.*

Proof. Let ν_B be a probability distribution over N_{XY} representing a mixed strategy for player B. To prove the claim, it suffices to show that there is a protocol $\Pi \in P_N$ such that

$\mathbb{E}_{\nu_B}[P_B(\Pi,\mu)] < \varepsilon + \delta$. Let $\bar{\mu}$ be the distribution corresponding to averaging over ν_B, that is

$$\bar{\mu}(x,y) = \mathbb{E}_{\nu_B}\mu(x,y).$$

Let $\mu' \in N_{XY}$ be a distribution that is δ-close to $\bar{\mu}$, and $\Pi' \in P_N$ the corresponding protocol. We will show that Π' is also good for $\bar{\mu}$. We have

$$\begin{aligned} P_e(\Pi',\bar{\mu}) &\leq P_e(\Pi',\mu') + \delta \\ &\leq \varepsilon + \delta, \end{aligned}$$

in which the first inequality follows from the fact that $\bar{\mu}$ and μ' are δ-close and the second inequality from the fact that $\Pi' \in P_N$ is the protocol corresponding to $\mu' \in N_{XY}$, i.e. $\Pi' \in \mathscr{T}^r(T,\mu',\varepsilon-\delta_2)$. Then

$$\begin{aligned} \mathbb{E}_{\nu_B}[P_B(\Pi',\mu)] &= \mathbb{E}_{\nu_B}[P_e(\Pi',\mu)] \\ &= P_e(\Pi',\bar{\mu}) \\ &< \varepsilon + \delta, \end{aligned}$$

in which the first equality is by definition, the second by linearity of expectation, and the inequality is by the above results about Π'. This concludes the proof of the claim. □

By the minimax theorem for zero-sum games, the above claim implies that there exists a probability distribution ν_A over P_N representing a mixed strategy for player A and such that the value of the game for player B is at most $1+\delta_1$. That is, for all $\mu \in N_{XY}$,

$$\mathbb{E}_{\nu_A}(P_B(\Pi,\mu)) < 1 + \delta_1.$$

Let $\bar{\Pi} = \mathbb{E}_{\nu_A}(\Pi)$ be the M-message protocol in the randomized model obtained by publicly averaging over ν_A. This is the protocol we are looking for. Indeed, it holds that

$\mathbb{E}_{\nu_A}(P_B(\Pi,\mu)) = P_e(\bar{\Pi},\mu)$ is the average error of $\bar{\Pi}$ on $\mu \in N_{XY}$. Letting μ run over all distributions with support of size one, it follows that $\bar{\Pi} \in \mathscr{T}^M(T,\varepsilon+\delta)$. The unbounded round result follows in the same way. □

We can now prove the direct sum result for worst-case classical tasks.

Theorem 4.6.5. *For a product classical task (T_i,ε_i), any $\varepsilon \in (0,1/2)$ and any number of message M,*

$$QCC^M(\otimes_i(T_i,\varepsilon_i)) \geq \frac{\varepsilon^2}{32M^2}\sum_i \left(QCC^M(T_i,\varepsilon_i+\varepsilon) - M(\frac{242M^2}{\varepsilon^2}+17)\right).$$

Proof. Similarly to the distributional case, we first have $QCC^M(\otimes_i(T_i,\varepsilon_i)) \geq \sum_i QIC^M(T_i,\varepsilon_i)$. Then, we use the fact that quantum information complexity is at least as large as max-distributional quantum information complexity, and use compression for each fixed input distribution in the optimization to obtain

$$\begin{aligned}\sum_i QIC^M(T_i,\varepsilon_i) &\geq \sum_i QIC^M_D(T_i,\varepsilon_i)\\ &= \sum_i \max_{\mu_i} QIC^M(T_i,\mu_i,\varepsilon_i)\\ &\geq \sum_i \max_{\mu_i} \frac{\varepsilon^2}{32M^2}\left(QCC^M(T_i,\mu_i,\varepsilon_i+\varepsilon) - M(\frac{242M^2}{\varepsilon^2}+17)\right).\end{aligned}$$

The result then follows by the quantum version of Yao's Min-Max theorem, Lemma 4.6.3, since δ can be taken arbitrarily small in it. □

4.7 Bounded-Round Disjointness

A further application of the quantum information complexity paradigm is to obtain powerful lower bounds on quantum communication complexity of specific functions. We provide an example by proving a tight lower bound, up to polylogarithmic terms, on the bounded round quantum communication complexity of disjointness.

4.7.1 Reduction from Disjointness to AND

Recall that the disjointness function computes, when viewing Alice's and Bob's inputs as subsets of $\{1,2,\cdots,n\}$, whether the sets are disjoint. That is, it is defined as follows: for $(x,y) \in \{0,1\}^n \times \{0,1\}^n$, $DISJ_n(x,y) = 1$ if for all $i \in [n]$, $x_i \wedge y_i = 0$, and 0 otherwise. In this section, we see how to reduce the quantum communication complexity of disjointness to the quantum information complexity of AND, for $AND(a,b) = a \wedge b$ the conjunction of the two bits a and b. We start with the following definition.

Definition 4.7.1. *For all $M \in \mathbb{N}, \varepsilon \in [0,1]$,*

$$QIC_0^M(AND,\varepsilon) = \inf_{\Pi \in \mathscr{T}^M(AND,\varepsilon)} \max_{\mu_0} QIC(\Pi,\mu_0),$$

in which the maximum ranges over all μ_0 satisfying $\mu_0(1,1) = 0$.

We can obtain a low-information protocol for AND from a protocol for disjointness.

Lemma 4.7.1. *For any n, M, $\varepsilon \in [0,1]$,$\Pi_D \in \mathscr{T}^M(DISJ_n,\varepsilon)$ and μ_0 such that $\mu_0(1,1) = 0$, there exists $\Pi_A \in \mathscr{T}^M(AND,\varepsilon)$ such that*

$$QIC(\Pi_A,\mu_0) = \frac{1}{n} QIC(\Pi_D,\mu_0^{\otimes n}).$$

Proof. Fix μ_0 and ε. We prove the result by induction on n. The base case is trivial since $DISJ_1 = \neg AND$, and so a protocol to compute $DISJ_1$ with error ε can be used to compute AND with error ε. For the induction step, suppose the result holds for $n-1$, we will use Lemma 4.3.4 to go from $DISJ_n$ to $DISJ_1$ and $DISJ_{n-1}$. Indeed, given Π_D computing $DISJ_n$ with error ε, we can use Lemma 4.3.4 with $\rho_1 = \mu_0$, $\rho_2 = \mu_0^{\otimes n-1}$ and then protocols Π^1 from the lemma computes $DISJ_1$ with error ε and Π^2 computes $DISJ_{n-1}$ with error ε, and they satisfy

$$QIC(\Pi_D,\mu_0^{\otimes n}) = QIC(\Pi^1,\mu_0) + QIC(\Pi^2,\mu_0^{\otimes n-1}).$$

By applying the induction hypothesis to Π^2, we obtain a protocol Π_A^2 computing AND

with error ε, and such that

$$QIC(\Pi_D, \mu_0^{\otimes n}) = n\left(\frac{1}{n}QIC(\Pi^1, \mu_0) + \frac{n-1}{n}QIC(\Pi_A^2, \mu_0)\right).$$

To obtain the claimed protocol Π_A, we then apply Lemma 4.3.7 to Π^1 and Π_A^2 with $p = 1/n$, and the average protocol then satisfy the conditions of the lemma. □

The following lemma is very similar to Theorem 4.3.1. The only difference is that the distributions we consider are restricted and on the right hand side the error of the protocol is measured in the worst case. Since the error is worst case, there is no loss in the error, and the payoff function would be simply $P_B(\Pi, \mu) = QIC(\Pi, \mu)/I$.

Lemma 4.7.2. *For all $M \in \mathbb{N}$,*

$$QIC_0^M(AND, \varepsilon) = \max_{\mu_0, \mu_0(1,1)=0} \inf_{\Pi \in \mathscr{T}^M(AND,\varepsilon)} QIC(\Pi, \mu_0)$$

Lemma 4.7.3. *For all M, $n \in \mathbb{N}$,*

$$QCC^M(DISJ_n, 1/3) \geq n \cdot QIC_0^M(AND, 1/3)$$

Proof. We have the following chain of inequality:

$$\begin{aligned} QCC^M(DISJ_n, 1/3) &\geq QIC^M(DISJ_n, 1/3) \\ &\geq \max_{\mu_0} \inf_{\Pi_D \in \mathscr{T}^M(DISJ_n, 1/3)} QIC(\Pi_D, \mu_0^{\otimes n}) \\ &\geq \max_{\mu_0} \inf_{\Pi_A \in \mathscr{T}^M(AND, 1/3)} n \cdot QIC(\Pi_A, \mu_0) \\ &\geq n \cdot QIC_0^M(AND, 1/3). \end{aligned}$$

The first inequality is by Lemma 4.3.19, the second since, on the right hand side, the maximization is over a smaller set of product distributions with $\mu_0(1,1) = 0$ and the minimization over a larger set of protocols, the third is because Lemma 4.7.1 implies the minimization is over a larger set of protocols, and the last is by Lemma 4.7.2. □

4.7.2 Optimal Bounds: Reducing Back to Disjointness

At this point, we need to show a $\tilde{\Omega}(1/M)$ lower bound on $QIC_0^M(AND, 1/3)$ to obtain the desired lower bound for disjointness. We only give a high-level overview of the technical arguments required from this point on, without providing proofs. We refer the interested reader to Ref. [40] for technical details. The main idea can be split into two parts.

First, we give a protocol for disjointness that runs the protocol for *AND* n times. But since running the AND protocol on an arbitrary input could incur a high quantum information cost, we start by doing some low communication preprocessing on the input using a trick from Ref. [38]. By subsampling a sublinear amount of coordinates and trying to find an intersection in them using distributed search as in Refs [1, 42], we can either say, after a communication in $o(\sqrt{n})$, that the set do intersect, or, except with small probability, that the number of intersection is small. Then, the prior-free quantum information cost of disjointness is upper bounded by n times the quantum information cost of AND on inputs with small weight on $(1,1)$. If we restrict the AND protocol to consist of M messages, this can only increase the information cost. But then, for M message protocols, we can use a continuity argument and boosting of the error to obtain a bound in term of $QIC_0^M(AND, 1/3)$.

Lemma 4.7.4. *For any* $\varepsilon > 0$ *and* $M \in \mathbb{N}$*,*

$$QIC_0^M(AND, \varepsilon) \leq O\big(\log 1/\varepsilon \cdot QIC_0^M(AND, 1/3)\big).$$

Proof. This can be proved in the same way as Lemma 4.3.22. □

Lemma 4.7.5. *Suppose we have a* M*-round protocol* Π *for AND. Then,*

$$QIC(\Pi, \mu) \leq QIC(\Pi, \mu_0) + O(MH(w)), \tag{4.7.1}$$

where $w = \mu(1,1) \leq 1/2$*,* $\mu_0(1,1) = 0$*, and* $\mu_0(x_i, y_i) = \frac{1}{1-w}\mu(x_i, y_i)$ *otherwise.*

Proof. This just follows from the proof of lemma 4.3.10, since the input size is constant.

□

Note that this is this only place in the argument where the number of message M appears explicitly in the bounds. We obtain the following result.

Lemma 4.7.6. *For any n, $M \in \mathbb{N}$,*

$$QIC(DISJ_n, 1/3) \leq O\big(n \cdot \log n \cdot QIC_0^M(AND, 1/3)\big) + n \cdot M \cdot H(w) + o(\sqrt{n}),$$

with $w = \frac{\log^4 n}{n}$.

The second part consists of proving a $\Omega(\sqrt{n})$ lower bound on $QIC(DISJ_n, 1/3)$. Note that $QIC_0^M(AND, 1/3) \in \tilde{\Omega}(1/M)$ then follows from this and the above result by an appropriate choice of $n \in \tilde{\theta}(M^2)$. It is probably possible to obtain such a result from the strong direct product theorem for the quantum communication complexity of disjointness of Klauck, Spalek and de Wolf [89], but the corresponding reduction would be much more complicated than the one we provide. Instead, we prove a stronger result, building on a threshold direct product theorem of Sherstov [119] for quantum communication complexity of Boolean function for which a tight bound is obtained through the generalized discrepancy method. This is one of the strongest method known to lower bound quantum communication complexity. We prove that the generalized discrepancy method also lower bounds quantum information complexity. Given a Boolean function f, denote by $GDM(f)$ the lower bound on $QCC(f, 1/3)$ obtained using the generalized discrepancy method. Note that $GDM(DISJ_n) \in \Omega(\sqrt{n})$. Since the asymptotic compression result for quantum information cost assumes a known input distribution, there is a lot of technical work going into the reduction from an average case to a worst case compression result satisfying the condition of Sherstov's threshold direct product theorem. We obtain the following result.

Lemma 4.7.7. *For any boolean function f, it holds that*

$$QIC(f, 1/3) \in \Omega(GDM(f)).$$

Note that a corollary of the above result is the first round independent upper bound on quantum communication complexity of Boolean functions in terms of their quantum information complexity.

Corollary 4.7.1. *For any Boolean function f, it holds that*

$$QCC(f,1/3) \in 2^{O(QIC(f,1/3)+1)}.$$

4.8 Conclusion

We conclude with a discussion of our results and further directions for this research program.

4.8.1 Discussion

We have defined a new notion of quantum information cost and a corresponding notion of quantum information complexity. In contrast to previously defined notions, these directly provide a lower bound on the communication, independently of round complexity. To define the quantum information cost of a protocol on an input quantum state, we take a detour through classical information cost and provide a different perspective on it, relating it to channel simulation with side information at the receiver, a variant of the classical reverse Shannon theorem studied in information theory. This provides a different proof that the information complexity is an achievable rate for amortized communication complexity, one that preserves the round complexity. Moreover, this characterization of the information cost in terms of the sum of the asymptotic communication cost for each channel simulation can be generalized to the quantum setting by considering the appropriate asymptotic task, which turns out to be equivalent to quantum state redistribution. Using this quantum generalization, we provide an operational interpretation in the distributional setting for quantum information complexity as the amortized quantum communication complexity, and in this sense provide the right quantum generalization of the classical information complexity. Along the way, we also show many important properties of the newly defined quantities. These are from Ref. [129]

and from joint work with Mark Braverman, Ankit Garg, Young Kun Ko and Jieming Mao [40]. For quantum protocols with classical inputs, we provide, in joint work with Mathieu Laurière [94], an alternative characterization of quantum information cost that avoids referring to purification registers, and quantifies the cost of forgetting classical information.

We use these new notions to obtain the first general direct sum theorem for quantum communication complexity that holds for multiple rounds of communication. This had been an open question since the first works on the direct sum question for quantum communication [82], and it was reiterated recently [7]. The approach we take is to exploit the link between this new, fully quantum notion of quantum information complexity that we introduce, and the task of quantum state redistribution. Protocol compression then builds upon a one-shot state redistribution protocol. This is from Refs [23, 129]. There is possibly still room for improvement in the dependence on the number of rounds for the direct sum theorem that we prove, but new techniques will probably be required in order to get substantial improvement over the parameters that we obtain. The fact that we are doing compression in a message-by-message fashion, with non-negligible error for each message compression, and at fixed length encoding, imposes severe limitations on the direct sum results that we can obtain.

However, the applicability of this notion of quantum information complexity to such a general, multi-round direct sum theorem, holding for all relations, provides further evidence that it is the correct quantum generalization of classical information complexity to consider in the standard communication complexity setting. Along with the operational interpretation as the amortized communication complexity and its application to lower bounding the bounded round quantum communication complexity of the disjointness function, it now appears clear that this is indeed the case. Finding such a quantum generalization was one of the open questions stated by Braverman in Ref. [33]. Another open question in Ref. [33] was to find the right generalization of information complexity to the multipartite setting. Note that for many issues that seemed to impose a limit on the development of a notion of quantum information complexity, like reversible protocols that can reveal to the players nothing but the output, there are similar issues that

seem to apply to a notion of classical multiparty information complexity. Thus, this new interpretation of (bipartite) classical information cost might prove useful to shed light on an appropriate notion of multiparty (classical) information cost.

Two of the main areas of success of classical information complexity is in obtaining direct sum and direct product theorems, and in obtaining communication complexity lower bounds, in particular on composite functions built from simpler component functions. A spectacular example of this is the recent result of Braverman, Garg, Pankratov, and Weinstein, proving the exact classical communication complexity of disjointness up to second order [38]. Quantum information complexity also satisfies a direct sum property for such composite functions. Indeed, on suitably chosen input distributions, the quantum information complexity of disjointness on n bits is at least equal to n times the quantum information complexity of the AND function on 2 bits. This is one of the properties we use in order to prove, in joint work with Mark Braverman, Ankit Garg, Young Kun Ko and Jieming Mao [40], a near-optimal bound on the bounded-round quantum communication complexity of disjointness. Along the way to proving this result, we prove that quantum information complexity also satisfies many important properties that make classical information complexity such a useful notion. A particular difference that we find in contrast with the classical notion is in continuity in the input, with a factor of the number of messages M proving to be necessary for some quantum protocols [40].

4.8.2 Open Questions

One potential bottleneck to obtaining other tight lower bounds using quantum information complexity is the fact that our notion of quantum information cost is defined in terms of a fully quantum conditional mutual information, a quantity that is much less understood than its classical counterpart. Obtaining meaningful lower bounds on quantum conditional mutual information is a notoriously hard problem in quantum information theory [98], with some progress in recent years [24, 29, 97]. A recent breakthrough result by Fawzi and Renner [60], yielding a lower bound on the quantum conditional mutual information in terms of the best recovery map acting on the conditioning system, will hopefully find application in the context of quantum information complexity.

We are confident that this new notion of quantum information complexity will stimulate interesting developments in quantum communication complexity, as well as in quantum information theory to obtain tools that would prove helpful for such developments. Interesting directions for this research program is first to obtain interesting lower bounds on the quantum information complexity of other functions, possibly by developing further techniques for lower bounding the conditional mutual information.

Other potential applications of this notion of quantum information complexity is in obtaining time-space trade-offs for quantum streaming algorithms [79], and obtaining the exact, up to second order, communication complexity of some problems, like the result in the classical setting that was recently obtained for the disjointness function [38]. Also, it would be interesting to investigate the general direct sum question in an unlimited round setting, and to try to obtain general direct product theorems, for which it is still an open question whether such theorems hold even for the simplest case of a single round of quantum communication.

CHAPTER 5

INTERACTIVE QUANTUM CAPACITY

5.1 Introduction

After introducing classical interactive coding and motivating the development of its quantum analogue, we discuss immediate difficulties in generalizing classical results to the quantum setting. We then present our approach to overcome these problems and prove the first results showing that quantum communication complexity is robust under noisy communication.

5.1.1 Classical Interactive Capacity

Quantum information theory is well developed for information transmission over noisy quantum channels, dating back to the work of Holevo in the 70's [71, 72], for the transmission of classical information [73, 116], quantum information [53, 99, 122], and even if we allow for pre-shared entanglement between sender and receiver [18, 19]. 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 [121, 124] by providing explicit and efficient constructions [44, 45, 65, 124], and by providing bounds on their existence [45, 61, 108].

Quantum communication complexity has also been studied in depth since Yao's seminal paper introduced the field in 1993 [139]. It is an idealized setting in which local computation is deemed free and communication noiseless but expensive. Quantum communication, even more so than classical communication, is prone to transmission errors in the real world. With the ubiquity of distributed computing nowadays, it has become increasingly important to develop information and coding theory for interactive protocols.

In the realm of classical communication, Schulman initiated the field with his pioneering works [112–114], showing that it is possible to simulate over a noisy channel any protocol designed to be run on a noiseless 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, the hidden constant of proportionality does not go to 1 asymptotically. For adversarial errors, Schulman also shows how to withstand corruption up to a rate of $\frac{1}{240}$. Recent work by Braverman and Rao [35] 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 [62] 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 simulations use *tree codes*, which were introduced by Schulman. Tree codes 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 proposed by Gelles, Moitra and Sahai [64], and they provided 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). In a worst-case adversarial scenario, the decoding might still take exponential time with potent tree codes. It was only recently that an alternative coding strategy developed by Brakerski and Kalai [27] was able to address the adversarial error case efficiently. Their strategy is to cleverly split the communication into blocks of logarithmic length in which tree encoding is used. In addition, they send, in between the blocks, some history information that enables efficient decoding. This construction was further improved by Brakerski and Naor [28]. A survey article by Braverman [32] 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 was recently investigated by Kol and Raz [90]. For this channel they find that indeed, in the low noise regime, the communication capacity behaves differently in the asymptotic limit of long interactive protocols than in the data transmission case.

5.1.2 Difficulties with Interactive Quantum Coding

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 inconsequential. In contrast, the no-cloning theorem of quantum theory [55, 136] rules out copying of quantum messages. As a result, if the information in some communication register is destroyed, 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. However, in a random error model, a constant dilation of each round would not be sufficient to achieve constant fidelity in the worst case of one-qubit transmission per round, 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 unless the number of rounds is constant: the adversary can always disrupt a whole block. The properties of classical information made it possible for Schulman and his successors to design clever classical simulation protocols that can withstand constant error rates at constant communication rates, and succeed in simulating classical protocols designed for noiseless channels over noisy channels by reproducing the whole transcript of the noiseless protocol. However, it was not obvious that it is possible, given an arbitrary 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 for protocols in the Cleve-Buhrman model, in which the communication is classical, it is not clear that we can achieve results similar to those for classical protocols. Indeed, a quantum measurement is in general irreversible. If such a measurement is performed on the shared entangled state and the players later realize that the measurement was based

on wrong classical information, the naive adaptation of the classical simulation to the Cleve-Buhrman model fails.

5.1.3 Overview of Results

Most of the results in this chapter are based on a collaboration with Gilles Brassard, Ashwin Nayak, Alain Tapp and Falk Unger [31]. We show that despite the above obstacles, 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 is analogous to Yao's model, and all communication in it is over noisy quantum channels, but the parties do not pre-share entanglement. The other is analogous to the Cleve-Buhrman model, and all communication in it is over noisy classical channels and parties are allowed to pre-share noiseless entanglement. We call these models the *quantum* and *shared entanglement models*, respectively. We also consider a further variation on the shared entanglement model in which entanglement is also noisy.

The main focus is on the model with perfect shared entanglement but adversarial noise on the classical communication. In such a context, the number of errors is defined to be the Hamming distance between the transcript of sent messages and the transcript of possibly corrupted received messages. Messages are over a constant size alphabet, and the error rate is the ratio between the number of errors introduced by the adversary in the worst-case and the number of such messages sent, i.e. the transcript length. Note that in this model, it is possible for the honest parties to generate a secret key unknown to the adversary by measuring their shared entanglement. Most of our technical contributions go into showing that a constant dilation factor on the communication suffices to withstand an adversarial error rate of $\frac{1}{2} - \varepsilon$ in the shared entanglement model, for arbitrarily small $\varepsilon > 0$. This is optimal, and matches the highest tolerable error rate in the analogous shared secret key model for classical interactive communication [62]. There are two main components going into establishing this result.

First, we need to establish a framework for simulating quantum protocols over noisy channels. To avoid losing quantum information, the approach we take is to teleport [15] the quantum communication register back and forth. When the register is in some party's

possession, this party 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 (but known) Pauli operators. Of particular importance to our work is the notion of tree codes as introduced by Schulman, which the players use to transmit classical information.

As described in a recent paper on efficient interactive coding [28], the high-level logic of all solutions proposed until now for classical protocol simulation can be summarized as follows: the parties try to evolve the protocol, and if they later realize there has been some error, they try to go back to the point where they last agreed (in a protocol tree representation, this would be their least common ancestor). In our approach for quantum protocols, the parties try to follow roughly the same idea, but for two reasons are not able to do this passively. First, there is no underlying transcript (or protocol tree) that the parties try to synchronize, except that they wish to evolve the correct sequence of unitaries. By the no-cloning theorem [55, 136], the parties cannot restart with a copy of the quantum information received up to some earlier point. Instead they have to actively rewind previous unitaries and wrong teleportation decodings until a suitable point in the protocol. Second, when they try to synchronize in this manner, they actively teleport, potentially leading to more errors on the joint quantum register.

An important ingredient in our simulation is the representation for noisy quantum protocols that we develop. As said before, in quantum protocols there is no direct analogue of a protocol tree representation that enables one to keep track exactly and explicitly of the evolution of the noiseless protocol simulation. The cleaned-up form (5.5.2) of our representation provides in some sense a quantum analogue of a protocol tree representation. As the classical representation, it enables an exact and explicit assessment of the evolution of the noiseless protocol simulation, as well as that of the departure from it due to noise.

At this point, it might look like we have reduced our problem to the classical case, since the parties only transmit classical information—the teleportation measurement outcomes. This enables us to reuse tools from classical interactive coding, most notably tree

codes, but the design of the quantum simulation protocol needs extra care. Unlike the classical case, agreement by the two parties on a common classical transcript is not sufficient. This transcript consists mostly of random teleportation measurement outcomes and is useless by itself. We additionally need to maintain a joint quantum state that eventually evolves according to the original protocol.

Once we realize the importance of teleportation in the context of noisy communication, and carefully design the simulation protocol, it may not come as a surprise that the simulation incurs only a constant factor overhead. The need for backtracking in the quantum simulation, however, seems to impose serious constraints on the tolerable error rate. *A priori* it is entirely unclear that we could hope to circumvent the low error tolerance seen in simulations with backtracking. The second part of our main contribution is to develop the necessary techniques to prove that we *can* tolerate an error rate as high as $\frac{1}{2} - \varepsilon$.

Indeed, all recent classical schemes tolerating high error rates have the property that the parties always go forward with the communication by using the tree structure of classical protocols. In comparison, in the original Schulman tree code based scheme there is some form of backtracking, due to which the scheme could only tolerate a much lower adversarial error rate of $\frac{1}{240}$. This is due to the fact that in a protocol with backtracking [114], the fraction of good rounds, in which both players correctly decode the tree code transmission, must be higher for the simulation to succeed than in a protocol that always goes forward by transmitting edges of a pointer jumping problem [35, 62]. There also is some form of backtracking in the outer level of the computationally efficient protocol of Ref. [27], thus limiting the overall error rate that can be tolerated to a fourth of that of the inefficient protocol used at the inner level. In light of these results, it is clear that previously used techniques would not suffice to tolerate error rates as high as $\frac{1}{2} - \varepsilon$ for our protocol, which requires backtracking. The new techniques we develop are thus necessary.

To achieve higher error tolerance, we follow Ref. [62] and use a *blueberry code* to effectively turn most adversarial errors into erasures. Concatenating such a code on top of a tree code yields a tree code with an erasure symbol. Since general transmission errors

are twice as harmful as erasures for the tree code condition, which is stated in terms of Hamming distance, it was shown in Ref. [62] that if the error rate is below $\frac{1}{2} - \varepsilon$, then the large number of rounds in which both parties correctly decode a long enough prefix is sufficient to imply success of the simulation. Once again due to backtracking, this condition is not sufficient for our purpose and in particular blueberry codes by themselves are not sufficient to improve error tolerance up to $\frac{1}{2}$ here. For us, 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. To circumvent this possibility, we need to bound the number of rounds with bad tree code decoding. Previously known bounds on this [114] can be used to show success of our simulation, but are far from enabling us to tolerate up to $\frac{1}{2}$ error rate. We develop a new bound on tree codes with an erasure symbol, Lemma 5.6.2, which might be of independent interest for classical interactive coding. This bound enables us to tightly control the number of rounds with bad decoding. Once we control this quantity, it is also important to insure that even when there is a corruption detected as an erasure in a round, as long as there is no bad decoding, the protocol will not need to spend a good round to correct for this previous erasure round.

We can adapt the techniques that we develop in the shared entanglement model for the quantum communication model: we first distribute a linear amount of entanglement using standard quantum information and coding theory techniques. This leads to a tolerable adversarial error rate of up to $\frac{1}{6}$ in the quantum model, close to the best achievable for quantum data transmission with zero error at $\frac{1}{4}$. This is better than the factor of two drop that might be expected if we compare classical interactive coding to unidirectional coding. 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 exists a family of binary functions for which a dilation factor proportional to $\frac{1}{C}$ is neces-

sary. When considering noisy entanglement in the form of noisy EPR pairs in a Werner state [132], we give, for any non-separable Werner state, simulation protocols with linear noisy classical communication and noisy EPR pair consumption. The techniques developed in this case can be adapted to show that the use of depolarizing channels in both directions enables the simulation to succeed whenever the quantum capacity with two-way classical communication, Q_2, is strictly positive. For some range of the depolarizing parameter, $Q = 0$ but $Q_2 > 0$, so this proves that Q does not characterize a quantum channel's capacity for interactive quantum communication.

Due to the use of tree codes, the protocols presented in this paper are not computationally efficient. However, it is possible to extend classical results on efficient interactive coding tolerating maximum error to noisy quantum communication. The representation of noisy protocols mentioned above is quite powerful and will be used in forthcoming papers to adapt classical results on computationally efficient interactive computation over adversarial channels [27] and on the interactive capacity of random noise channels [90] to the quantum regime.

Organization: This chapter is structured as follows: in Section 5.2, we set up the notation and state the relevant definitions for the classical part of our simulation protocols. In Section 5.3, we define the different models of noisy communication. In Section 5.4, we define the notion of interactive quantum capacity, and in the following section prove that it is strictly positive for many channels. In section 5.6, we state and prove our main result for the adversarial case in the shared entanglement model. Section 5.7 shows how to adapt the result of the previous section to obtain various other interesting results, in particular for the quantum model, the noisy shared entanglement model, and in the case of a random error model. We conclude with a discussion of our results and further research directions.

5.2 Classical Communication Protocols and Online Codes

Our simulation protocols contain an important classical component, so we set the notation for noiseless classical communication that we use, and define the online codes

that we use.

5.2.1 Noiseless Communication

In our setting, we are interested in protocols in which each party sends a message from some message set $[d] = \{1, 2, \cdots, d-1, d\}$ of size d in alternation, for some fixed number of rounds N' (actually, $\frac{N'}{2}$ 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 messages the parties wish to transmit, without the coding.

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 messages $a_1, a_2, \cdots, a_{i-1} \in [d]$ and $b_1, b_2, \cdots, b_{i-1} \in [d]$ Alice and Bob sent in the previous rounds, respectively. We refer to these sequences of messages (at the end of round i) as Alice's history $s_{\mathsf{A}} = a_1 \cdots a_i \in [d]^i$ and Bob's history $s_{\mathsf{B}} = b_1 \cdots b_i \in [d]^i$, respectively. Note that these histories are updated in each round, and that each history, 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 history tree. The whole (noiseless) communication can be extracted from the information in these two histories.

When the communication is noisy, in some rounds the parties make errors when trying to determine the other party's history. When comparing the history $s = s_1 \cdots s_i \in [d]^i$ of a party in round i of the protocol without coding, with the other party's best guess $s^i = s^i_1 \cdots s^i_i \in [d]^i$ for that history, the least common ancestor of s and s^i is the node at depth $i - \ell$ such that $s_1 \cdots s_{i-\ell} = s^i_1 \cdots s^i_{i-\ell}$ but $s_{i-\ell+1} \neq s^i_{i-\ell+1}$. We call ℓ the *magnitude* of the error of such a guess s^i, and in general for two histories $s, s^i \in [d]^i$ satisfying the above (with least common ancestor at depth $i - \ell$) we write $L(s, s^i) = \ell$. Note that we can compute ℓ as $i - \max\{t : (\forall j \leq t)[s_j = s^i_j]\}$.

5.2.2 Interactive Error Correcting Codes

Standard error correcting codes are designed for data transmission and therefore are not particularly well suited for interactive communication over noisy channels. We use two kinds of codes designed to work in an interactive scenario.

5.2.2.1 Tree Codes

In his breakthrough papers [113, 114], Schulman defined tree codes, which are particular codes designed for such interactive communication. Indeed, these tree codes can perform encoding and decoding round by round (following Ref. [62], 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 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 Σ and two words $\bar{e} = e_1 \cdots e_t \in \Sigma^t$ and $\bar{e}' = e'_1 \cdots e'_t \in \Sigma^t$ 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 5.2.1. (Tree codes [114]) *Given a message set $[d]$ of size $d > 1$, a number of rounds of communication $N' \in \mathbb{N}$, a distance parameter $\alpha \in (0,1)$ and a transmission alphabet Σ of size $|\Sigma| > d$, a d-ary* tree code *of depth N' and distance parameter α over alphabet Σ is defined by an encoding function $\mathcal{E} : [d]^{\leq N'} \to \Sigma$, and a decoding function $\mathcal{D} : \Sigma^{\leq N'} \to [d]^{\leq N'}$.*

Let $\bar{\mathcal{E}} : [d]^{\leq N'} \to \Sigma^{\leq N'}$ denote the extension of $\mathcal{E}$ to strings, i.e., for any $t \leq N'$, and $a = a_1 \cdots a_t \in [d]^t$,

$$\bar{\mathcal{E}}(a) = \mathcal{E}(a_1)\,\mathcal{E}(a_1 a_2) \cdots \mathcal{E}(a_1 \cdots a_{t-1})\,\mathcal{E}(a_1 \cdots a_t) \ ,$$

which is a string in Σ^t.

The encoding function satisfies the following distance property, called the tree code property. For any $t \leq N'$, and $a, a' \in [d]^t$,

$$L(a,a') = \ell \quad \Longrightarrow \quad \Delta(\bar{\mathcal{E}}(a), \bar{\mathcal{E}}(a')) \geq \alpha \cdot \ell \ .$$

In other words, if the least common ancestor of a, a' is at depth $t - \ell$, then the corresponding codewords are at distance at least $\alpha\ell$.

The decoding function satisfies the property that for any $t \leq N'$, and $\bar{e} \in \Sigma^t$,

$$\mathcal{D}(\bar{e}) \in \{a : a \in [d]^t \text{ minimizes } \Delta(\bar{\mathcal{E}}(a), \bar{e})\} \ .$$

We later consider decoding of tree codes with an erasure symbol $\perp$ that is not used by the encoding function, but may occur in the output of a channel. The decoding algorithm extends *verbatim* to received words with erasure symbols: it outputs a message sequence whose tree encoding is closest in Hamming distance to the received word.

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 $\mathcal{D}(\bar{e})$ whenever $|\{a : a \in [d]^t \text{ minimizes } \Delta(\bar{\mathcal{E}}(a), \bar{e})\}| > 1$. Also note that we can view tree codes in the following alternative way, connecting them with the history tree representation defined above. Starting with a history tree, we can label the arcs out of each node by a symbol from Σ corresponding to the encoding of that path in the tree code. The encoding function $\bar{\mathcal{E}}$ 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 history tree, and to the number of errors during these corresponding $L(a,a')$ last transmissions. The following was proved in Ref. [114] about the existence of tree codes. Let $\mathrm{H}(\alpha) = -\alpha \cdot \log \alpha - (1-\alpha) \cdot \log (1-\alpha)$ denote the binary entropy function.

Lemma 5.2.1. *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 Σ with $|\Sigma| = 2\lfloor (2 \cdot 2^{H(\alpha)} \cdot d)^{\frac{1}{1-\alpha}} \rfloor - 1$ suffices to label the arcs of some tree code, i.e., there exists an encoding function $\mathcal{E}$ satisfying the tree code property, and the required alphabet*

size is independent of N', the number of rounds of communication.

In fact, the result due to Schulman is even stronger: there exists an unbounded depth tree code with Σ 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 Ref. [62].

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 $t-\ell+1$ (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 ℓ rounds. Then, incorrect decoding (i.e., decoding to a message different from the one encoded) implies that there were 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 $\bar{e} = \bar{\mathcal{E}}(\bar{a}) \in \Sigma^t$, received as $\bar{e}'' \in \Sigma^t$, and decoded as $\bar{a}' = \mathcal{D}(\bar{e}'') \in [d]^t$, with $\bar{e}' = \mathcal{E}(\bar{a}')$, 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 $\Delta(e_{t-\ell+1} \cdots e_t, e''_{t-\ell+1} \cdots e''_t) \geq \frac{1}{2} \cdot \alpha \cdot \ell$ (Note $e_1 \cdots e_{t-\ell} = e'_1 \cdots e'_{t-\ell}$). This property is extremely useful for interactive communication: even if the decoding of a message is incorrect in some round, if there are sufficiently many error-free subsequent transmissions, we can later correct that error. This property is essential to our analysis of the simulation protocol, and to our proof of Lemma 5.6.2.

5.2.2.2 Blueberry Codes

Another kind of online code we need to withstand the highest possible error rates are randomized error detection codes called blueberry codes in Ref. [62]. 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 Ref. [93]. 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, transmission errors are twice as harmful as erasures in the tree decoding. (We can view the erasure flag $\perp$ as a special symbol in Σ never used in the encoding, but which helps in decoding.) When incorrect decoding occurs, the two parties might perform operations on the quantum registers that need to be corrected later. On the other hand, when an erasure occurs, it is visible to the recipient and this prevents him from performing such incorrect operations. Hence, concatenating a blueberry code with the tree code enables significant improvement in the allowed error rates.

These blueberry codes were defined in Ref. [62] for the purpose of authenticating streams of classical messages and for the simulation of interactive classical protocols. Below we summarize their definition and important properties.

Definition 5.2.2. (Blueberry codes [62]) *For $i \geq 1$ let $\mathcal{B}_i : \Gamma \to \Gamma$ be a random and independent permutation. The* blueberry code *maps a string $e \in \Sigma^t \subset \Gamma^t$ of arbitrary length t to $\mathcal{B}(e) = \mathcal{B}_1(e_1)\mathcal{B}_2(e_2)\cdots\mathcal{B}_t(e_t)$. We denote such a code as $\mathcal{B} : \Sigma^* \to \Gamma^*$, 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 5.2.3. *Assume that at some time i, $d_i = \mathcal{B}_i(e_i)$ is transmitted and $d'_i \neq d_i$ is received. If $\mathcal{B}_i^{-1}(d'_i) \notin \Sigma$, we mark the transmission as an erasure, and the decoding algorithm (for the Blueberry code) outputs $\perp$. Otherwise, this event is called an error.*

Corollary 5.2.1. *Let $e \in \Sigma^t$ and assume $\mathcal{B}(e)$ is communicated over a noisy channel. Every symbol corrupted by the channel causes either an error with probability ε_β, or an erasure with probability β.*

Lemma 5.2.2. *Assume a blueberry code $\mathcal{B} : \Sigma^* \to \Gamma^*$ 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 $(1 - 2\varepsilon_\beta)$-fraction of the ct* corrupted *transmissions are marked as erasures.*

Corollary 5.2.2. *If out of t received transmissions, ct were marked as erasures while decoding a blueberry code $\mathcal{B} : \Sigma^* \to \Gamma^*$, then except with probability $2^{-\Omega(t)}$ over the shared randomness, the adversarial corruption rate is at most $c/(1-2\varepsilon_\beta)$.*

5.3 Quantum Simulators and Interactive Channels

In this section, we describe the kind of protocols we want to simulate over noisy channels, and then formally define the models of noisy communication we study.

5.3.1 Noiseless Communication Model

In the *noiseless quantum communication model* that we want to simulate, we consider protocols in the hybrid model of a special form: throughout the whole protocol, the A register is held by Alice, the B register by Bob, the C register, which is the communication register, is exchanged back-and-forth between Alice and Bob and initially held by Alice, and finally the E register purifies the initial (and then also the final) state of the ABC registers and might be held by Eve, a potential adversary. For protocols in this form, we refer to N as their length and to each of the $N/2$ back-and-forth exchange between Alice and Bob as rounds. The initial state $|\psi_{\text{init}}\rangle^{ABCE} \in \mathscr{D}(A \otimes B \otimes C \otimes E)$ is chosen arbitrarily from the set of possible inputs, 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, which may be entangled with some reference system E. A protocol Π is then defined by the sequence of unitaries $U_1, U_2, \cdots, U_{N+1}$, with U_i for odd i known at least to Alice (or given to her in a black box) and acting on registers AC, and U_i for even i known at least to Bob (or given to him in a black box) and acting on registers BC. For simplicity, we assume that N is even. We can modify any protocol to satisfy this property, while increasing 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 final 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. The output state of the protocol is $\Pi(|\psi_{\text{init}}\rangle) = \text{Tr}_E(|\psi_{\text{final}}\rangle\langle\psi_{\text{final}}|^{ABCE})$, and by a slight abuse of notation we also represent the induced quantum channel from $ABCE$ to ABC simply by Π. 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. We restrict ourselves to the case of a single-qubit communication register C, which is the worst case for noisy interactive communication. Every protocol can be converted into such a form by increasing the communication by a factor of at most two but possibly at the expense of much more interaction: if a party has to speak when it is not his turn, he sends a qubit in state $|0\rangle$. Note that both the Yao and the Cleve-Buhrman 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 are an essential part of the Cleve-Buhrman model.

We later 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 D, $\tilde{D}$, let $\text{SWAP}_{D\leftrightarrow\tilde{D}}$ denote the quantum unitary that swaps the $D,\tilde{D}$ registers. Recall that N is assumed to be even. In a noiseless protocol embedding, for $i \in \{1,2,\cdots N-1\}$, we leave U_i untouched. We replace U_N by $(\text{SWAP}_{B\leftrightarrow\tilde{B}}U_N)$ and U_{N+1} by $(\text{SWAP}_{AC\leftrightarrow\tilde{A}\tilde{C}}U_{N+1})$. Finally, for $i \in \{N+2,N+3,\cdots N'+1\}$, we define $U_i = \text{I}$, 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 and C registers of the protocol to be simulated are subsystems of A', and B is one of B'. 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 $\mathscr{M}_1^{XA'C'}$, $\mathscr{M}_2^{YB'C'}$, and the output of the protocol on input $|\psi\rangle \in \mathscr{H}(A \otimes B \otimes C \otimes E)$ is the state of the ABC subsystem of $\mathscr{M}_2\mathscr{M}_1(|\psi\rangle)$, and is denoted $U(|\psi\rangle)$. By abuse of notation, the induced quantum channel from $ABCE$ to ABC is also denoted U.

5.3.2 Noisy Communication Model

There are many possible models for noisy communication. We consider two in particular: one analogous to the Yao model with no shared entanglement but noisy quantum communication, which we call the *quantum model*, and one analogous to the Cleve-Buhrman 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 5.7.4. For simplicity, we formally define in this section what we sometimes refer to as *alternating* communication models, in which Alice and Bob alternately transmit the communication register to each other, and this is the model in which most of our protocols are defined. Our definitions easily adapt to somewhat more general models which we call *oblivious* communication models, following Ref. [35]. In these models, Alice and Bob do not necessarily transmit their messages in alternation, but nevertheless in a fixed order and of fixed sizes known to all (Alice, Bob and Eve) depending only on the round, and not on the particular input or the actions of Eve. Communication models with a dependence on inputs or actions of Eve are called *adaptive* communication models.

5.3.2.1 Quantum Model

The following notions will be required to define adversaries in the quantum model. When considering a quantum system A of dimension q, we fix an orthonormal basis $\{|i\rangle\}_{i\in\{0,1,\cdots q-1\}}$ for A and use the following generalizations of Pauli operators: for $j,k \in \{0,1,\cdots q-1\}$, $\mathrm{X}^j|k\rangle = |(k+j) \mod q\rangle$ and $\mathrm{Z}^j|k\rangle = e^{\mathrm{i}2\pi\frac{jk}{q}}|k\rangle$. The operators in the set $\{\mathrm{X}^j\mathrm{Z}^k\}_{j,k\in\{0,1,q-1\}}$ are known as the Heisenberg-Weyl operators, and form a basis

for the linear vector space of operators on A, and the operators in

$$\mathscr{F}_{q,N} = \{\mathrm{X}^{j_1}\mathrm{Z}^{k_1} \otimes \cdots \otimes \mathrm{X}^{j_N}\mathrm{Z}^{k_N}\}_{j_\ell k_\ell \in \{0,1,\cdots q-1\}^2, \ell \in [N]} \tag{5.3.1}$$

form a basis for the space of operators on $A^{\otimes N}$. For $E \in \mathscr{F}_{q,N}$, we denote by $\mathrm{wt}(E)$ the weight of E, i.e., the number of A subsystems on which E acts non-trivially. For $\delta \in [0,1]$, the set

$$\mathscr{E}_{\delta,q,N} = \{E \in \mathscr{F}_{q,N} : \mathrm{wt}(E) \leq \delta N\} \tag{5.3.2}$$

is the subset of elements of $\mathscr{F}_{q,N}$ of weight less than or equal to δN.

Now, 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 implement a noiseless protocol Π as a black-box, the A and C_{A} parts correspond to the registers of the noiseless communication protocol, while $\tilde{A}$ and $\tilde{C}_{\mathsf{A}}$ 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 Π 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, passing through Eve; it is held by Alice at both the beginning and the end of the simulation protocol. A simulation protocol Q in the quantum model of length N' is defined by a sequence of quantum instruments $\mathscr{M}_1^{XA'C'}$, $\mathscr{M}_2^{YB'C'}$, $\cdots$, $\mathscr{M}_{N'+1}^{XA'C'}$ such that, on input a state $\left|\psi'_{\mathrm{init}}\right\rangle^{A'B'C'E'} = |\psi_{\mathrm{init}}\rangle^{ABC_{\mathsf{A}}E} \otimes |0\rangle$, given black-box access to a noiseless

protocol Π, and against an adversary $\mathcal{A}$ defined by a sequence of quantum instruments $\mathscr{N}_1^{ZE'C'}, \cdots, \mathscr{N}_{N'}^{ZE'C'}$, the protocol outputs the $\tilde{A}\tilde{B}\tilde{C}$ subsystems of

$$\rho_{\text{final}} = \mathscr{M}_{N'+1}^{\Pi} \mathscr{N}_{N'} \mathscr{M}_{N'}^{\Pi} \cdots \mathscr{M}_2^{\Pi} \mathscr{N}_1 \mathscr{M}_1^{\Pi}(|\psi'_{\text{init}}\rangle\langle\psi'_{\text{init}}|). \tag{5.3.3}$$

(Here, the superscript Π emphasizes the black-box access to the protocol.) We denote this output by $Q^{\Pi}(\mathcal{A}(|\psi_{\text{init}}\rangle))$, and the induced quantum channel from $ABCE$ to $\tilde{A}\tilde{B}\tilde{C} \cong ABC$ by $Q^{\Pi}(\mathcal{A})$. 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 5.3.1. *A simulation protocol Q in the quantum model of length N' succeeds with error ε at simulating all length N noiseless protocols against all adversaries in some class $\mathscr{A}$ if, for all noiseless protocols Π of length N, for all adversaries $\mathcal{A} \in \mathscr{A}$, $\|\Pi - Q^{\Pi}(\mathcal{A})\|_{\diamond} \leq \varepsilon$. The communication rate R_{Q} of Q is $R_{\mathrm{Q}} = \frac{N}{N' \log q}$ for $q \geq 2$ the alphabet size of the communication register C'.*

Note that the adversary only has to make the simulation fail on some particular protocol, and on some particular input, to characterize the simulation protocol as ineffective against her.

In a random error model (analogous to that studied in quantum information theory, à la Shannon), Eve is a non-malicious passive environment, and $\mathscr{N}_i = \mathscr{N}^{\mathcal{Q}}$ for some fixed quantum channel $\mathscr{N}^{\mathcal{Q}}$, and the class $\mathscr{A}$ contains a single element $\mathscr{N}^{C' \otimes N'}$, (with trivial Z, E' registers). For simplicity, we then say that the simulation succeeds over $\mathscr{N}^{\mathcal{Q}}$. 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 $\mathscr{A}^{\mathrm{Q}}_{\delta,q,N'}$ for some parameter δ such that $0 \leq \delta \leq 1$. The class $\mathscr{A}^{\mathrm{Q}}_{\delta,q,N'}$ contains all adversaries with a bound δ on the fraction of communications of the C' register they corrupt, in the following sense. Here, $\mathscr{F}_{q',1}, \mathscr{E}_{\delta,q,N'}$ are defined in Eqs. (5.3.1) and (5.3.2), respectively.

Definition 5.3.2. *The class $\mathscr{A}^{\mathrm{Q}}_{\delta,q,N'}$ of adversaries in the quantum model with* error rate *bounded by δ, $0 \leq \delta \leq 1$, contains adversaries of the following kind. Each adversary*

is specified by a sequence of instruments $\mathscr{N}_1^{ZE'C'_1}$, $\cdots$, $\mathscr{N}_{N'}^{ZE'C'_{N'}}$ *with arbitrary local quantum register* E' *of dimension* $q' \in \mathbb{N}$ *and local classical register* Z *with classical state set* $\mathscr{Z}$*, with* $|\mathscr{Z}| \in \mathbb{N}$*. All these adversaries act on a quantum communication register* C' *of dimension* $q \in \mathbb{N}$*, on protocols of length* $N' \in \mathbb{N}$*. For any* $\rho = \sum_{z_0 \in \mathscr{Z}} p_{Z_0}(z_0)|z_0\rangle\langle z_0|^Z \otimes \rho(z_0)^{E'C'^{\otimes N'}} \in \mathscr{D}(Z \otimes E' \otimes C'^{\otimes N'})$*, the action of such an adversary is*

$$\mathscr{N}_{N'}^{ZE'C'_{N'}} \cdots \mathscr{N}_1^{ZE'C'_1}(\rho) = \sum_{i,z,z_0} p_{Z_0}(z_0)|z\rangle\langle z|^Z \otimes G_{i,z,z_0}^{E'C'^{\otimes N'}} \rho(z_0) G_{i,z,z_0}^{\dagger E'C'^{\otimes N'}},$$

for i *ranging over some finite set,* z*,* $z_0 \in \mathscr{Z}$*, with each* G_{i,z,z_0} *of the form*

$$G_{i,z,z_0} = \sum_{H \in \mathscr{E}_{\delta,q,N'}, F \in \mathscr{F}_{q',1}} \alpha_{H,F,i,z,z_0} F^{E'} \otimes H^{C'^{\otimes N'}},$$

also subject to the requirement that for any $z_0 \in \mathscr{Z}$*,* $\sum_{i,z} G_{i,z,z_0}^\dagger G_{i,z,z_0} = \mathrm{I}^{E'C'^{\otimes N'}}$.

This adapts to an interactive communication model the formal definition of adversarial channel given in Ref. [96] in a unidirectional communication model. Note that this allows for adaptive, probabilistic, entangled strategies for Eve, but such that any Kraus operator G_{i,z,z_0} is a linear combination of operators which act on at most a δ fraction of the C' registers non-trivially. We therefore say that the fraction of errors is bounded by δ for all adversaries in $\mathscr{A}^{\mathrm{Q}}_{\delta,q,N'}$.

5.3.2.2 Shared Entanglement Model

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 $|\psi_{\mathrm{init}}\rangle^{ABCE}$, Alice and Bob also share entanglement to be consumed during the simulation in the form of a large state $|\phi\rangle^{T_\mathsf{A}T_\mathsf{B}}$ with the registers T_A, T_B held by Alice and Bob, respectively. In general, the entanglement registers have a product decomposition $T_\mathsf{A} = T_\mathsf{A}^1 \otimes \cdots \otimes T_\mathsf{A}^{N'}$, $T_\mathsf{B} = T_\mathsf{B}^1 \otimes \cdots \otimes T_\mathsf{B}^{N'}$. 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, passing through

Eve; it is held by Alice at both the beginning and the end of the simulation protocol. A simulation protocol S in the shared entanglement model of length N' is defined by a sequence of quantum instruments $\mathscr{M}_1^{XA'T_\mathsf{A}C''}, \mathscr{M}_2^{YB'T_\mathsf{B}C''}, \cdots, \mathscr{M}_{N'+1}^{XA'T_\mathsf{A}C''}$ such that, on input a state $\left|\psi'_{\text{init}}\right\rangle^{A'B'C''E'} = |\psi_{\text{init}}\rangle^{ABC_\mathsf{A}E} \otimes |0\rangle$, given black-box access to a noiseless protocol Π, and against an adversary $\mathcal{A}$ defined by a sequence of quantum instruments $\mathscr{N}_1^{ZE'C''}$, $\cdots$, $\mathscr{N}_{N'}^{ZE'C''}$, the protocol outputs the $\tilde{A}\tilde{B}\tilde{C}$ subsystems of the state ρ_{final} given by

$$\rho_{\text{final}} = \mathscr{M}_{N'+1}^{\Pi} \mathscr{N}_{N'} \mathscr{M}_{N'}^{\Pi} \cdots \mathscr{M}_2^{\Pi} \mathscr{N}_1 \mathscr{M}_1^{\Pi}(|\psi'_{\text{init}}\rangle\langle\psi'_{\text{init}}|). \tag{5.3.4}$$

(Again, the superscript Π emphasizes the black-box access to the protocol by the simulator.) We denote this output by $S^{\Pi}(\mathcal{A}(|\psi_{\text{init}}\rangle))$, and the induced quantum channel from $ABCE$ to $\tilde{A}\tilde{B}\tilde{C} \cong ABC$ by $S^{\Pi}(\mathcal{A})$. 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 5.3.3. *A simulation protocol S in the shared entanglement model of length N' succeeds with error ε at simulating all length N noiseless protocols against all adversaries in some class $\mathscr{A}$ if, for all noiseless protocols Π of length N, for all adversaries $\mathcal{A} \in \mathscr{A}$, $\|\Pi - S^{\Pi}(\mathcal{A})\|_\diamond \leq \varepsilon$. The communication rate R_C of S is $R_\mathrm{C} = \frac{N}{N' \log q}$ for $q \geq 2$ the alphabet size of the classical communication register C'', and the entanglement consumption rate R_E is $R_\mathrm{E} = \frac{\log(\max(\dim T_\mathsf{A}, \dim T_\mathsf{B}))}{N \log q}$ for $T_\mathsf{A}, T_\mathsf{B}$ 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 $\mathscr{N}_i = \mathscr{N}^S$ for some fixed classical channel $\mathscr{N}^S$, and the class $\mathscr{A}$ contains a single element $\mathscr{N}^{C''^{\otimes N'}}$ (with trivial Z, E' registers). For simplicity, we then say that the simulation succeeds over $\mathscr{N}^S$. 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 $\mathscr{A}^{\mathrm{S}}_{\delta,q,N'}$ for some parameter $0 \leq \delta \leq 1$. The class $\mathscr{A}^{\mathrm{S}}_{\delta,q,N'}$ contains all adversaries with a bound δ on the fraction of communications of the C'' register they corrupt, in the following sense. Here, for two strings c, c_0 over a finite alphabet, $\Delta(\cdot,\cdot)$ is the Hamming distance function counting the number of positions in which c and c_0 differ; see section

5.2.2.1 for a formal definition.

Definition 5.3.4. *The class $\mathscr{A}^{\mathrm{S}}_{\delta,q,N'}$ of adversaries with* error rate *bounded by δ, $0 \leq \delta \leq 1$, in the shared entanglement model contains adversaries of the following kind. Each adversary is specified by instruments $\mathscr{N}_1^{ZE'C''_1}$, $\cdots$, $\mathscr{N}_{N'}^{ZE'C''_{N'}}$ with arbitrary local quantum register E' of dimension $q' \in \mathbb{N}$ and local classical register Z with classical state set $\mathscr{Z}$, with $|\mathscr{Z}| \in \mathbb{N}$. All these instruments act on a classical communication register C'' of dimension $q \in \mathbb{N}$, on protocols of length $N' \in \mathbb{N}$. For any $\rho = \sum_{z_0 \in \mathscr{Z}} p_{Z_0}(z_0)|z_0\rangle\langle z_0|^Z \otimes \rho(z_0)^{E'C''^{\otimes N'}} \in \mathscr{D}(Z \otimes E' \otimes C''^{\otimes N'})$, the action of such an adversary is*

$$\mathscr{N}_{N'}^{ZE'C''_{N'}} \cdots \mathscr{N}_1^{ZE'C''_1}(\rho) = \sum_{c,c_0,z,z_0} p_{Z_0}(z_0)|z\rangle\langle z|^Z \otimes G^{E'C''^{\otimes N'}}_{c,c_0,z,z_0} \rho(z_0) G^{\dagger E'C''^{\otimes N'}}_{c,c_0,z,z_0} ,$$

for c, $c_0 \in \{0,1,\cdots,q-1\}^{N'}$ satisfying $\Delta(c,c_0) \leq \delta N'$, z, $z_0 \in \mathscr{Z}$, with each G_{c,c_0,z,z_0} of the form

$$G_{c,c_0,z,z_0} = \sum_{F \in \mathscr{F}_{q',1}} \alpha_{F,c,c_0,z,z_0} F^{E'} \otimes |c\rangle\langle c_0|^{C''^{\otimes N'}} ,$$

also subject to the requirement that for any $c_0 \in \{0,1,\cdots,q-1\}^{N'}$, $z_0 \in \mathscr{Z}$, $\sum_{c,z} G^{\dagger}_{c,c_0,z,z_0} G_{c,c_0,z,z_0} = \mathrm{I}^{E'} \otimes |c_0\rangle\langle c_0|^{C''^{\otimes N'}}$.

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), final transcript c on the communication register, inputs z_0, c_0, at most a δ fraction of the actions of Eve have acted non-trivially on the C'' register, even though she can copy all classical transmissions in the Z registers. We therefore say that the fraction of error is bounded by δ for all adversaries in $\mathscr{A}^{\mathrm{Q}}_{\delta}$.

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.

5.4 Definition of Interactive Quantum Capacity

In classical information theory, a single quantity characterizes a noisy channel's capacity to transmit information. Indeed, whether we consider the task of sending an arbitrary long message or the task of distributing randomness, the optimal rate of asymptotic communication, for probability of error tending to zero, is characterized by the same quantity, the noisy channel's capacity. This holds whether we allow the sender and receiver to share randomness or not at the outset, and even if we allow the receiver to send back-communication to the sender. In all of these tasks and in all of these settings, the channel capacity is the same, and evaluates to the maximum over all input distributions of the mutual information between the input and the output to the channel.

The situation is already much more complicated if we want to characterize a noisy channel's capacity to implement interactive communication. From the work of Schulman [114], we know that a channel with non-zero capacity for data transmission will have non-zero capacity to implement any interactive task, but it was only recently shown that there exists settings and tasks for which the capacity of a particular channel with low noise to implement an interactive task is strictly smaller than its capacity to implement unidirectional data transmission. However, still in the low noise regime, it is possible to have capacity close to one for the binary symmetric channel [90]. However, depending on the kind of task that we want to implement, a different model of communication might have to be considered, since in the noisy interactive setting, it is conceivable that Alice and Bob lose synchronisation and both try to speak simultaneously over the noisy channel.

In quantum information theory, even in the unidirectional setting, the question is much more complicated than for unidirectional classical information theory. First, we could be interested in the task of sending either classical or quantum information over

a noisy channel, or even a combination of both. Then, for each of these, we could be interested in transmitting different kind of information: for quantum information, the data to be transmitted could be an arbitrary, unknown quantum state, part of some larger entangled state, half of a maximally entangled state, etc. It turns out that in most settings, these tasks often define only two distinct capacities: the capacity to transmit classical or quantum information, however these are defined. But then, there are also a variety of settings to consider, and these yield much different quantities: the capacity of a channel without any assistance, when assisted by entanglement, when assisted by two-way classical communication, etc. There are even channels for which the capacity to transmit quantum information without assistance is zero, but becomes strictly positive if we allow classical back-communication or entanglement assistance.

In the interactive setting, we then also will have to distinguish between the kind of assistance that we allow, and the kind of task that we want to implement. We define the interactive capacity of a quantum channel (or of a set of channel as in the adversarial setting), with respect to an infinite set of task, classical or quantum, that we want to implement. This set is parametrized by a parameter n, such that for all n there exists a task in the set with quantum communication complexity at least n. We also define the interactive capacity with respect to the side resources that we want to allow. Note that if we allow for free noiseless classical communication, a quantum channel's capacity for interactive communication is equal to its capacity for quantum information transmission, which follows from teleportation. Hence, we will rather be interested in the case of no assistance for quantum channels, and of entanglement-assistance, either perfect or noisy, for either classical or quantum channels. We have the following definitions.

Definition 5.4.1. *The* quantum communication complexity of some task T over a noisy channel $\mathscr{M}$ assisted by some resources $\mathscr{Q}$ *is defined as the minimum number of uses of* $\mathscr{M}$*, in either direction and with assistance of resources* $\mathscr{Q}$*, required to successfully implement the task* T*, up to some error parameter implicit in* T*. It is denoted as* $QCC^{\mathscr{Q}}_{\mathscr{M}}(T)$*. If it is not possible to implement task* T *over* $\mathscr{M}$ *assisted by* $\mathscr{Q}$*, we define* $QCC^{\mathscr{Q}}_{\mathscr{M}}(T) = \infty$*.*

Definition 5.4.2. *The interactive capacity of channel* $\mathscr{M}$ *assisted by some resources* $\mathscr{Q}$

with respect to the set of task $\mathcal{T}$ is defined as

$$C_{\mathcal{T}}^{\mathcal{Q}}(\mathscr{M}) = \liminf_{n\to\infty}\{\frac{n}{QCC_{\mathscr{M}}^{\mathcal{Q}}(T)} : T \in \mathcal{T}, QCC(T) = n\}$$

if $QCC_{\mathscr{M}}^{\mathcal{Q}}(T) < \infty$ for all $T \in \mathscr{T}$, and 0 otherwise.

The focus in this chapter is in the very noisy regime, for which we show that a positive capacity for unidirectional transmission implies a positive capacity for interactive communication. This holds for an arbitrary set of quantum tasks. However, the simulation protocols could be much more interactive than the original protocols. Ideas developed here can be applied, together with classical techniques and appropriate data structures, to obtain high communication rates, and high capacity, in the low noise regime; see Section 5.8 for a further discussion on this.

5.5 Positive Interactive Quantum Capacity

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. This protocol contains the essential ideas of the optimal protocol of section 5.6, but the description and analysis are simplified because we do not have the additional blueberry code layer. Moreover, this protocol succeeds with perfect fidelity, provided the number of errors is below a certain threshold.

5.5.1 Result

We focus on the shared entanglement model. Techniques to distribute entanglement in both random [53, 99, 122] and adversarial [45, 61, 108] error models are well-studied. We can combine our findings with these entanglement distribution techniques to translate results in the shared entanglement model to the quantum model. We first focus on an adversarial model of error, and then adapt these results to a random error model. Such extensions to other models of communication are studied in Section 5.7.

It is already clear from the results in this section and Section 5.7 that for any set of tasks, any classical channel with positive capacity for data transmission will have positive interactive quantum capacity when assisted by perfect entanglement. It is also clear that a quantum channel with positive capacity for distributing entanglement will have positive capacity for interactive quantum communication without requiring any assistance.

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 5.6, 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.

Given an adversarial channel in the shared entanglement model with low enough error rate, we show how to simulate perfectly any noiseless protocol of length N over this channel using a number of transmissions linear in N, and consuming a linear number of EPR pairs. More precisely, we prove the following.

Theorem 5.5.1. *There exist a constant error rate $\delta > 0$, communication rate $R_{\mathrm{C}} > 0$, transmission alphabet size $q \in \mathbb{N}$, and entanglement consumption rate $R_{\mathrm{E}} \in \mathbb{R}^{+}$ 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 at least R_{C}, transmission alphabet size q, entanglement consumption rate at most R_{E}, which succeeds with zero error at simulating all noiseless protocols of length N against all adversaries in $\mathscr{A}^{\mathrm{S}}_{\delta,q,N'}$.*

Specific values for the constants posited in the theorem are given at the end of Section 5.5.4.

5.5.2 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 $T_\mathsf{A}T_\mathsf{B}$ to do so, as well as the noisy classical channels to transmit their teleportation measurement outcomes. 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 they later realize 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 points to notice here. First, the sequence of operations acting on the joint register is a sequence of reversible unitaries. Hence, if the parties keep track of the sequence of operations on the joint register, at least one of the parties can reverse any of his 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 $\pm \mathrm{X}^x\mathrm{Z}^z$ is applied at some point, but only one knows the correct value of $xz \in \{0,1\}^2$, and similarly both know U_j^M (with $U_j^{+1} = U_j$, $U_j^{-1} = U_j^\dagger$, $U_j^0 = \mathrm{I}$) is applied at some point, but only one knows the correct values of $j \in \{1, \cdots, N'+1\}$ and $M \in \{-1, 0, +1\}$. This is the classical information they try to transmit to each other so that both know exactly the sequence of operations that have been applied on the joint register. The tree codes due to Schulman are particularly well suited for protecting against noise 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. The goal is that the parties know exactly where they are in the simulation of the protocol (i.e., the sequence of unitaries that have been applied to the virtual protocol registers) when they are able to correctly decode the classical messages sent by the other party. To enable a party to learn exactly what action

was taken by the other party in the earlier rounds, the message sent in each round is in $\{0,1\}^2 \times \{-1,0,+1\} \times \{0,1\}^2$, encoded with a tree code. The first pair of bits corresponds to the teleportation decoding operation done at the beginning of a party's turn. The trit is associated with the evolution in the noiseless protocol: $+1$ stands for going forward with the protocol, i.e., for a unitary operator of the noiseless protocol that was applied to the joint state of the party's local register and the communication register; -1 stands for going backwards with the protocol, i.e., for the inverse of a unitary of the noiseless protocol that was applied by that party to the joint state; 0 stands for holding the protocol idle, i.e., no action is taken by that party to evolve the protocol in that round. Note that the index j of the unitary U_j^M 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 *history* at some point the sequence of these triplets of messages he transmitted up to that point (see section 5.2). If a party succeeds in correctly decoding the history of the other party, he then possesses all the information about the operations that 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 eventually decode this message correctly), the other party still needs to be informed that the decoding of the teleportation in that particular round was different from what it should have been. 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 need be. We indeed use this property in the simulation (especially in the simulation for maximal error tolerance in Section 5.6.)

5.5.3 Description of the Simulator

All communication is done with a tree encoding over some alphabet Σ. To later simplify the analysis, we fix the distance parameter to $\alpha = \frac{39}{40}$. The message set consists 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 5.2.1, we know that there exists a $q \in \mathbb{N}$ independent of N' such that an alphabet Σ of size q suffices to label the arcs of a tree code of any depth $N' \in \mathbb{N}$. Both parties agree before the protocol begins on such a tree code of depth N' with corresponding encoding and decoding functions $\mathcal{E}$ and $\mathcal{D}$ (each party uses a separate instance of the same tree code to transmit her/his messages to the other party). The goal is to tolerate error rate up to $\delta = \frac{1}{80}$.

The convention we use for the variables describing the protocol is the following. On Alice's side, in round i, $x_i^{\mathsf{AD}} z_i^{\mathsf{AD}} \in \{0,1\}^2$ correspond to the bits she uses for the teleportation decoding on the X and Z Pauli operators, respectively, $x_i^{\mathsf{AM}} z_i^{\mathsf{AM}} \in \{0,1\}^2$ correspond to the bits of the teleportation measurement on the corresponding Pauli operators, $j_i^{\mathsf{A}} \in \mathbb{Z}$ and $M_i^{\mathsf{A}} \in \{-1,0,+1\}$ correspond respectively to the index of the unitary she uses in round i and to whether she uses $U_{j_i^{\mathsf{A}}}^{+1} = U_{j_i^{\mathsf{A}}}$, its inverse $U_{j_i^{\mathsf{A}}}^{-1} = U_{j_i^{\mathsf{A}}}^{\dagger}$, or simply applies the identity channel $U_{j_i^{\mathsf{A}}}^{0} = \mathrm{I}$ on the AC quantum register, and the counter c_i^{A} keeps track of the sum of all previous messages M_ℓ^{A}, $l \leq i$. On Bob's side, we use a similar set of variables, with superscript B instead of A. All Pauli operators are applied on the virtual C register. 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. Similarly, the superscript i is used when discussing other variables that are round dependent.

The actions Alice and Bob take in round i are based on their best guess for the state $|\psi_i\rangle$ of the joint register at the beginning of round i. The state $|\psi_i\rangle$ can be classically computed from the information in Alice's and Bob's histories. The analysis rests on the following two representations for the state $|\psi_i\rangle$. The first one can be directly computed,

up to irrelevant operations of Eve on the E register, as

$$|\psi_i\rangle^{ABCE} = \prod_{\ell=1}^{i-1} \left(\mathrm{X}^{x_\ell^{\mathsf{BM}}} \mathrm{Z}^{z_\ell^{\mathsf{BM}}} U_{j_\ell^{\mathsf{B}}}^{M_\ell^{\mathsf{B}}} \mathrm{Z}^{z_\ell^{\mathsf{BD}}} \mathrm{X}^{x_\ell^{\mathsf{BD}}} \mathrm{X}^{x_\ell^{\mathsf{AM}}} \mathrm{Z}^{z_\ell^{\mathsf{AM}}} U_{j_\ell^{\mathsf{A}}}^{M_\ell^{\mathsf{A}}} \mathrm{Z}^{z_\ell^{\mathsf{AD}}} \mathrm{X}^{x_\ell^{\mathsf{AD}}} \right) |\psi_{\text{init}}\rangle^{ABCE} . \tag{5.5.1}$$

Here, from the history s_{A} of Alice's history tree, we can directly obtain from the ℓth message sent by Alice, for $\ell = 1 \cdots i-1$, the two bits $x_\ell^{\mathsf{AD}} z_\ell^{\mathsf{AD}}$ used to decode the teleportation, the trit M_ℓ^{A} corresponding to the evolution of the protocol performed in round ℓ, and then the two bits $x_\ell^{\mathsf{AM}} z_\ell^{\mathsf{AM}}$ corresponding to the outcome of the teleportation measurement. We then use counters c_ℓ^{A}'s that maintain the sums of the M_ℓ^{A}'s to compute the indices j_ℓ^{A}'s of the noiseless protocol unitaries used by Alice in round ℓ: $c_0^{\mathsf{A}} = 0, c_\ell^{\mathsf{A}} = c_{(\ell-1)}^{\mathsf{A}} + M_\ell^{\mathsf{A}}, j_\ell^{\mathsf{A}} = 2c_{(\ell-1)}^{\mathsf{A}} + M_\ell^{\mathsf{A}}$. Note that j_i^{A} depends only on the sequence of messages $M_1^{\mathsf{A}}, M_2^{\mathsf{A}}, \cdots, M_{(i-1)}^{\mathsf{A}}, M_i^{\mathsf{A}}$. Similarly, the history s_{B} of Bob's history tree is used to obtain $x_\ell^{\mathsf{BD}} z_\ell^{\mathsf{BD}}, x_\ell^{\mathsf{BM}} z_\ell^{\mathsf{BM}}$, as well as M_ℓ^{B}, and to compute $c_0^{\mathsf{B}} = 0, c_\ell^{\mathsf{B}} = c_{(\ell-1)}^{\mathsf{B}} + M_\ell^{\mathsf{B}}, j_\ell^{\mathsf{B}} = 2c_{(\ell-1)}^{\mathsf{B}} + M_\ell^{\mathsf{B}} + 1$. We define $U_j^M = \mathrm{I}$ whenever $j \leq 0$ or $M = 0$. Note that if $M_\ell^{\mathsf{A}} \neq 0$, j_ℓ^{A} is odd and $U_{j_\ell^{\mathsf{A}}}^M$ acts on Alice's side. Similarly, if $M_\ell^{\mathsf{B}} \neq 0$, j_ℓ^{B} is even and $U_{j_\ell^{\mathsf{B}}}^M$ acts on Bob's side. Also note that $j \leq N' + 1$ so the U_j's are well-defined, by the noiseless protocol embedding described in Section 5.3.1.

From this first representation of the state $|\psi_i\rangle$, we can classically compute a second one by recursively "cleaning up" the first representation. The clean-up is performed by combining as many of the operators as possible, as follows. We multiply all consecutive Pauli operators acting on the C register, and simplify consecutive pairs of operators U_ℓ, U_ℓ^{-1} acting on the same set of qubits, to obtain a state of the form:

$$|\psi_i\rangle^{ABCE} = \hat{\sigma}^i \, \tilde{U}_{t_i}^i \, \tilde{\sigma}_{t_i}^i \, \tilde{U}_{t_i-1}^i \, \tilde{\sigma}_{t_i-1}^i \, \cdots \, \tilde{U}_2^i \, \tilde{\sigma}_2^i \, \tilde{U}_1^i \, \tilde{\sigma}_1^i \, U_{r_i} U_{r_i-1} \cdots U_2 U_1 \, |\psi_{\text{init}}\rangle^{ABCE} \tag{5.5.2}$$

with $\hat{\sigma}^i = \pm \mathrm{X}^{\hat{x}^i} \mathrm{Z}^{\hat{z}^i}$, and for $\ell \in \{1, \ldots, t_i\}$, $\tilde{\sigma}_\ell^i = \mathrm{X}^{x_\ell^i} \mathrm{Z}^{z_\ell^i}$ for $\hat{x}^i \hat{z}^i, x_\ell^i z_\ell^i \in \{0,1\}^2$, and $\tilde{U}_\ell^i = U_{\ell'}^{\pm 1}$ for some $r_i - 2t_i \leq \ell' \leq r_i + 2t_i$. The rules used recursively to perform the clean-up are the following: in the case that $\tilde{\sigma}_\ell^i = \mathrm{I}$, we require that for two consecutive unitary operators acting on the same set of qubits, if $\ell > 1$, then $\tilde{U}_\ell^i \neq (\tilde{U}_{\ell-1}^i)^{-1}$, and if $\ell = 1$,

then $\tilde{U}_1^i \neq U_{r_i+1}$ and $\tilde{U}_1^i \neq U_{r_i}^{-1}$. This last rule is what determines the cut between U_{r_i} and $\tilde{U}_1^i\tilde{\sigma}_1^i$. 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 evolve the state as in the noiseless protocol. Note that this is well-defined: there is a unique representation in the form (5.5.2) corresponding to any in the form (5.5.1).

To decide which action to take in round i, Alice starts by decoding the possibly corrupted messages $f'_1,\ldots,f'_{i-1} \in \Sigma$ received from Bob up to this point to obtain her best guess $s_{\mathsf{B}}^i = \mathcal{D}(f'_1,\ldots,f'_{i-1})$ for the history s_{B} of his history tree. Along with the history s_{A} of her history tree, she uses it to compute her best guess of the form (5.5.2) of the joint state. If her decoding of Bob's history is *good* (error-free), then she has all the information she needs to compute 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 $t_i = 0$ or not, according to her best guess for the state $|\psi_i\rangle$). 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 $M_i^{\mathsf{A}} = +1$ 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 $M_i^{\mathsf{A}} = 0$. 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 $\tilde{U}_{t_i} = U_{\ell'}^{M'}$ for ℓ' odd, then she sets $M_i^{\mathsf{A}} = -M' \in \{\pm 1\}$ (note that in this case it holds that $j_i^{\mathsf{A}} = \ell'$), else she sets $M_i^{\mathsf{A}} = 0$ and she hopes Bob will next correct $\tilde{U}_{t_i}$. In all cases, with $\hat{\sigma}_i^C = \pm \mathrm{X}^{\hat{x}_i}\mathrm{Z}^{\hat{z}_i}$, she sets $x_i^{\mathsf{AD}} = \hat{x}_i, z_i^{\mathsf{AD}} = \hat{z}_i$ and computes $c_i^{\mathsf{A}} = c_{(i-1)}^{\mathsf{A}} + M_i^{\mathsf{A}}, j_i^{\mathsf{A}} = 2c_{(i-1)}^{\mathsf{A}} + M_i^{\mathsf{A}}$. Note that she does not care about the global phase factor ± 1 appearing in $\hat{\sigma}_i$ during the clean-up from the form (5.5.1) to the form (5.5.2). This phase arises because the Pauli operators X and Z anticommute, and is irrelevant.

After this classical preprocessing, she can now perform her quantum operations on the AC registers: she first decodes the teleportation operation (and possibly some other Pauli errors remaining on the C register) by applying $\mathrm{Z}^{z_i^{\mathsf{AD}}}\mathrm{X}^{x_i^{\mathsf{AD}}}$ on the $T_{\mathsf{A}}^{2(i-1)}$ register

before swapping registers $T_{\mathsf{A}}^{2(i-1)}$ and C_{A}, effectively putting the virtual C register into C_{A}. (Note that in round 1, Alice already possesses the C register so this part is trivial: we let $T_{\mathsf{A}}^{0} = C_{\mathsf{A}}$ and set $x_1^{\mathsf{AD}} z_1^{\mathsf{AD}} = 00$.) She then performs $U_{j_i^{\mathsf{A}}}^{M_i^{\mathsf{A}}}$ 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_{A}^{2i-1} register, obtaining measurement outcome $x_i^{\mathsf{AM}} z_i^{\mathsf{AM}} \in \{0,1\}^2$. She updates her history s_{A} by following the edge $a_i = (x_i^{\mathsf{AD}} z_i^{\mathsf{AD}}, M_i^{\mathsf{A}}, x_i^{\mathsf{AM}} z_i^{\mathsf{AM}})$ in the history tree, and transmits message $e_i = \mathcal{E}(a_1 \cdots a_i)$ over the noisy classical channel, with $\mathcal{E}$ the encoding function of the tree code.

Upon receiving the message e_i', a possibly corrupted version of e_i, Bob obtains his best guess s_{A}^i for Alice's history s_{A} by computing, with previous messages $e_1' \cdots e_{i-1}'$, $s_{\mathsf{A}}^i = \mathcal{D}(e_1' \cdots e_i')$. He uses it along with his own history s_{B} to compute his best guess of the representation of the state

$$\left(\mathrm{X}^{x_i^{\mathsf{AM}}} \mathrm{Z}^{z_i^{\mathsf{AM}}} U_{j_i^{\mathsf{A}}}^{M_i^{\mathsf{A}}} \mathrm{Z}^{z_i^{\mathsf{AD}}} \mathrm{X}^{x_i^{\mathsf{AD}}} \right) |\psi_i\rangle \tag{5.5.3}$$

analogous to that in (5.5.1). He then cleans this up to obtain a representation analogous to that in (5.5.2), and based on this latest representation chooses in the same way as Alice his $x_i^{\mathsf{BD}} z_i^{\mathsf{BD}}, M_i^{\mathsf{B}}$, and then uses M_i^{B} and c_{i-1}^{B} to compute $c_i^{\mathsf{B}}, j_i^{\mathsf{B}}$. After this classical preprocessing, he can then perform his quantum operations: he first decodes the teleportation operation by applying $\mathrm{Z}^{z_i^{\mathsf{BD}}} \mathrm{X}^{x_i^{\mathsf{BD}}}$ on the T_{B}^{2i-1} register and by swapping it with C_{B}, creating a virtual C register, then performs $U_{j_i^{\mathsf{B}}}^{M_i^{\mathsf{B}}}$ 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 $x_i^{\mathsf{BM}} z_i^{\mathsf{BM}}$. He updates his history s_{B} by following the edge $b_i = (x_i^{\mathsf{BD}} z_i^{\mathsf{BD}}, M_i^{\mathsf{B}}, x_i^{\mathsf{BM}} z_i^{\mathsf{BM}})$, and transmits message $f_i = \mathcal{E}(b_1 \cdots b_i)$ over the channel. The round is completed 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 5.3.1), and use them as their respective outcomes for the protocol. If the simulation is successful the output quantum state corresponds to the ABC subsystem of $|\psi_{\text{final}}\rangle^{ABCE}$ specified by the original noiseless protocol. We later prove that the protocol is successful

if the error rate is below $\frac{1}{80}$.

We summarize the protocol below. Alice and Bob start with the state $|\psi_{\text{init}}\rangle$ in the registers $ABC_{\mathsf{A}}E$, the register C_{B} initialized to $|0\rangle$, the registers $T_{\mathsf{A}}T_{\mathsf{B}}$ initialized to N' EPR pairs $\left[\frac{1}{\sqrt{2}}(|00\rangle+|11\rangle)\right]^{\otimes N'}$, with one qubit from each EPR pair held by Alice, one by Bob, and the qubits in registers $\tilde{A},\tilde{B},\tilde{C}$ initialized to $|0\rangle$ (cf. the noiseless protocol embedding described in Section 5.3.1). They also have access to a suitable amount of classical workspace for local computations required for the simulation. They repeat the following for $i = 1, \ldots, \frac{N'}{2}$:

1. If $i > 1$, Alice computes $s_{\mathsf{B}}^i = \mathcal{D}(f_1' \cdots f_{i-1}')$, and extracts $b_\ell^i = (x_\ell^{i\mathsf{BD}} z_\ell^{i\mathsf{BD}}, M_\ell^{i\mathsf{B}}, x_\ell^{i\mathsf{BM}} z_\ell^{i\mathsf{BM}})$ for $\ell = 1, \ldots, i-1$. These are her best guesses for Bob's messages. She computes the corresponding $c_\ell^{i\mathsf{B}}, j_\ell^{i\mathsf{B}}$. For $i = 1$, the values of the parameters Alice needs for the simulation are straightforward.

2. Also using s_{A}, she computes her best guess for the form (5.5.2) of the state $|\psi_i\rangle$ of the joint register, and the corresponding $x_i^{\mathsf{AD}} z_i^{\mathsf{AD}}$, M_i^{A}, c_i^{A}, j_i^{A}, as described earlier in this section.

3. If $i > 1$, she completes the teleportation operation by applying $\mathrm{Z}^{z_i^{\mathsf{AD}}} \mathrm{X}^{x_i^{\mathsf{AD}}}$ to register $T_{\mathsf{A}}^{2(i-1)}$ and swaps this with the C_{A} register.

4. She applies $U_{j_i^{\mathsf{A}}}^{M_i^{\mathsf{A}}}$ to the AC_{A} register, in an attempt to evolve the original protocol.

5. She teleports the C_{A} register to Bob using entanglement in register T_{A}^{2i-1} and gets outcomes $x_i^{\mathsf{AM}} z_i^{\mathsf{AM}}$.

6. Alice updates her state s_{A} by following edge $a_i = (x_i^{\mathsf{AD}} z_i^{\mathsf{AD}}, M_i^{\mathsf{A}}, x_i^{\mathsf{AM}} z_i^{\mathsf{AM}})$ and transmits message $e_i = \mathcal{E}(a_1 \cdots a_i)$ using the noisy classical channel to Bob, who receives e_i', a possibly corrupted version of e_i.

7. Bob computes $s_{\mathsf{A}}^i = \mathcal{D}(e_1' \cdots e_i')$ and also using s_{B}, performs actions on his side analogous to Alice's. He completes the teleportation operation, swaps register T_{B}^{2i-1} with C_{B}, applies the appropriate unitary operation to the register $C_{\mathsf{B}}B$, uses the T_{B}^{2i} register to teleport the C_{B} register to Alice, and finally transmits f_i. Round i is completed when Alice receives 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.

5.5.4 Analysis

The analysis is done conditional on some overall classical state (and in particular, some respective views of Alice, Bob and Eve 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 return to this issue later.

The total number of rounds is $\frac{N'}{2}$, with two transmissions per round, for a total of N' transmissions. We define two kinds of rounds: *good* rounds in which both parties decode correctly the other party's history, and *bad* rounds in which at least one party makes a decoding error. To analyse the protocol, we define a "potential function" $P(i) \in \mathbb{Z}$ which increases at least by some (strictly positive) amount in good rounds, and decreases by at most some other (bounded) amount in bad rounds. The potential function is such that we know the simulation succeeds whenever $P(\frac{N'}{2}+1) \geq N+1$. 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 (5.5.2) for the form of the quantum state of the joint registers at the beginning of round i (or equivalently, at the end of round $i-1$). Recall that 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. Define

$$P(i) = r_i - 2t_i. \tag{5.5.4}$$

The factor of 2 in front of t_i is to account for the worst case scenario for the simulation in round i. As will be apparent from our analysis below, in the worst case, all remaining $\tilde{U}_l^i$'s are applied by the same party who applied U_{r_i-1} and $\tilde{U}_{t_i}^i = U_{r_i-1-2(t_i-1)}^{-1}$. Then, if $P(\frac{N'}{2}+1) \geq N+1$, the operators $\tilde{U}_\ell^i$ in (5.5.2) at the end of the simulation (i.e., with $i =$

$N'+1$) may only be equal to the identity operator, as ensured by the noiseless protocol embedding. Thus the output of the simulation is correct. We now prove the following technical lemma which bounds $P(i)$ as a function of the number of good and bad rounds.

Lemma 5.5.1. *At the end of round i, define*

$$\begin{aligned} N_{\mathrm{g}}^{i} &= |\{j : j \leq i, \text{round } j \text{ was good}\}|, \\ N_{\mathrm{b}}^{i} &= |\{j : j \leq i, \text{round } j \text{ was bad}\}|. \end{aligned}$$

Then $P(i+1) \geq N_{\mathrm{g}}^{i} - 4N_{\mathrm{b}}^{i}$.

Proof. We prove Lemma 5.5.1 by induction. For the base case, $|\psi_1\rangle = |\psi_{\mathrm{init}}\rangle$, so $P(1) = 0$ and the statement holds.

To get a flavor of the induction step, let us look at $P(2)$ at the end of round 1. In round 1, Alice applies U_1 then teleports the virtual C register. If Bob decodes the message correctly, he applies U_2 and teleports the virtual register C back, leading to a joint state of the form $\hat{\sigma}U_2U_1|\psi_{\mathrm{init}}\rangle$. In this case $N_{\mathrm{b}}^{1} = 0$, so $P(2) = 2 \geq 1 = N_{\mathrm{g}}^{1}$. If there is a decoding error, at worst Bob applies the incorrect Pauli operation to complete the teleportation step, and still applies U_2. The joint state is then of the form $\hat{\sigma}U_2\tilde{\sigma}U_1|\psi_{\mathrm{init}}\rangle$. In this case, $N_{\mathrm{g}}^{1} = 0$, and $P(2) = 1 - 2 = -1 \geq -4 = -4N_{\mathrm{b}}^{1}$.

For the induction step, given the state $|\psi_i\rangle$ at the end of round $i-1$, we consider two cases. First, suppose that the ith round is good, so that $N_{\mathrm{g}}^{i} = N_{\mathrm{g}}^{i-1} + 1$ and $N_{\mathrm{b}}^{i} = N_{\mathrm{b}}^{i-1}$. Both Alice and Bob correctly reconstruct the state as in (5.5.2). If $t_i = 0$, by the simulation rules, at least one of Alice or Bob can advance the original noiseless protocol, and $t_{i+1} = t_i = 0$ and $r_{i+1} \geq r_i + 1$. (If r_i is odd, only Bob advances the protocol, otherwise both do.) If $t_i \geq 1$, again, at least one of Alice or Bob can invert the unitary operation $\tilde{U}_{t_i}^{i}$ (depending on the parity of ℓ, where $\tilde{U}_{t_i}^{i} = U_{\ell}^{\pm 1}$). Then $t_{i+1} \leq t_i - 1$, and $r_{i+1} \geq r_i$. So in

all cases

$$\begin{aligned} P(i+1) &= r_{i+1} - 2t_{i+1} \\ &\geq r_i - 2t_i + 1 \\ &= P(i) + 1 \\ &\geq N_{\mathrm{g}}^{i-1} - 4N_{\mathrm{b}}^{i-1} + 1 \\ &= N_{\mathrm{g}}^{i} - 4N_{\mathrm{b}}^{i} \ . \end{aligned}$$

In the second case, the ith round is bad, so that $N_{\mathrm{g}}^{i} = N_{\mathrm{g}}^{i-1}$ and $N_{\mathrm{b}}^{i} = N_{\mathrm{b}}^{i-1} + 1$. At worst, both Alice and Bob decode the received messages incorrectly. With an incorrect guess for the state in (5.5.2), Alice's actions in this round either decrease r_i by one, or increase t_i by one, or leave both unchanged. The same holds for Bob. At worst, $t_{i+1} = t_i + 2$ and $r_{i+1} = r_i$. The other possibilities such as $t_{i+1} = t_i + 1, r_{i+1} = r_i - 1$ or $t_{i+1} = t_i, r_{i+1} = r_i - 2$ lead to a smaller decrease in the potential function P. So

$$\begin{aligned} P(i+1) &= r_{i+1} - 2t_{i+1} \\ &\geq r_i - 2t_i - 4 \\ &= P(i) - 4 \\ &\geq N_{\mathrm{g}}^{i-1} - 4N_{\mathrm{b}}^{i-1} - 4 \\ &= N_{\mathrm{g}}^{i} - 4N_{\mathrm{b}}^{i}. \end{aligned}$$

In all cases, $P(i+1) \geq N_{\mathrm{g}}^{i} - 4N_{\mathrm{b}}^{i}$ which proves the claim. □

Corollary 5.5.1. *If $P(\frac{N'}{2}+1) \geq N+1$, then the simulation succeeds with zero error.*

Proof. For notational convenience, in this proof let $r = r_{\frac{N'}{2}+1}$, $t = t_{\frac{N'}{2}+1}$. We also let the superscript $\frac{N'}{2}+1$ be implicit in all the operators $\tilde{U}_{\ell}^{\pm 1}$ that occur in the proof below.

The only unitary operations from the original protocol that Alice applies are of the form $U_{\ell}^{\pm 1}$, for odd ℓ. Moreover, Alice knows her history at all times. Thus, even in a bad round i, she either applies $U_{\ell+2}$, I or U_{ℓ}^{-1}, where U_{ℓ} is the last unitary operation she applied in the representation (5.5.2). A similar statement holds for Bob. Thus, the

subscripts, in the original protocol, of two consecutive unitary operators applied by the same party in (5.5.2) do not differ by more than 2.

We have $P(\frac{N'}{2}+1) = r-2t \geq N+1$, so $r \geq N+1+2t$ with $t \geq 0$. In particular, we have $r \geq N+1$. Once U_r has been applied, the noiseless protocol embedding ensures that the final state of the noiseless protocol in registers ABC is safely stored in local registers $\tilde{A}\tilde{B}\tilde{C}$ that are never changed by $U_{N+2}\cdots U_{N'+1}$, or by the Pauli operations on the virtual C register. It remains to be verified that that all the operators $\tilde{U}_\ell$, $0 \leq \ell \leq t$, have indices strictly higher than $N+1$.

The indices (in the original protocol) of the operators $\tilde{U}_\ell$ applied by Alice may decrease by at most two at once, and similarly for Bob. So the worst case is if all the operators $\tilde{U}_\ell$ are applied by the same party, and are inverses of the noiseless protocol unitaries. Without loss of generality, we consider only this case. If the party who applied U_r also applies all the operators $\tilde{U}_\ell$, then $\tilde{U}_1 = U_r^{-1}$, $\tilde{U}_2 = U_{r-2}^{-1}$, $\ldots$, $\tilde{U}_t = U_{r-2(t-1)}^{-1}$ and $r-2(t-1) > r-2t = P(\frac{N'}{2}+1) \geq N+1$. So the simulation generates the correct output. Similarly if the party who applied U_{r-1} also applies all the operators $\tilde{U}_\ell$, then $\tilde{U}_1 = U_{r-1}^{-1}$, $\tilde{U}_2 = U_{r-3}^{-1}$, $\ldots$, $\tilde{U}_t = U_{r-2t+1}$ and $r-2t+1 > r-2t = P(\frac{N'}{2}+1) \geq N+1$. In all cases, the safe registers $\tilde{A}\tilde{B}\tilde{C}$ 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 show that if the number of errors as a fraction of N', which is the total number of classical symbols transmitted over the adversarial channel, is bounded by a particular constant $\delta > 0$, we are guaranteed 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 from the more general result in Lemma 5.6.2, which we prove in the next section when studying a protocol designed to tolerate the highest possible error rate. The implication we use here is the following: if the error rate is bounded by δ (so there are at most $\delta N'$ errors) and the tree code distance of both Alice and Bob's tree code is at least α, then the number of bad rounds N_{b} is bounded as $N_{\text{b}} \leq (2\delta + \varepsilon_\alpha)N'$, where $\varepsilon_\alpha = 1-\alpha$.

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

$$\begin{aligned} P\left(\frac{N'}{2}+1\right) &\geq N_{\mathrm{g}} - 4N_{\mathrm{b}} \\ &= \frac{N'}{2} - 5N_{\mathrm{b}} \\ &\geq \frac{N'}{2} - 5(2\delta+\varepsilon_\alpha)N' \\ &= N'\left(\frac{1}{2} - \frac{10}{80} - \frac{5}{40}\right) \\ &= \frac{1}{4}N' \\ &= N+1 \quad . \end{aligned}$$

Here, the first inequality is from Lemma 5.5.1, the first equality is by definition of N_{g}, N_{b}, i.e., $\frac{N'}{2} = N_{\mathrm{g}} + N_{\mathrm{b}}$, and the second inequality is from our bound on N_{b} due to Lemma 5.6.2. The fact that the simulation succeeds is then immediate from Corollary 5.5.1.

Note that the form of 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 and her view of the transcript, as allowed by the model), the simulation succeeds regardless of her choice of action. As long as the corruption rate is bounded by δ, our analysis holds in each branch of the adversary's probabilistic computation . We use the definition of the class $\mathscr{A}^{\mathrm{S}}_{\delta,q,N'}$ to prove that indeed, the simulation succeeds with zero error.

For $|\psi\rangle \in \mathscr{H}(A \otimes B \otimes C \otimes E \otimes D)$, with D a purifying system of the same size as $A \otimes B \otimes C \otimes E$, with have that

$$(\Pi \otimes \mathrm{I}^D)(|\psi\rangle) = \mathrm{Tr}_E(U_N \cdots U_1 |\psi\rangle\langle\psi| U_1^\dagger \cdots U_N^\dagger) \quad ,$$

where Π is the protocol that is being simulated. For any adversary in $\mathcal{A} \in \mathscr{A}^{\mathrm{S}}_{\delta,q,N'}$, the

simulation yields state

$$(S^{\Pi}(\mathcal{A}) \otimes \mathrm{I}^{D})(|\psi\rangle) = \mathrm{Tr}_{\neg(\tilde{A}\tilde{B}\tilde{C}D)}(\mathscr{M}^{\Pi}_{N'+1}\mathscr{N}_{N'}\mathscr{M}^{\Pi}_{N'}\cdots\mathscr{M}^{\Pi}_{2}\mathscr{N}_{1}\mathscr{M}^{\Pi}_{1}(|\psi\rangle\langle\psi|)),$$

in which the $\neg(\tilde{A}\tilde{B}\tilde{C}D)$ subscript for the partial trace means that we trace all except the $\tilde{A}\tilde{B}\tilde{C}D$ registers, and the instrument $\mathscr{M}^{\Pi}_{\ell}$ is the simulation step for the ℓth local computation by the corresponding party. Then we can rewrite

$$\begin{aligned}&(S^{\Pi}(\mathcal{A}) \otimes \mathrm{I}^{D})(|\psi\rangle) \\ &\qquad = \sum_{x_{\mathrm{T}}y_{\mathrm{T}}z} p_{X_{\mathrm{T}}Y_{\mathrm{T}}Z}(x_{\mathrm{T}}, y_{\mathrm{T}}, z|\,|\psi\rangle) \quad |x_{\mathrm{T}}\rangle\langle x_{\mathrm{T}}|^{X_{\mathrm{T}}} \otimes |y_{\mathrm{T}}\rangle\langle y_{\mathrm{T}}|^{Y_{\mathrm{T}}} \otimes |z\rangle\langle z|^{Z} \otimes \rho(x_{\mathrm{T}}, y_{\mathrm{T}}, z)\end{aligned}$$

where X_{T}, Y_{T} are the registers containing the views x_{T}, y_{T} of the transcript as seen by Alice and Bob, respectively, Z is the adversary's classical register, $\rho(x_{\mathrm{T}}, y_{\mathrm{T}}, z)$ are some quantum states, and $p_{X_{\mathrm{T}}Y_{\mathrm{T}}Z}$ a probability distribution conditional on the input $|\psi\rangle$. By definition of the class $\mathscr{A}^{\mathrm{S}}_{\delta,q,N'}$, we have that, conditioned on some classical state z of Eve, $\rho(x_{\mathrm{T}}, y_{\mathrm{T}}, z)$ suffers at most $\delta N'$ corruptions by Eve, for any possible transcript views x_{T}, y_{T}. So, by the above analysis, its $\tilde{A}\tilde{B}\tilde{C}D$ subsystems contains $\mathrm{Tr}_{E}(U_N \cdots U_1 |\psi\rangle\langle\psi| U_1^{\dagger} \cdots U_N^{\dagger})$, a perfect copy of $(\Pi \otimes \mathrm{I}^{D})(|\psi\rangle)$ for any views $x_{\mathrm{T}}, y_{\mathrm{T}}$ of the transcripts of Alice and Bob, respectively. Hence, tracing over all subsystems but $\tilde{A}\tilde{B}\tilde{C}D$, we obtain $(\Pi \otimes \mathrm{I}^{D})(|\psi\rangle)$, and the simulation protocol succeeds with zero probability of error at simulating any noiseless protocol of length N against all adversaries in $\mathscr{A}^{\mathrm{S}}_{\delta,q,N'}$.

We have thus established the following. We use a tree code of arity $d = 48$ and distance parameter $\alpha = 1 - \varepsilon_{\alpha} = \frac{39}{40}$. With $q = |\Sigma|$ chosen according to Lemma 5.2.1, $R_{\mathrm{C}} = \frac{N}{N' \log q} = \frac{1}{4(1+\frac{1}{N})\log q} \geq \frac{1}{8 \log q}$, $R_{\mathrm{E}} = \frac{1}{\log q}$, and $\delta = \frac{1}{80}$, we have that for all N, there exists a universal simulation protocol in the shared entanglement model that, given black-box access to any two-party quantum protocol of length N in the noiseless model, succeeds with zero probability of error at simulating the noiseless protocol on any input (independent of the contents of the purifying register held by Eve) while transmitting $\frac{1}{R_{\mathrm{C}} \log q} N$ symbols from an alphabet Σ of size q over any adversarial channel with error

rate δ, and consuming $\frac{R_E}{R_C}N$ EPR pairs. This proves Theorem 5.5.1.

5.6 Tolerating Maximal Error Rates

We show how we can modify the basic protocol described in the last section such that it tolerates up to $\frac{1}{2} - \varepsilon$ error rate, for arbitrarily small $\varepsilon > 0$, 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 this model. More formally, we prove the following results.

Theorem 5.6.1. *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 worst-case error than the best unidirectional protocol. This result holds in the oblivious as well as the 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 of length N' with the above parameters, there exists an adversary $\mathcal{A} \in \mathscr{A}^S_{\frac{1}{2},q,N'}$ and an unidirectional protocol U such that for all noiseless protocols Π of length N, $\|S^\Pi(\mathcal{A}) - \Pi\|_\diamond \geq \|U - \Pi\|_\diamond$.*

Theorem 5.6.2. *Given an adversarial channel in the shared entanglement model with constant 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 transmissions linear in N, and consuming a linear number of EPR pairs. More precisely, there exists a constant $c > 0$ such that for arbitrarily small constant $\varepsilon > 0$, 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 $\mathscr{A}^S_{\frac{1}{2}-\varepsilon,q,N'}$.*

5.6.1 Proof of Optimality

To prove Th. 5.6.1, we observe that the argument of Ref. [62] 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's. In particular, he could replace all of Bob's transmissions by a fixed message, and leave Alice's messages unchanged. Effectively Bob does not transmit any information to Alice, and this protocol can be simulated in the unidirectional model. Indeed, suppose that for a fixed register E, transmission alphabet Σ of size q, noiseless protocol length N, and simulation protocol length N', the adversary $\mathcal{A}_{\frac{1}{2}}$ 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 Π of length N. We construct $\mathscr{M}_1^U$ which is the composition of all operations of Alice in S while replacing all messages of Bob by e_0. In the unidirectional protocol U, Alice applies the instrument $\mathscr{M}_1^U$ to Alice's share of the joint state in the simulation protocol. The quantum communication from Alice to Bob is the concatenation of all the messages from Alice in the simulation protocol, along with Bob's share of the initial joint state. Bob would then apply the instrument $\mathscr{M}_2^U$, which is the sequential application of all his operations in the simulation protocol S. This unidirectional protocol simulates S running against the adversary $\mathcal{A}_{\frac{1}{2}}$ for any noiseless protocol and any input, and then produces the same output.

The above proof also applies in an oblivious model for noisy communication. 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. An adversary can choose to disrupt all the messages of the party who communicates at most half the number of symbols. Hence, the proof also extends to the case of oblivious, but not necessarily alternating, communication. In such a case, the simulation protocol would also define a function Speak : $[N'] \to \{A, B\}$ known to all (Alice, Bob and Eve) which specifies whose turn it is to speak and is independent of both the input and of the action of Eve.

We can further 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 to decide which party's communication she is going to corrupt. It would be a reasonable assumption if the communication is classical, that Eve can see who speaks *before* she decides whether or not to corrupt a message. In this case, the statement is changed to $\|S^{\Pi}(\mathcal{A}) - \Pi\|_{\diamond}$ 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 5.7.2 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.

5.6.2 Proof of Achievability

We first describe the modification to the simulation of Section 5.5 that are required to tolerate maximum error, and then describe how to adapt the analysis to obtain such optimal result.

5.6.2.1 Description of the Simulation

The proof of achievability is somewhat more involved. It follows ideas similar to that of the basic simulation, but protocol is carefully analysed and optimized. We start by setting up new notation that enables us to do so. The intuition given in section 5.5.2 still applies here, but parameters which were fixed in the basic case now depend on the parameter ε when we wish to tolerate an error-rate of $\frac{1}{2} - \varepsilon$. In particular, the distance parameter $\alpha = 1 - \varepsilon_{\alpha}$ now changes, as well as the length of the protocol $N' = lN$. 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 $\varepsilon_\beta = \frac{|\Sigma|-1}{|\Gamma|-1}$, with Σ the tree code alphabet and Γ the blueberry code alphabet. Each of the tree code transmission alphabet symbols are further encoded with the blueberry code before transmission over the noisy channel. A corruption 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 attempt to continue the simulation (as in the previous section) in that round. The corresponding trit sent is 0, and the teleportation decoding bits are 00. Otherwise, the structure of the protocol is mainly unchanged.

We summarize the optimized protocol below. Alice and Bob start with the state $|\psi_{\text{init}}\rangle$ in the registers $ABC_{\mathsf{A}}E$, the register C_{B} initialized to $|0\rangle$, the registers $T_{\mathsf{A}}T_{\mathsf{B}}$ initialized to N' EPR pairs $\left[\frac{1}{\sqrt{2}}(|00\rangle+|11\rangle)\right]^{\otimes N'}$, with one qubit from each EPR pair held by Alice, one by Bob, and the qubits in registers $\tilde{A},\tilde{B},\tilde{C}$ initialized to $|0\rangle$ (cf. the noiseless protocol embedding described in Section 5.3.1). They measure a suitable number of additional EPR pairs to produce a secret key unknown to the adversary. Using this, generate common blueberry codes $\mathcal{B}_1,\mathcal{B}_2,\ldots,\mathcal{B}_{N'}$ uniformly and independently from the set of permutations over Γ. They also have access to a suitable amount of classical workspace for local computations required for the simulation.

Alice and Bob repeat the following for $i = 1\cdots\frac{N'}{2}$:

1. For $i=1$, there is no message to be decoded, and the values of the parameters needed for the simulation are straightforward. Alice continues with step 3. If $i>1$, Alice decodes the blueberry encoding of Bob's possibly corrupted last transmission. If she detects an erasure, she sets $M_i^{\mathsf{A}} = 0, x_i^{\mathsf{AD}} = z_i^{\mathsf{AD}} = 0$ 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 continues with step 2.

2. Alice computes $s_{\mathsf{B}}^i = \mathcal{D}(f'_1\cdots f'_{i-1})$, and extracts $b_\ell^i = (x_\ell^{i\mathsf{BD}} z_\ell^{i\mathsf{BD}}, M_\ell^{i\mathsf{B}}, x_\ell^{i\mathsf{BM}} z_\ell^{i\mathsf{BM}})$ for $\ell = 1,\ldots,i-1$, her best guess for Bob's messages, and the corresponding $c_\ell^{i\mathsf{B}}, j_\ell^{i\mathsf{B}}$.

3. Using $s_{\mathsf{A}}, s_{\mathsf{B}}$, she computes her best guess for the state $|\psi_i\rangle$ of the joint register,

and the corresponding $x_i^{\mathsf{AD}} z_i^{\mathsf{AD}}, M_i^{\mathsf{A}}, c_i^{\mathsf{A}}, j_i^{\mathsf{A}}$.

4. She completes the teleportation by applying $\mathrm{Z}^{z_i^{\mathsf{AD}}} \mathrm{X}^{x_i^{\mathsf{AD}}}$ to register $T_{\mathsf{A}}^{2(i-1)}$ and swaps this with the C_{A} register.
5. She tries to make progress in the simulation by applying $U_{j_i^{\mathsf{A}}}^{M_i^{\mathsf{A}}}$ to the AC_{A} register.
6. She teleports the C_{A} register to Bob using entanglement in register T_{A}^{2i-1} and gets outcomes $x_i^{\mathsf{AM}} z_i^{\mathsf{AM}}$.
7. Alice updates her history s_{A} by following edge $a_i = (x_i^{\mathsf{AD}} z_i^{\mathsf{AD}}, M_i^{\mathsf{A}}, x_i^{\mathsf{AM}} z_i^{\mathsf{AM}})$, computes $e_i = \mathcal{E}(a_1 \cdots a_i)$ and transmits the blueberry encoding $\mathcal{B}_{2i-1}(e_i)$ of e_i over the noisy the channel to Bob.
8. Upon receiving 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 decodes the transmission as $e_i' \in \Sigma$, a possibly corrupted version of e_i.
9. Bob computes $x_i^{\mathsf{BD}} z_i^{\mathsf{BD}}, M_i^{\mathsf{B}}$ analogously to Alice, depending on whether or not he detects an erasure. In more detail, if Bob does not detect an erasure, he decodes $s_{\mathsf{A}}^i = \mathcal{D}(e_1' \cdots e_i')$ and also uses s_{B} to compute the above parameters. He then performs actions on his registers analogous to Alice's: he completes the teleportation step, swaps register T_{B}^{2i-1} with C_{B}, applies the operator $U_{j_i^{\mathsf{B}}}^{M_i^{\mathsf{B}}}$ to the registers BC_{B}, uses the T_{B}^{2i} register to teleport back the C_{B} register to Alice, computes f_i, and transmits the blueberry encoding $\mathcal{B}_{2i}(f_i)$ of f_i to Alice. Round i is completed when Alice receives a possibly corrupted version of this message.

After these $\frac{N'}{2}$ rounds, they both extract the output of the simulation from the $\tilde{A}\tilde{B}\tilde{C}$ registers specified by the noiseless protocol embedding.

5.6.2.2 Analysis

As in the proof in Section 5.5.4, the analysis is first carried conditional on some respective views of Alice, Bob and Eve of the transcript at each round. An additional component is the conditioning on some classical state z of the Z register of the adversary, Eve, and the averaging over the shared secret key used for the blueberry code. In

particular, if the adversary has an adaptive and probabilistic strategy, we condition on some strategy consistent with the transcript on which we have already conditioned. We return to this issue later.

We again define a function $P(i)$ such that the simulation succeeds whenever $P(\frac{N'}{2}+1) \geq N+1$. Using the notation and the form of the state $|\psi_i\rangle$ on the joint register $ABCE$ at the beginning of round i (or at the end of round $i-1$) rewritten as in (5.5.2), we let $P(i) = r_i - 2t_i$ (i.e., the same potential function works for the enhanced simulation as well). We now have three kinds of rounds: *good* rounds in which both parties decode correctly the other party's history, *bad* rounds in which at least one party makes a decoding error, and the *erasure* rounds, in which no party makes a decoding error, but at least one party decodes an erasure from the blueberry code.

We state an analogue of the technical Lemma 5.5.1 and its corollary.

Lemma 5.6.1. *At the end of round i, define*

$$\begin{aligned} N_{\mathrm{g}}^i &= |\{j : j \leq i, \text{round } j \text{ was good}\}|, \\ N_{\mathrm{b}}^i &= |\{j : j \leq i, \text{round } j \text{ was bad}\}|, \\ N_{\mathrm{e}}^i &= |\{j : j \leq i, \text{round } j \text{ was an erasure round}\}|. \end{aligned}$$

Then $P(i+1) \geq N_{\mathrm{g}}^i - 4N_{\mathrm{b}}^i$.

The proof of this lemma and its corollary below are omitted since they are nearly identical to the proofs in the basic simulation. The only difference is if in some round, which may be a bad round or an erasure round, at least one party detects an erasure. We sketch the argument in the case that round i is an erasure round. The only unitary operation applied by a party that detects an erasure, is a Pauli operator on the virtual communication register C. If both parties detect an erasure, $r_{i+1} = r_i$ and $t_{i+1} = t_i$. If any one party decodes correctly and the other detects an erasure, we have $r_{i+1} \geq r_i$ and $t_{i+1} \leq t_i$, so $P(i+1) \geq P(i)$. (The function only increases if the party that decoded correctly can apply U_{r_i+1} or $\tilde{U}_{t_i}^{-1}$ as defined by the simulation, i.e., that party holds the registers on which the said unitary operation acts.) In both cases, the quantity $N_{\mathrm{g}}^i - 4N_{\mathrm{b}}^i =$

$N_{\mathrm{g}}^{i-1} - 4N_{\mathrm{b}}^{i-1} \leq P(i)$, so $P(i+1) \geq N_{\mathrm{g}}^{i} - 4N_{\mathrm{b}}^{i}$.

Corollary 5.6.1. *If $P(\frac{N'}{2}+1) \geq N+1$, then the simulation succeeds with zero error.*

Hence, it suffices to bound the ratio of bad to good rounds as a function of the corruption rate in order 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 make 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, there is now a third kind of round, and we would also want to ensure that the ratio of good rounds versus erasure rounds does not get arbitrarily low, and that $P(\frac{N'}{2}+1) \geq N+1$.

We focus on the number $N_{\mathrm{g}} = N_{\mathrm{g}}^{\frac{N'}{2}+1}$, $N_{\mathrm{b}} = N_{\mathrm{b}}^{\frac{N'}{2}+1}$ and $N_{\mathrm{e}} = N_{\mathrm{e}}^{\frac{N'}{2}+1}$ 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 appeal to a corollary of the following technical lemma. The lemma derives a new bound on tree codes with an erasure symbol. Since this result only pertains to the structure of such codes independent of our application, it might have applications to classical interactive coding and other settings as well.

Lemma 5.6.2. *If there is a bound δ on the fraction of the total number of transmissions 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 as $N_{\mathrm{b}} \leq (2\delta + \varepsilon_\alpha)N'$, where $\varepsilon_\alpha = 1 - \alpha$, and α is the distance parameter of the tree code with an erasure symbol used by Alice and Bob.*

Proof. For any $1 \leq i \leq j \leq \frac{N'}{2}$, let $I_{\mathrm{e}}^{\mathsf{A}}(i,j), I_{\mathrm{b}}^{\mathsf{A}}(i,j), I_{\mathrm{g}}^{\mathsf{A}}(i,j)$ be the subset of rounds $i, i+1, \cdots, j-1, j$ in which the symbol Alice gets from the blueberry decoding is an erasure, an error (i.e., an incorrect symbol), or the original encoded symbol, respectively. Note that these are disjoint sets satisfying $I_{\mathrm{e}}^{\mathsf{A}}(i,j) \cup I_{\mathrm{b}}^{\mathsf{A}}(i,j) \cup I_{\mathrm{g}}^{\mathsf{A}}(i,j) = [i,j]$, where $[i,j]$ denotes the set $\{i, i+1, \cdots, j-1, j\}$. Similarly, let $J_{\mathrm{b}}^{\mathsf{A}}(i,j)$ and $J_{\mathrm{g}}^{\mathsf{A}}(i,j)$ be the subsets of $[i,j]$ in which the sequence of messages Alice gets from the tree decoding corresponds to a decoding error and the correct decoding, respectively. Again note that

$I_{\mathrm{e}}^{\mathsf{A}}(i,j) \cup J_{\mathrm{b}}^{\mathsf{A}}(i,j) \cup J_{\mathrm{g}}^{\mathsf{A}}(i,j) = [i,j]$, a disjoint union. We define analogous subsets for Bob with A's replaced by B's in the notation. Using this notation, we have

$$\begin{aligned} N_{\mathrm{b}} &= \left| J_{\mathrm{b}}^{\mathsf{A}}\left(1, \frac{N'}{2}\right) \cup J_{\mathrm{b}}^{\mathsf{B}}\left(1, \frac{N'}{2}\right) \right|, \quad \text{and} \\ \left| I_{\mathrm{b}}^{\mathsf{A}}\left(1, \frac{N'}{2}\right) \right| + \left| I_{\mathrm{b}}^{\mathsf{B}}\left(1, \frac{N'}{2}\right) \right| &\leq \delta N' \ . \end{aligned}$$

The statement we wish to prove is

$$\left| J_{\mathrm{b}}^{\mathsf{A}}\left(1, \frac{N'}{2}\right) \cup J_{\mathrm{b}}^{\mathsf{B}}\left(1, \frac{N'}{2}\right) \right| \leq 2\delta N' + \varepsilon_\alpha N'.$$

We prove the following stronger statements, which claim that the number of rounds in which a party makes a tree code decoding error is only slightly larger than the number of rounds in which she makes a blueberry code decoding error:

$$\left| J_{\mathrm{b}}^{\mathsf{A}}\left(1, \frac{N'}{2}\right) \right| \leq 2 \left| I_{\mathrm{b}}^{\mathsf{A}}\left(1, \frac{N'}{2}\right) \right| + \frac{1}{2}\varepsilon_\alpha N' \ , \tag{5.6.1}$$

and

$$\left| J_{\mathrm{b}}^{\mathsf{B}}\left(1, \frac{N'}{2}\right) \right| \leq 2 \left| I_{\mathrm{b}}^{\mathsf{B}}\left(1, \frac{N'}{2}\right) \right| + \frac{1}{2}\varepsilon_\alpha N' \ .$$

The proofs of the two statements are similar, so we only prove the statement for Alice's subsets. To lighten the notation, we drop the A superscripts. For any subset K of $[\frac{N'}{2}]$ and any two strings $\bar{e}, \bar{e}' \in \Sigma^t$ with $\bar{e} = e_1 \cdots e_t$ and $\bar{e}' = e'_1 \cdots e'_t$, and $t \leq N'/2$, define $\Delta_K(\bar{e}, \bar{e}') = |\{i \in K : i \leq t, e_i \neq e'_i\}|$. Note that with $\bar{K} = [\frac{N'}{2}] \setminus K$, $\Delta(\bar{e}, \bar{e}') = \Delta_K(\bar{e}, \bar{e}') + \Delta_{\bar{K}}(\bar{e}, \bar{e}')$, and $\Delta_K(\bar{e}, \bar{e}') \leq |K|$.

We are now ready to prove the statement (5.6.1). We prove by strong induction on the number of rounds t that $|J_{\mathrm{b}}(1,t)| \leq 2|I_{\mathrm{b}}(1,t)| + \varepsilon_\alpha t$. The base case, $t = 1$, is immediate: in the first round, Alice does not decode any message, so that the two sets $J_{\mathrm{b}}(1,1), I_{\mathrm{b}}(1,1)$ are empty.

For $t > 1$, assume that

$$|J_{\mathrm{b}}(1,j)| \quad \leq \quad 2|I_{\mathrm{b}}(1,j)| + \varepsilon_\alpha j \ ,$$

for all j with $0 \leq j < t$, where we define $J_{\mathrm{b}}(1,0) = I_{\mathrm{b}}(1,0) = \emptyset$. If in round t, $t > 1$, Alice detects an erasure or decodes correctly, then the induction step is immediate. Hence, for the induction step, we consider the case of incorrect decoding. Let $\bar{a} \in [d]^t$ be the sequence of transmitted messages, $\bar{e} = \bar{\mathcal{E}}(\bar{a}) \in \Sigma^t$ the corresponding sequence of transmissions, $\bar{e}' \in \Sigma^t$ the sequence of possibly corrupted receptions, $\bar{a}' = \mathcal{D}(\bar{e}') \in [d]^t$ the sequence of decoded messages, and $\bar{e}'' = \bar{\mathcal{E}}(\bar{a}')$ its encoding in the tree code. 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 $\Delta_{[1,t-\ell]}(\bar{e}'', \bar{e}) = 0$, as the encodings have the same prefix as well. Since $\bar{e}'' \neq \bar{e}$, note that $1 \leq \ell \leq t$. By the induction hypothesis,

$$|J_{\mathrm{b}}(1,t-\ell)| \quad \leq \quad 2|I_{\mathrm{b}}(1,t-\ell)| + \varepsilon_\alpha (t-\ell) \ .$$

By definition

$$\begin{aligned} |J_{\mathrm{b}}(1,t)| &= |J_{\mathrm{b}}(1,t-\ell)| + |J_{\mathrm{b}}(t-\ell+1,t)|, \\ |I_{\mathrm{b}}(1,t)| &= |I_{\mathrm{b}}(1,t-\ell)| + |I_{\mathrm{b}}(t-\ell+1,t)|, \end{aligned}$$

so it suffices to prove

$$|J_{\mathrm{b}}(t-\ell+1,t)| \quad \leq \quad 2|I_{\mathrm{b}}(t-\ell+1,t)| + \varepsilon_\alpha \ell \tag{5.6.2}$$

to complete the proof.

Let $K = I_{\mathrm{e}}(t-\ell+1,t)$, the set of rounds in which Alice detects an erasure. Since codewords in the tree code, in particular $\bar{e}''$ and $\bar{e}$, do not contain the erasure symbol, the decoding condition $\Delta(\bar{e}'', \bar{e}') \leq \Delta(\bar{e}', \bar{e})$ is equivalent to $\Delta_{\bar{K}}(\bar{e}'', \bar{e}') \leq \Delta_{\bar{K}}(\bar{e}', \bar{e})$. We

therefore have

$$\begin{aligned}\Delta(\bar{e}'',\bar{e}) &= \Delta_K(\bar{e}'',\bar{e})+\Delta_{\bar{K}}(\bar{e}'',\bar{e}) \\ &\leq |I_{\mathrm{e}}(t-\ell+1,t)|+\Delta_{\bar{K}}(\bar{e}'',\bar{e}) \\ &\leq |I_{\mathrm{e}}(t-\ell+1,t)|+\Delta_{\bar{K}}(\bar{e}'',\bar{e}')\Delta_{\bar{K}}(\bar{e}',\bar{e}) \\ &\leq |I_{\mathrm{e}}(t-\ell+1,t)|+2\Delta_{\bar{K}}(\bar{e}',\bar{e}) \\ &= |I_{\mathrm{e}}(t-\ell+1,t)|++2\,|I_{\mathrm{b}}(t-\ell+1,t)| \ . \end{aligned} \tag{5.6.3}$$

On the other hand, since $\bar{a}\neq\bar{a}'$, the tree code distance condition stipulates that $\Delta(\bar{e}'',\bar{e})\geq \alpha\ell=(1-\varepsilon_\alpha)\ell$. Along with (5.6.3), this gives us

$$\ell \ \leq \ \Delta(\bar{e}'',\bar{e})+\varepsilon_\alpha\ell \ \leq \ |I_{\mathrm{e}}(t-\ell+1,t)|+2\,|I_{\mathrm{b}}(t-\ell+1,t)|+\varepsilon_\alpha\ell \ . \tag{5.6.4}$$

We use this to bound the number of bad rounds for Alice, in terms of the number of blueberry decoding errors she encounters. We have

$$\begin{aligned}\ell &= |I_{\mathrm{e}}(t-\ell+1,t)|+|J_{\mathrm{b}}(t-\ell+1,t)|+\left|J_{\mathrm{g}}(t-\ell+1,t)\right| \\ &\geq |I_{\mathrm{e}}(t-\ell+1,t)|+|J_{\mathrm{b}}(t-\ell+1,t)| \ . \end{aligned} \tag{5.6.5}$$

Combining (5.6.4) and (5.6.5), we get the claimed bound, as in (5.6.2). □

Corollary 5.6.2. *If the corruption rate c of the channel satisfies $0\leq c<\frac{1}{2}$, then except with probability smaller than $2^{-\Omega(N')}$, where N' is the length of the simulation protocol, the total number of bad rounds in the simulation is bounded as $N_{\mathrm{b}}\leq(2\varepsilon_\beta+\varepsilon_\alpha)N'$, where $\varepsilon_\alpha=1-\alpha$, α is the distance parameter of the tree code, $\varepsilon_\beta=1-\beta$, and β is the erasure parameter of the blueberry code.*

Proof. Suppose the transmitted symbol is $g_i\in\Gamma$ after a blueberry encoding $\mathcal{B}_j$ (where $j\in\{2i-1,2i\}$), and that conditional on her classical state and some measurement outcomes z_k until round i, Eve chooses to corrupt g_i into a different $g_i'\in\Gamma$. This action is independent from the randomness used in B_j, and 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. We

consider two cases. First, suppose the corruption rate c is bounded as $\varepsilon_\beta \leq c < \frac{1}{2}$ (so that the corruption rate is at least a constant). By Lemma 5.2.2, with probability $1 - 2^{-\Omega(N')}$ at least a $(1-2\varepsilon_\beta)$-fraction of the cN' corrupted transmissions are detected as erasures. So the blueberry decoding gives at most $cN' - c(1-2\varepsilon_\beta)N' = 2c\varepsilon_\beta N' < \varepsilon_\beta N'$ transmission errors, except with probability negligible in N'. Taking $\delta = \varepsilon_\beta$ in the statement of Lemma 5.6.2 gives us the corollary. If $0 \leq c \leq \varepsilon_\beta$, then the corollary is immediate from Lemma 5.6.2, with $\delta = \varepsilon_\beta$. □

With the above result in hand, we can show that if the corruption rate is $\frac{1}{2} - \varepsilon$ with $\varepsilon > 0$, and we take $\varepsilon_\alpha = \frac{1}{20}\varepsilon, \varepsilon_\beta = \frac{1}{40}\varepsilon, N' \geq \frac{2}{\varepsilon}(N+1)$, then except with negligible probability, the simulation succeeds:

$$
\begin{aligned}
P\left(\frac{N'}{2}+1\right) \quad & \geq N_{\mathrm{g}} - 4N_{\mathrm{b}} \\
& = \frac{N'}{2} - N_{\mathrm{e}} - 5N_{\mathrm{b}} && \text{(By Lemma 5.6.1)} \\
& \geq \varepsilon N' - 5N_{\mathrm{b}} && (\text{since } N'/2 = N_{\mathrm{g}} + N_{\mathrm{b}} + N_{\mathrm{e}}) \\
& \geq \varepsilon N' - 5(2\varepsilon_\beta + \varepsilon_\alpha)N' && (\text{since } N_{\mathrm{e}} \leq (1/2 - \varepsilon)N') \\
& = N'\left(\varepsilon - \frac{10}{40}\varepsilon - \frac{5}{20}\varepsilon\right) && \text{(By Corollary 5.6.2)} \\
& = \frac{1}{2}\varepsilon N' \\
& \geq N+1 \quad .
\end{aligned}
$$

That the simulation succeeds is now immediate from Corollary 5.6.1.

The above statement holds conditional 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 5.6.2, we argue as in Section 5.5.4 to translate these results to the output state produced by the protocols, even when we consider inputs entangled with some reference register R. We do not repeat the whole analysis here, since it is nearly identical once we make the following observation. An arbitrary channel Eve fitting the framework of the shared entanglement model could have adaptive, probabilistic behaviour based on

previous measurement outcomes. However, these probabilistic choices are independent of the secret key generated by Alice and Bob for the blueberry code. As in section 5.5.4, the above result holds for each probabilistic choice of Eve. Summing over all such choices, we obtain the same result, proving Theorem 5.6.2.

5.7 Results in Other Models

By adapting the results in the shared entanglement model for an adversarial error model, we can obtain several 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 prove that the standard forward quantum capacity of the quantum channels used does not characterize their communication capacity in the interactive communication scenario. Finally, we consider a variation on the shared entanglement model in which, along with the noisy classical communication, the shared entanglement is also noisy.

5.7.1 Shared Entanglement Model with Random Errors

In this section we consider two-party protocols with prior shared entanglement and classical communication over binary symmetric channels. Given a two-party quantum protocol of length N in the noiseless model and any $C > 0$, we exhibit a simulation protocol in the shared entanglement model that is of length $\mathrm{O}(\frac{1}{C}N)$ and succeeds in simulating the original protocol with negligible error over classical binary symmetric channels of capacity C. More precisely,

Theorem 5.7.1. *There exist constants $c, l > 0$ such that given any classical binary symmetric channel $\mathscr{M}$ of capacity $C > 0$ and noiseless protocol length $N \in 2\mathbb{N}$, there exist a universal simulator S in the shared entanglement model of length N' with communication rate $R_{\mathrm{C}} \geq lC$, transmission alphabet of size 2, entanglement consumption rate $R_{\mathrm{E}} \leq 1$, which succeeds with error 2^{-cN} at simulating all noiseless protocols of length N over $\mathscr{M}$.*

We complement this with a lower bound for the communication rate. We exhibit 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 protocols of length $\mathrm{o}(\frac{1}{C}N)$ in the shared entanglement model with classical binary symmetric channels of capacity C fail at producing the final state with low error on some input. Moreover, the family of quantum protocol can be chosen to be one that computes a distributed binary function. More precisely,

Theorem 5.7.2. *There exists a sequence $\{\Pi_N\}_{N\in 2\mathbb{N}}$ of two-party quantum protocols such that for all $C > 0$, for any simulation protocol S in the shared entanglement model of length $N' \in \mathrm{o}(N/C)$ with communication rate $R_{\mathrm{C}} = \frac{N}{N'}$ and arbitrary entanglement consumption rate R_{E}, the simulation makes error at least $1 - \mathrm{o}(1)$ over binary symmetric channel of capacity C.*

5.7.1.1 Discussion of Optimality

The above results show that, in the regime where we use binary symmetric channels of classical capacity close to 0, we cannot do much better than what we achieve, up to a multiplicative constant on top of the $\frac{1}{C}$ 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 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 we 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 [90] 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.

5.7.1.2 Proof of Theorem 5.7.1

In Lemma 2 of Ref. [114], it is stated that, given a transmission alphabet Σ, there exists $d > 0$ and $\varepsilon \in (0, \frac{1}{90})$ such that given a binary symmetric channel $\mathscr{M}$ of capacity C, there is a $p \in \mathbb{N}, p \leq d\frac{1}{C}$, an encoding function $\mathcal{E} : \Sigma \to \{0,1\}^p$ and a decoding function $\mathcal{D} : \{0,1\}^p \to \Sigma$ such that $\Pr[\mathcal{D}(\mathscr{M}(\mathcal{E}(e))) \neq e] \leq \varepsilon$ for all e. We use this in conjunction with the result of Theorem 5.5.1 and the Chernoff bound to obtain the following result: with $\varepsilon < \frac{1}{80}$, Σ given by Lemma 5.2.1 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 $p \in \mathbb{N}, \mathcal{E}$ and $\mathcal{D}$, if all the Σ transmissions in the basic simulation protocol are done by reencoding over $\{0,1\}^p$ with $\mathcal{E}$ (and decoding with $\mathcal{D}$), then except with probability $2^{-\Omega(N'')}$ for $N'' = 4(1+\frac{1}{N})N$ the length of the basic simulation protocol over alphabet Σ, $N' = pN''$ 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 Σ symbols is below $\frac{1}{80}$. By Theorem 5.5.1 the simulation succeeds.

5.7.1.3 Proof of Theorem 5.7.2

It is known that for a classical discrete memoryless channel such as the binary symmetric channel, entanglement-assistance does not increase the classical capacity [19], and it is also known that allowing for classical feedback also does not lead to an increase in 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 [25]. 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. For some protocols of length N fitting our general framework in the noiseless model, like those accomplishing a quantum swap function or even a clas-

sical swap or bitwise XOR functions on inputs of size $\frac{N}{2}$, the parties effectively exchange their entire inputs to produce the correct output. Hence, a dilation factor proportional to the inverse of the capacity $\frac{1}{C}$ is necessary. What we wish to prove is even stronger: there exists a family of distributed binary functions such that this is necessary. We consider the inner product function $\mathrm{IP}_n : \{0,1\}^n \times \{0,1\}^n \to \{0,1\}$, defined as $\mathrm{IP}_n(x,y) = \oplus_{i=1}^n x_i \wedge y_i$, which has communication complexity in $\Theta(n)$ in both the Yao and the Cleve-Buhrman quantum communication complexity model [48, 103].

By a reduction due to Cleve, van Dam, Nielsen, and Tapp [48], any protocol evaluating the IP_n function with small error can be used to transmit n classical bits with small probability of error. Hence, any noise-tolerant simulation of such a protocol over a channel of classical capacity C can be used to transmit n-bit strings with some small probability of failure. As a consequence, for small enough error, the simulation requires at least $\frac{1}{C}n$ uses of the channel. 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 protocols compute a function in a coherent way is natural if we wish to compose quantum simulation protocols; then they may be run on arbitrary superpositions of inputs.

5.7.2 Quantum Model with Adversarial Errors

We turn our attention to two-party protocols where there is no prior entanglement, and the communication is over noisy *quantum* channels. Given an adversarial channel in the quantum model with error rate strictly smaller than $\frac{1}{6}$, we can simulate any noiseless protocol of length N over this channel using a number of transmissions linear in N. More precisely, we show the following.

Theorem 5.7.3. *There exists a constant $c > 0$ such that for arbitrary small $\varepsilon > 0$, there exist a communication rate $R_{\mathrm{C}} > 0$ 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* $\mathscr{A}^{\mathrm{Q}}_{\frac{1}{6}-\varepsilon,q,N'}$.

5.7.2.1 Proof of Theorem 5.7.3

The approach we take in the quantum model is to emulate the simulation in the shared entanglement model. First, we use the quantum channels available to distribute sufficient entanglement, and then use them effectively as classical channels along with the entanglement to run the simulation protocol from section 5.6. Thus the simulation consists of an entanglement distribution phase, followed by a protocol implementation phase.

In more detail, suppose we wish to emulate a simulation protocol of length N' in the shared entanglement model. Alice uses lN' transmissions, for a parameter l to be specified below, to distribute sufficient perfect entanglement to Bob through the use of a quantum error correcting code (QECC). They then run the simulation protocol in section 5.6. During this protocol implementation phase, before transmission and after reception of a quantum register through the channel, both the sender and the receiver measure the register. These measurements have the effect of transforming all possible quantum actions of Eve into classical actions. 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 information over a classical channel. If the size q of the communication register is larger than the alphabet size Γ of the transmissions, and Eve maps some of these classical messages outside of Γ, Alice and Bob mark these as erasures. So Eve does not gain anything by introducing errors outside Γ.

We start by pinning down the parameters of the quantum error correcting codes (QECCs) needed to distribute the necessary amount of entanglement. In the interest of simplicity, we do not attempt to optimize the parameters involved.

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 of the simulation in Section 5.6. Two maximally entangled states of size $2s$, i.e., states of the form $\sum_{j=0}^{2s-1} |j\rangle^{T_{\mathsf{A}}} |j\rangle^{T_{\mathsf{B}}}$, are used to generate the

secret keys and to create the EPR pairs required for teleportation in every round. For a given size q for the communication register, and simulation protocol in the shared entanglement model of length N', we distribute a maximally entangled state over $N' \log_q(2s)$ registers of size q.

In the entanglement distribution phase of the simulation in the quantum model, we encode the $N' \log_q(2s)$ registers into lN' registers of size q. For the encoding, we use a quantum error correcting code with alphabet size q, transmission rate $R_Q \geq \frac{1}{l} \log_q(2s)$, and maximum tolerable error rate δ to be determined shortly. We only consider exact quantum error correcting codes, but the analysis extends to approximate ones. (Approximate error correction allows for some deviation from perfect transmission.)

To determine the relationship between q, l, and δ required for the simulation to succeed, we first note that in the protocol implementation phase (the second phase of the simulation), we transmit classical messages chosen from a set of size $|\Gamma|$ over the quantum channel. For simplicity, we choose $q \geq |\Gamma|$. To ensure that this second phase succeeds, the number of corruptions in it should be bounded by $(\frac{1}{2} - \varepsilon)N'$. An adversary could choose to put all of the allowed corruptions in the first (entanglement distribution) phase, so the QECC should be able to recover from the same number of errors. In other words, we require $\delta l N' \geq \frac{N'}{2} - \varepsilon N'$. The length of the message in the entanglement distribution phase satisfies $l \geq \frac{1-2\varepsilon}{2\delta}$. In summary, the entire simulation tolerates $\frac{N'}{2} - \varepsilon N'$ adversarial errors during a total of $(l+1)N'$ transmissions of size q registers provided a suitable QECC exists. The error rate tolerated is $\frac{1-2\varepsilon}{2(l+1)}$.

The above analysis applies to the oblivious communication model. If we restrict ourselves to the alternating communication model, we have twice as much communication, i.e., $2lN'$ size-q registers, in the entanglement transmission phase. The adversary can choose to corrupt the transmissions of one party alone, so $l \geq \frac{1-2\varepsilon}{2\delta}$ as before. The total number of transmissions is however $(2l+1)N'$, so the error rate tolerated is $\frac{1-2\varepsilon}{2(2l+1)}$.

We now appeal to a high-dimensional quantum Gilbert-Varshamov bound [8, 61] stating that for arbitrarily small $\varepsilon' > 0$, there exist 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 allow for

perfect decoding of the quantum state. Using these codes with $\varepsilon' = \varepsilon$, we get $\delta = \frac{1}{4} - \varepsilon$, $l \geq \frac{1-2\varepsilon}{2\delta} = \frac{2(1-2\varepsilon)}{1-4\varepsilon}$ and net error rate $\frac{1-2\varepsilon}{2(l+1)} = \frac{(1-2\varepsilon)(1-4\varepsilon)}{6-16\varepsilon} \geq \frac{1}{6} - \varepsilon$ that the simulation protocol can tolerate in an oblivious model of communication. In an alternating model of communication, we are able to tolerate an error rate of $\frac{1}{10} - \varepsilon$.

The above choice of parameters ensures that the error rate in the entanglement distribution phase is bounded by $\frac{1}{4} - \varepsilon$, and the received quantum state can be decoded perfectly. This establishes a shared maximally entangled state of the required dimension. Moreover, the corruption rate of the adversary during the protocol implementation phase is lower than $\frac{1}{2} - \varepsilon$. Recall that Alice and Bob measure the states received over the quantum channel in the standard basis to convert it to a classical channel. Given any strategy of the adversary, which is necessarily independent of the secret key used for the blueberry codes, for any choice of measurement outcomes for Alice and Bob, the simulation succeeds with probability exponentially close to 1 (in terms of N'). The remainder of the analysis goes as in section 5.6.2.2, proving Theorem 5.7.3.

5.7.2.2 Discussion of 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. With the approach of first distributing entanglement and then using the $\frac{1}{2} - \varepsilon$ error rate simulation protocol in the shared entanglement model, we get an overall tolerable error rate for the simulation of less than $\frac{1}{6}$. Crépeau, Gottesman and Smith [50] showed how we can tolerate up to $\frac{1}{2}$ error rate asymptotically for data transmission if we consider approximate quantum error correcting codes. Using these we could get $\frac{1}{4} - \varepsilon$ 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 in the original protocol. This would lead to a communication rate of 0 asymptotically in the simulation. 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 would like the tolerable error rate to be. Using this kind of a

code, if we break up the simulation in two phases, an entanglement distribution part and then a protocol implementation part, the above is the best we can do. We might hope to develop a fully quantum analogue of tree codes that does not entail the two phase simulation, in order to achieve higher error rates. The putative quantum codes would require some properties for fault-tolerant computation, so that we may coherently apply the noiseless protocol unitary operations in the simulation. This issue does not occur in the fully classical setting, since we can copy classical information and perform the computation on the copy.

Finally, we note that the proof of Theorem 5.6.1 applies here as well. It establishes a bound of $\frac{1}{2}$ on the maximum error rate tolerable in an oblivious communication model: no simulation protocol in the quantum model can succeed with arbitrarily small error against all adversaries in $\mathscr{A}^{\mathrm{Q}}_{\frac{1}{2},q,N'}$, for any $q, N' \in \mathbb{N}$.

5.7.3 Quantum Model with Random Errors

We shift our focus to quantum communication over depolarizing channels. Given a two-party quantum protocol of length N in the noiseless model and any $Q > 0$, we devise a simulation protocol in the quantum model that is of length $\mathrm{O}(\frac{1}{Q}N)$ and succeeds in simulating the original protocol with arbitrarily small error over quantum depolarizing channels of quantum capacity Q. More precisely,

Theorem 5.7.4. *There exist a constant $l > 0$ and a function $f : \mathbb{N} \to \mathbb{R}^+$ with $\lim_{N\to\infty} f(N) = 0$ such that given any depolarizing channel $\mathscr{M}$ of quantum capacity $Q > 0$ and noiseless protocol length $N \in 2\mathbb{N}$, there exists a universal simulator P in the quantum model of length N' with communication rate $R_{\mathrm{Q}} \geq lQ$, transmission alphabet size 2, which succeeds with error $f(N)$ at simulating all noiseless protocols of length N over $\mathscr{M}$.*

We point out that quantum capacity with feedback is a lower bound on the dilation needed to simulate protocols over depolarizing channels. There exist a sequence of two-party quantum protocols of increasing length N in the noiseless model such that for all $Q_{\mathrm{B}} > 0$, any corresponding sequence of simulation protocols of length $\mathrm{o}(\frac{1}{Q_{\mathrm{B}}}N)$ in the

quantum model with quantum depolarizing channels of quantum capacity Q_{B} with classical feedback fail at producing the final state with low error on some input. Moreover, the family of quantum protocol can be chosen to be one computing a distributed binary function.

Theorem 5.7.5. *There exists a sequence $\{\Pi_N\}_{N\in 2\mathbb{N}}$ of two-party quantum protocols such that for all $Q_B > 0$, for any simulation protocol S in the shared entanglement model of length $N' \in \mathrm{o}(N/Q_B)$ with communication rate $R_{\mathrm{C}} = \frac{N}{N'}$, the simulation makes error at least $\Omega(1)$.*

It turns out that quantum capacity does not capture the ability to transmit information in an interactive setting. Given a two-party quantum protocol of length N in the noiseless model, there exist 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.

Theorem 5.7.6. *There exist constants $c, R_{\mathrm{Q}} > 0$ such that given a particular depolarizing quantum channel $\mathscr{M}_0$ 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 $\mathscr{M}_0$.*

5.7.3.1 Proof of Theorem 5.7.4

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

It suffices to adapt the result from section 5.5 for a basic simulation protocol of length N'' over some large alphabet Σ. We then only need to distribute N'' maximally entangled states of the appropriate size. For any depolarizing channel of quantum capacity $Q > 0$, we use standard coding results from quantum Shannon theory [133] to distribute

entanglement at a rate of $\frac{d}{Q}$ for some $d > 0$ with low error. Then, for the protocol implementation phase, we appeal to two properties. First, the classical capacity C of a quantum channel is at least as large as its quantum capacity. Second, a classical capacity achieving strategy for the depolarizing channel is to simulate a binary symmetric channel (BSC) of capacity C for each transmission by measuring the output in the computational basis, and then to block code over the corresponding BSC (see, e.g., Ref. [133] for details). We can then translate the proof of Theorem 5.7.1 to design our classical strategy. This succeeds with overwhelming probability assuming perfect entanglement, and the output is arbitrarily close to the noiseless protocol one. Combining the bound on the error from the two phases, 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.

5.7.3.2 Proof of Theorem 5.7.5

The idea for this proof is to use the fact that distributing an EPR pair over a quantum depolarizing channel produces a Werner state, which is symmetric in interchange of Alice and Bob (see section 5.7.4 for a definition of Werner states). Moreover, if Bob uses the free classical feedback to teleport to Alice with these Werner states, this creates a virtual depolarizing channel from him to Alice, with the same parameter as the actual channel from Alice to him. Hence, a quantum depolarizing channel from Alice to Bob along with free classical feedback is sufficient to simulate depolarizing channels in both directions, and the total number of uses of the depolarizing channel is the same in both cases.

Similar to what was argued in the proof of Theorem 5.7.2 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. We apply a coherent version of the idea to use the inner product protocol to communicate, as in the proof of Theorem 5.7.2. This allows us to use the depolarizing channel to distribute quantum entanglement, and then also to teleport (again with the inner product protocol used this time to communicate classical information). For this, it is sufficient to note that what we achieved in the proof of Theorem 5.7.2 using the protocol for IP_n is actualy stronger than $\Theta(N)$ bits of classical communication: we had a coherent bit channel [67] for $\Theta(N)$ cobits (coherent bits), which can be used to distribute $\Theta(N)$ ebits (EPR pairs). Note that we once again make 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.

5.7.3.3 Proof of Theorem 5.7.6

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 property 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 Σ, Alice distributes N'' EPR pairs to Bob by sending $\frac{4N''}{(1-p)}$ halves of such states over the quantum erasure channel. Then, except with negligible probability, at least N'' of them are received intact, and Bob knows which these are. The feedback consists of informing Alice which N'' pairs were received intact and can be used in the protocol. This can be done over the quantum erasure channel, with probability negligibly smaller than 1, with a classical message of length linear in N''.

Then, given a message set Σ 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 $\frac{1}{90}$. Except with negligible probability, the fraction of N'' transmissions of symbols of Σ transmitted in this way is below $\frac{1}{80}$. We can then use a reasoning similar

to that in the proof of Theorem 5.7.3 to argue that the output is arbitrarily close to the noiseless protocol one.

Now for the depolarizing channel, the reasoning 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 Ref. [17] 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 Ref. [16] along with their hashing method. The recurrence method 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. The hashing method is a one-way protocol with positive yield in the perfect fidelity limit, but which does not work on highly noisy entanglement. 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. (We could probably use a derandomization argument to avoid communicating the random strings in this protocol.) 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 5.7.4 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. 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. (The rotation takes the states back to the symmetric Werner form; $\log 12$ bits of information per pair is sufficient for this purpose [17].) We now use these EPR pairs along with

teleportation to effectively obtain a depolarizing channel of quantum capacity $Q > 0$. We use standard coding from Quantum Shannon theory [133] over this quantum channel to distribute N'' near perfect EPR pairs. This new step only requires a linear amount of classical communication. After the initial very noisy entanglement distribution step, we thus only have three classical messages to send over the depolarizing channel of classical capacity $C > 0$. We generate near perfect entanglement using the depolarizing channel a linear number of times, and then go on to the protocol implementation phase as before. Note that we are not yet guaranteed 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 in error, we adapt the above protocol. Before using teleportation and QECC to distribute near-perfect 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 number 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 property that some classical capacity achieving strategy for the depolarizing channel also has negligible error, we get the stated exponential decay in the error.

5.7.3.4 Discussion of Optimality

It is known that for some range of the depolarizing parameter, the quantum capacity Q_{B} with classical feedback of the depolarizing channel is strictly larger than its unassisted forward quantum capacity Q [17]. In particular, there exist 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 classical information. This proves that the standard forward quantum capacity of the quantum channels used does not characterize their communication capacity in the

interactive communication scenario. Note that $Q_{\mathrm{B}} > 0$ if and only if the depolarizing parameter $\varepsilon' < \frac{2}{3}$, and so $Q_{\mathrm{B}} > 0$ 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 $Q_{\mathrm{B}} > 0$, we can modify the method used in the proof of Theorem 5.7.6. We iteratively use the recurrence method a constant number of times on the noisy distributed EPR pairs, until the depolarizing channels induced through teleportation over the noisy distilled EPR pairs have non-zero forward quantum capacity. (Here the constant depends upon the depolarizing parameter, but not on N.) Then we 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 $\mathrm{O}(\frac{1}{Q_{\mathrm{B}}}N)$. 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.

5.7.4 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.

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 Ref. [16], 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 $\mathscr{T}_{\varepsilon'}$ 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 Ref. [16] 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.

5.7.4.1 Adversarial Errors in the Classical Channel

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 value of the parameter F. As described in the proof of Theorem 5.7.6, 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 non-zero quantum capacity Q. Standard quantum error correction techniques enable 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 goal. The protocol sketched above requires the communication of $2l_c + 1$ messages to distill near-perfect entanglement, independent of N, followed by a phase of simulating the message transmissions from the original protocol. The simulation protocol tolerates a constant error rate, though inversely proportional in l_c. It requires a constant rate of noisy entanglement consumption, which is exponential in l_c since each round of the recurrence method consumes at least half of the noisy EPR pairs. The protocol has a constant, positive rate of communication, though inversely proportional in the number of consumed noisy EPR pairs.

5.7.4.2 Random Errors in the Classical Channel

The case of noisy communication through binary symmetric channels once again is immediate from the adversarial error case by a concentration of measure argument. The communication rate of the resulting protocol is inversely proportional in the classical capacity C, and also in the number of noisy EPR pairs consumed.

5.8 Conclusion

We conclude with a discussion of our results and further research directions.

5.8.1 Discussion

In this chapter, we proposed a simulation of interactive quantum protocols intended for noiseless communication over noisy channels. Our approach is to replace irreversible measurements by reversible pseudo-measurements in the Cleve-Buhrman model (with shared entanglement and classical communication). Then, in the noisy version of the model, we teleport back and forth the corresponding quantum communication register to avoid losing quantum information. We develop a representation for such noisy quantum protocols that gives an analogue of Schulman's protocol tree representation for classical protocols. We prove that with this approach, it is possible to simulate the evolution of quantum protocols designed for noiseless quantum channels over noisy classical channels with only a linear dilation factor.

In the case of adversarial channel errors in which the parties are allowed to preshare a linear amount of entanglement, we prove that the error rate of $\frac{1}{2} - \varepsilon$ that our simulation tolerates is optimal for oblivious protocols. To get the tolerable error rate as high as $\frac{1}{2} - \varepsilon$, we develop new techniques along with a new bound on tree codes with an erasure symbol, Lemma 5.6.2. To simplify the exposition, we chose not to optimize different parameters, such as communication and entanglement consumption rates and communication register size.

We adapt our findings to a random error model in which parties are allowed to share entanglement but communicate over binary symmetric channels of capacity $C > 0$. We obtain communication rates proportional to C. We show that, up to a hidden constant, this is optimal for some family of distributed binary functions, for example the inner product functions $IP_n : \{0,1\}^n \times \{0,1\}^n \to \{0,1\}$, defined as $IP_n(x,y) = \oplus_{i=1}^n x_i \cdot y_i$. Our findings can also be adapted to obtain similar (though not optimal) results for the quantum model (the noisy version of Yao's model). Here, the simulation protocols run in two phases. In the first, a preprocessing phase, 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. This is followed by a simulation phase in which the actions of the parties parallel those in the shared entanglement model. In the case of adversarial noise, we show that we can tolerate an error rate of $\frac{1}{6} - \varepsilon$ in the quantum model. In the case of random noise in which the parties communicate over depolarizing channels of capacity $Q > 0$, we obtain rates proportional to Q. Perhaps surprisingly, we 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 $Q = 0$. This proves that Q does not characterize a quantum channel's capacity for interactive quantum communication. We extend our ideas to perform simulation in an extension of the shared entanglement model in which not only is the classical communication noisy, but also the entanglement is noisy.

5.8.2 Open Questions

A direction of research that immediately falls out of this work is characterizing the communication rates in all of the models discussed. In particular, the precise interactive capacity of the depolarizing channel with a specified noise parameter remains open. The question of interactive capacity for the binary symmetric channel was raised in the classical context by Schulman [114] and brought back to attention recently by Braverman in a survey article on the topic of interactive coding [32]. Recent developments provide lower and upper bounds for this quantity [90]. 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 for the study of such a quantity. Since every interactive quantum protocol can be mapped onto our general problem, it would be natural 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 that for classical protocols [90], up to a factor of two for teleportation? We show in upcoming articles that for small bit flip probability ε, the lower bound of $\frac{1}{2} - O(\sqrt{H(\varepsilon)})$ holds, and even extends to a lower bound of $1 - O(\sqrt{H(\varepsilon)})$ for depolarizing channels. Do the techniques developed in Ref. [90] adapt to the quantum setting

to obtain matching upper bounds of $\frac{1}{2} - \Omega(\sqrt{H(\varepsilon)})$ and $1 - \Omega(\sqrt{H(\varepsilon)})$, respectively? What about other channels?

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, it is likely that a fully quantum approach with new kinds of quantum codes needs to be developed. In particular, ideas from fault-tolerant quantum computation might be necessary. Furthermore, the important question of integrating our results into a larger fault-tolerant framework, in which the local operations are also noisy, remains open. Yet another important question for interactive quantum coding is what would happen in a shared entanglement setting if along with the noisy classical communication, the entanglement provided were also noisy; we investigated this question for a depolarizing noise model for the entanglement, but other models would also be interesting to study. In particular, what about adversarial noise on the shared EPR pairs above the unidirectional binary error rate limit? Note that below that bound, we can adapt the techniques we use here for distillation. Finally, the question of computationally efficient simulation also remains open, and we will show in upcoming works how to merge the techniques developed here with those of Brakerski and Kalai [27] to efficiently process the classical communication in our simulation protocols.

CHAPTER 6

CONCLUSION

6.1 Discussion

With the advent of quantum communication networks hopefully arriving in a relatively near future, it is important to develop a theory of information for interactive quantum protocols. The aim of this thesis was to present in a coherent way the foundation for such a theory. We studied the interactive quantum analogues of source coding and noisy channel coding. We also studied a unidirectionnal compression task, quantum state redistribution, which is strongly tied to the notion of quantum information complexity of interactive tasks.

In Chapter 3, we proved the first smooth entropy bounds on the amount of quantum communication required to implement quantum state redistribution; this is joint work with Mario Berta and Matthias Christandl [23]. In the asymptotic iid limit, we recover the previously known optimal rates for this task [54, 100, 140]. An additional result that we show is that our converse bounds even hold if we allow for feedback from the receiver to the sender, hence in the iid setting the conditional quantum mutual information lower bound on quantum communication is robust under interactive communication.

In Chapter 4, we introduced new, fully quantum notions of quantum information cost and complexity, and provided an operational interpretation for them as the amortized quantum communication complexity; this is from Ref. [129]. We proved that these quantities satisfy most of the important properties of their classical counterparts; this is from Ref. [129] and from joint work with Mark Braverman, Ankit Garg, Young Kun Ko and Jieming Mao [40]. In the case of classical inputs, we also provided an alternate characterization of quantum information cost that quantifies the cost of forgetting classical information; this is work in progress with Mathieu Laurière [94]. An application of the quantum information complexity paradigm, which also requires a protocol compression result, is the first general multi-round direct sum theorem for quantum communication

complexity; this is from Ref. [129]. Another application is to prove an optimal lower bound, up to polylogarithmic terms, on the bounded round quantum communication complexity of the disjointness function; this is joint work with Mark Braverman, Ankit Garg, Young Kun Ko and Jieming Mao [40].

In Chapter 5, we proposed the first schemes for implementing interactive quantum communication over noisy channels with only constant overhead, proving that quantum communication complexity is robust under noisy communication. We even show that it is possible to withstand a maximal fraction $\frac{1}{2} - \varepsilon$ of adversarial noise in this setting with perfect shared entanglement but noisy classical communication. This required the development of new bounds on classical interactive codes. Since the distribution of entanglement can be implemented with standard, unidirectional quantum communication, this shows that any channel with positive capacity for data transmission also has positive interactive quantum capacity. The idea that we develop, using teleportation over a noisy classical channel in order to evolve the simulation of a protocol over this channel, seems quite general and is applicable in other noise regimes. Perhaps surprisingly, we also show that some channels with zero unassisted quantum capacity have strictly positive unassisted interactive quantum capacity, proving that unassisted quantum capacity does not characterize a channel's capacity to implement interactive quantum communication.

6.2 Open Questions

Many interesting research directions fall out of these works. First, for quantum state redistribution in an interactive setting, it might be interesting to study this task in different models of communication than the ones we consider. For example, by allowing arbitrary pre-shared entanglement and variable length classical communication, would it be possible to implement this task in a one-shot setting with average communication cost close to the conditional quantum mutual information, without the $\frac{1}{\varepsilon}$ multiplicative factor inherent when we consider worst-case communication cost? For the analogous problem in the classical setting, Braverman and Rao [36] proved that this can indeed be done, although with a lot of interaction, at an average cost exactly equal, up to second order,

to the conditional mutual information. Note that in the quantum setting, even for the simpler case of source coding, such a result would be quite surprising since the variable length classical message cannot leak much information about the underlying quantum state.

In the context of interactive communication, when many successive rounds of communication have quantum information cost much lower than one qubit, is it possible to compress them globally at a low communication cost, which would be proportional to their total quantum information cost but not to the number of messages? In the classical setting, Barak, Braverman, Chen and Rao perform such a task in order to obtain an unbounded round direct sum theorem for communication complexity [10].

Another important question, which was settled only recently by Braverman and Schneider [37] in the classical setting, is whether quantum information complexity is computable. Indeed, it is known in the classical setting that to asymptotically reach the infimum in the definition of information complexity, an infinite sequence of protocols with increasing number of rounds can be required [38]. In Ref. [37], the authors provide a bound on the rate of convergence in terms of the number of rounds of the underlying protocols. Can we prove something similar in the quantum setting?

Apart from these foundational questions on quantum information complexity, it would also be interesting to find further applications to concrete lower bounds. Given its early success in settling the bounded round quantum communication complexity of disjointness, it is reasonable to be hopeful that this notion will find many such applications. Moreover, the recent breakthrough result of Fawzi and Renner, providing a powerful lower bound on the conditional quantum mutual information, should be helpful for these potential applications of quantum information complexity.

In the context of noisy interactive quantum coding, a first question would be to try to obtain good characterization of the interactive quantum capacity in the low noise regime for some well-studied channels, like the depolarizing and erasure channels, as well as for the binary symmetric channel with perfect entanglement assistance. In the entanglement-assisted setting, it is possible to use the techniques we developed in Chapter 5 along with ideas from classical interactive coding to obtain good rates of communication in this

regime. For quantum channels in the low noise regime, is it possible to generalize these ideas without requiring entanglement assistance? We believe that ideas from coherent communication [67] should lead to interesting results in this setting.

For adversarial quantum errors, it would be interesting to develop a fully quantum analogue of tree codes. These could be useful in order to avoid having two-phase protocols that first distribute entanglement before they implement interactive communication. New challenges arise in the quantum setting due to the fact that we cannot copy states at the different stages of the protocol, and thus we probably will have to perform computation of some sort on encoded data. Hence, ideas from fault-tolerant quantum computing will probably arise naturally in order to develop such quantum tree codes. A somewhat related question would be to study the question of noisy interactive quantum communication when the local quantum computation is also noisy. Can we still implement interactive communication with positive communication rates in such a setting?

All in all, many interesting questions remain in this young field of interactive quantum information theory, and hopefully the material in this thesis can serve as an appropriate introduction in order to study them.

BIBLIOGRAPHY

[1] Scott Aaronson and Andris Ambainis. Quantum search of spatial regions. In *Proceedings of the 44th Annual IEEE Symposium on Foundations of Computer Science*, pages 200–209, 2003.

[2] Harold Abelson. Lower bounds on information transfer in distributed computations. In *Proceedings of the 19th Annual IEEE Symposium on Foundations of Computer Science*, pages 151–158, 1978.

[3] Anura Abeyesinghe, Igor Devetak, Patrick Hayden, and Andreas Winter. The mother of all protocols: Restructuring quantum information's family tree. *Proceedings of the Royal Society of London. Series A*, 465(2108):2537–2563, 2009.

[4] Dorit Aharonov, Alexei Kitaev, and Noam Nisan. Quantum circuits with mixed states. In *Proceedings of the 30th Annual ACM Symposium on Theory of Computing*, pages 20–30, 1997.

[5] Robert Alicki and M. Fannes. Continuity of quantum conditional information. *Journal of Physics A: Mathematical and General.*, 37:L55, 2004.

[6] Anurag Anshu, Vamsi Krishna Devabathini, and Rahul Jain. Near optimal bounds on quantum communication complexity of single-shot quantum state redistribution. *arXiv:quant-ph/1410.3031*, 2014.

[7] Anurag Anshu, Rahul Jain, Priyanka Mukhopadhyay, Ala Shayeghi, and Penghui Yao. A new operational interpretation of relative entropy and trace distance between quantum states. *arXiv:quant-ph/1404.1366*, 2014.

[8] Alexei Ashikhim and Emanuel Knill. Non-binary quantum stabilizer codes. *IEEE Transactions on Information Theory*, 47(7):3065–3072, 2001.

[9] Ziv Bar-Yossef, T. S. Jayram, Ravi Kumar, and D. Sivakumar. An information statistics approach to data stream and communication complexity. In *Proceedings*

of the 43rd Annual IEEE Symposium on Foundations of Computer Science, pages 209–218, 2002.

[10] Boaz Barak, Mark Braverman, Xi Chen, and Anup Rao. How to compress interactive communication. In *Proceedings of the 42nd Annual ACM Symposium on Theory of Computing*, pages 67–76, 2010.

[11] Howard Barnum, Claude Crépeau, Daniel Gottesman, Adam Smith, and Alain Tapp. Authentication of quantum messages. In *Proceedings of the 43rd Annual IEEE Symposium on Foundations of Computer Science*, pages 449–458, 2002.

[12] Avraham Ben-Aroya, Oded Regev, and Ronald de Wolf. A hypercontractive inequality for matrix-valued functions with applications to quantum computing and LDCs. In *Proceedings of the 49th Annual IEEE Symposium on Foundations of Computer Science*, pages 477–486, 2008.

[13] Charles H. Bennett and Gilles Brassard. Quantum cryptography: Public key distribution and coin tossing. In *Proceedings of IEEE International Conference on Computers, Systems and Signal Processing*, pages 175–179, 1984.

[14] Charles H. Bennett and Stephen J. Wiesner. Communication via one- and two-particle operators on Einstein–Podolsky–Rosen states. *Physical Review Letters*, 69(20):2881–2884, 1992.

[15] Charles H. Bennett, Gilles Brassard, Claude Crépeau, Richard Jozsa, Asher Peres, and William K. Wooters. Teleporting an unknown quantum state via dual classical and Einstein–Podolsky–Rosen channels. *Physical Review Letters*, 70(13):1895–1899, 1993.

[16] 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*, 73(5):722–725, 1996.

[17] 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):3824–3851, 1996.

[18] 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):3081–3084, 1999.

[19] 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. *IEEE Transactions on Information Theory*, 48(10):2637–2655, 2002.

[20] Charles H. Bennett, Igor Devetak, Aram W. Harrow, Peter W. Shor, and Andreas Winter. Quantum reverse Shannon theorem. *IEEE Transactions on Information Theory*, 60(5):2926–2959, 2014.

[21] Mario Berta. Single-shot quantum state merging. *Master's thesis, ETH Zurich*, 2008.

[22] Mario Berta, Matthias Christandl, and Renato Renner. The quantum reverse Shannon theorem based on one-shot information theory. *Commununications in Mathematical Physics*, 306:579–615, 2011.

[23] Mario Berta, Matthias Christandl, and Dave Touchette. Smooth entropy bounds on one-shot state redistribution. *arXiv:quant-ph/1409.4338*, 2014.

[24] Mario Berta, Kaushik Seshadreesan, and Mark M. Wilde. Rényi generalizations of the conditional quantum mutual information. *Journal of Mathematical Physics*, 56(2):022205, 2015.

[25] Garry Bowen. Quantum feedback channels. *IEEE Transactions on Information Theory*, 50(10):2429–2434, 2004.

[26] Michel Boyer, Gilles Brassard, Peter Hoyer, and Alain Tapp. Tight bounds on quantum searching. *Fortschritte der Physik*, 46:493–506, 1998.

[27] Zvika Brakerski and Yael Tauman Kalai. Efficient interactive coding against adversarial noise. In *Proceedings of the 53rd Annual IEEE Symposium on Foundations of Computer Science*, pages 160–166, 2012.

[28] Zvika Brakerski and Moni Naor. Fast algorithms for interactive coding. In *Proceedings of the 24th Annual ACM-SIAM Symposium on Discrete Algorithms*, pages 443–456, 2013.

[29] Fernando G. S. L. Brandao, Matthias Christandl, and Jon Yard. Faithful squashed entanglement. *Commununications in Mathematical Physics*, 306:805–830, 2011.

[30] Gilles Brassard. Quantum communication complexity. *Foundations of Physics*, 33(11):1593–1616, 2003.

[31] Gilles Brassard, Ashwin Nayak, Alain Tapp, Dave Touchette, and Falk Unger. Noisy interactive quantum communication. In *Proceedings of the 55th Annual IEEE Symposium on Foundations of Computer Science*, pages 296–305, 2014.

[32] Mark Braverman. Coding for interactive computation: Progress and challenges. In *Proceedings of the 50th Annual Allerton Conference on Communication, Control, and Computing*, pages 1914–1921, 2012.

[33] Mark Braverman. Interactive information complexity. In *Proceedings of the 44th Annual ACM Symposium on Theory of Computing*, pages 505–524, 2012.

[34] Mark Braverman. Interactive information and coding theory. In *Proceedings of the International Congress of Mathematicians*, pages 535–559, 2014.

[35] Mark Braverman and Anup Rao. Towards coding for maximum errors in interactive communication. In *Proceedings of the 43rd Annual ACM Symposium on Theory of Computing*, pages 159–166, 2011.

[36] Mark Braverman and Anup Rao. Information equals amortized communication. In *Proceedings of the 52nd Annual IEEE Symposium on Foundations of Computer Science*, pages 748–757, 2011.

[37] Mark Braverman and Jon Schneider. Information complexity is computable. *ECCC: TR15-023*, 2015.

[38] Mark Braverman, Ankit Garg, Denis Pankratov, and Omri Weinstein. From information to exact communication. In *Proceedings of the 45th Annual ACM Symposium on Theory of Computing*, pages 151–160, 2013.

[39] Mark Braverman, Anup Rao, Omri Weinstein, and Amir Yehudayoff. Direct products in communication complexity. In *Proceedings of the 54th Annual IEEE Symposium on Foundations of Computer Science*, pages 746–755, 2013.

[40] Mark Braverman, Ankit Garg, Young Kun Ko, Jieming Mao, and Dave Touchette. Near-optimal bounds on bounded-round quantum communication complexity of disjointness. *Accepted to FOCS'15*, 2015.

[41] Harry Buhrman and Ronald de Wolf. Communication complexity lower bounds by polynomials. In *The 16th Annual IEEE Conference on Computational Complexity*, pages 120–130, 2001.

[42] Harry Buhrman, Richard Cleve, and Avi Wigderson. Quantum vs. classical communication and computation. In *Proceedings of the 30th Annual ACM Symposium on Theory of Computing*, pages 63–68, 1998.

[43] Francesco Buscemi and Nilanjana Datta. The quantum capacity of channels with arbitrarily correlated noise. *IEEE Transactions on Information Theory*, 56:1447–1460, 2010.

[44] A. Robert Calderbank and Peter W. Shor. Good quantum error-correcting codes exist. *Physical Review A*, 54(2):1098–1105, 1996.

[45] A. Robert Calderbank, Eric M. Rains, Peter W. Shor, and Neil. J. A. Sloane. Quantum error correction via codes over GF(4). *IEEE Transactions on Information Theory*, 44(4):1369–1387, 1998.

[46] A. Chakrabarti, Yaoyun Shi, A. Wirth, and Andrew C.-C. Yao. Informational complexity and the direct sum problem for simultaneous message complexity. In *Proceedings of the 42nd Annual IEEE Symposium on Foundations of Computer Science*, pages 270–278, 2001.

[47] Richard Cleve and Harry Buhrman. Substituting quantum entanglement for communication. *Physical Review A*, 56(2):1201–1204, 1997.

[48] Richard Cleve, Wim van Dam, Michael Nielsen, and Alain Tapp. Quantum entanglement and the communication complexity of the inner product function. In *Quantum Computing and Quantum Communications. Springer Berlin Heidelberg*, pages 61–74, 1999.

[49] Thomas M. Cover and Joy A. Thomas. *Elements of Information Theory*. John Wiley and Sons, Inc., 1991.

[50] Claude Crépeau, Daniel Gottesman, and Adam Smith. Approximate quantum error-correcting codes and secret sharing schemes. In *Advances in Cryptology-EUROCRYPT 2005. Springer Berlin Heidelberg*, pages 285–301, 2005.

[51] Nilanjana Datta and Min-Hsiu Hsieh. The apex of the family tree of protocols: optimal rates and resource inequalities. *New Journal of Physics*, 13:093402, 2011.

[52] Nilanjana Datta, Min-Hsiu Hsieh, and Jonathan Oppenheim. An upper bound on the second order asymptotic expansion for the quantum communication cost of state redistribution. *arXiv:quant-ph/1409.4352*, 2014.

[53] Igor Devetak. The private classical capacity and quantum capacity of a quantum channel. *IEEE Transactions on Information Theory*, 51(1):44–55, 2005.

[54] Igor Devetak and Jon Yard. Exact cost of redistributing multipartite quantum states. *Physical Review Letters*, 100(23):230501, 2008.

[55] Dennis Dieks. Communication by EPR devices. *Physics Letters A*, 92(6):271–272, 1982.

[56] David P. DiVincenzo, Peter W. Shor, and John A. Smolin. Quantum-channel capacity for very noisy channels. *Physical Review A*, 57(2):830–839, 1998.

[57] Lukas Drescher and Omar Fawzi. On simultaneous min-entropy smoothing. In *IEEE International Symposium on Information Theory Proceedings*, pages 161–165, 2013.

[58] Frédéric Dupuis. The decoupling approach to quantum information theory. *Ph.D. thesis, Université de Montréal.*, 2009.

[59] Frédéric Dupuis, Mario Berta, Jurg Wullschleger, and Renato Renner. One-shot decoupling. *Communications in Mathematical Physics*, 328:251, 2014.

[60] Omar Fawzi and Renato Renner. Quantum conditional mutual information and approximate Markov chains. *arXiv:quant-ph/1410.0664*, 2014.

[61] Keqin Feng and Zhi Ma. A finite Gilbert-Varshamov bound for pure stabilizer quantum codes. *IEEE Transactions on Information Theory*, 50(12):3323–3325, 2004.

[62] Matthew Franklin, Ran Gelles, Rafail Ostrovsky, and Leonard Schulman. Optimal coding for streaming authentication and interactive communication. In *Advances in Cryptology-CRYPTO 2013*, pages 1–20, 2013.

[63] Christopher A. Fuchs and Jeroen van de Graaf. Cryptographic distinguishability measures for quantum mechanical states. *IEEE Transactions on Information Theory*, 45:1216, 1999.

[64] Ran Gelles, Ankur Moitra, and Amit Sahai. Efficient and explicit coding for interactive communication. In *Proceedings of the 52nd Annual IEEE Symposium on Foundations of Computer Science*, pages 768–777, 2011.

[65] Daniel Gottesman. A class of quantum error-correcting codes saturating the quantum hamming bound. *Physical Review A*, 54(3):1862–1868, 1996.

[66] Lov K. Grover. A fast quantum mechanical algorithm for database search. In *Proceedings of the 28th Annual ACM Symposium on Theory of Computing*, pages 212–219, 1996.

[67] Aram W. Harrow. Coherent communication of classical messages. *Physical Review Letters*, 92(9):1–4, 2004.

[68] Prahladh Harsha, Rahul Jain, David McAllester, and Jaikumar Radhakrishnan. The communication complexity of correlation. In *The 22nd Annual IEEE Conference on Computational Complexity*, pages 10–23, 2007.

[69] Matthew B. Hastings. A counterexample to additivity of minimum output entropy. *Nature Physics*, 5(4):255–257, 2009.

[70] Patrick Hayden, Michał Horodecki, Jon Yard, and Andreas Winter. A decoupling approach to the quantum capacity. *Open Systems and Information Dynamics*, 15: 7–19, 2008.

[71] Alexander S. Holevo. Towards the mathematical theory of quantum communication channels. *Probl. Peredachi Inform. (in Russian)*, 8:63–71, 1972.

[72] Alexander S. Holevo. Bounds for the quantity of information transmitted by a quantum communication channel. *Probl. Peredachi Inform. (in Russian)*, 9:177–183, 1973.

[73] Alexander S. Holevo. The capacity of the quantum channel with general signal states. *IEEE Transactions on Information Theory*, 44(1):269–273, 1998.

[74] Michał Horodecki, Jonathan Oppenheim, and Andreas Winter. Partial quantum information. *Nature*, 436:673–676, 2005.

[75] Michał Horodecki, Jonathan Oppenheim, and Andreas Winter. Quantum state merging and negative information. *Commununications in Mathematical Physics*, 269:107, 2007.

[76] Peter Høyer and Ronald de Wolf. Improved quantum communication complexity bounds for disjointness and equality. In *Proceedings of the International Symposium on Theoretical Aspects of Computer Science*, pages 299–310, 2002.

[77] Rahul Jain and Hartmut Klauck. New results in the simultaneous message passing model via information theoretic techniques. In *The 24th Annual IEEE Conference on Computational Complexity*, pages 369–378, 2009.

[78] Rahul Jain and Ashwin Nayak. Short proofs of the quantum substate theorem. *IEEE Transactions on Information Theory*, 58(6):3664–3669, 2012.

[79] Rahul Jain and Ashwin Nayak. The space complexity of recognizing well-parenthesized expression. *IEEE Transactions on Information Theory*, 60(10): 1–23, 2014.

[80] Rahul Jain, Jaikumar Radhakrishnan, and Pranab Sen. Privacy and interaction in quantum communication complexity and a theorem about the relative entropy of quantum states. In *Proceedings of the 43rd Annual IEEE Symposium on Foundations of Computer Science*, pages 429–438, 2002.

[81] Rahul Jain, Jaikumar Radhakrishnan, and Pranab Sen. A lower bound for bounded round quantum communication complexity of set disjointness. In *Proceedings of the 44th Annual IEEE Symposium on Foundations of Computer Science*, pages 220–229, 2003.

[82] Rahul Jain, Jaikumar Radhakrishnan, and Pranab Sen. Prior entanglement, message compression and privacy in quantum communication. In *The 20th Annual IEEE Conference on Computational Complexity*, pages 285–296, 2005.

[83] Rahul Jain, Jaikumar Radhakrishnan, and Pranab Sen. A new information-theoretic property about quantum states with an application to privacy in quantum communication. *Journal of the ACM*, 56(6):33, 2009.

[84] Rahul Jain, Attila Pereszlényi, and Penghui Yao. A direct product theorem for bounded-round public-coin randomized communication complexity. In *Proceedings of the 53rd Annual IEEE Symposium on Foundations of Computer Science*, pages 167–176, 2012.

[85] Rahul Jain, Attila Pereszlényi, and Penghui Yao. A parallel repetition theorem for entangled two-player one-round games under product distributions. In *The 29th Annual IEEE Conference on Computational Complexity*, pages 209–216, 2014.

[86] Bala Kalyanasundaram and Georg Schnitger. The probabilistic communication complexity of set intersection. *SIAM Journal on Discrete Mathematics*, 5(4): 545–557, 1992.

[87] Iordanis Kerenidis, Mathieu Lauriere, Francois Le Gall, and Mathys Rennela. Privacy in quantum communication complexity. *arXiv:quant-ph/1409.8488*, 2014.

[88] Hartmut Klauck, Ashwin Nayak, Amnon Ta-Shma, and David Zuckerman. Interaction in quantum communication. *IEEE Transactions on Information Theory*, 53 (6):1970–1982, 2007.

[89] Hartmut Klauck, Robert Špalek, and Ronald de Wolf. Quantum and classical strong direct product theorems and optimal time-space tradeoff. *SIAM Journal on Computing*, 36:1472–1493, 2007.

[90] Gillat Kol and Ran Raz. Interactive channel capacity. In *Proceedings of the 45th Annual ACM Symposium on Theory of Computing*, pages 715–724, 2013.

[91] Ilan Kremer. Quantum communication. *Master's thesis, Hebrew University, Computer Science Department*, 1995.

[92] Eyal Kushilevitz and Noam Nisan. *Communication Complexity*. Cambridge University Press, 1997.

[93] Michael Langberg. Private codes or succinct random codes that are (almost) perfect. In *Proceedings of the 45th Annual IEEE Symposium on Foundations of Computer Science*, pages 325–324, 2004.

[94] Mathieu Laurière and Dave Touchette. The quantum cost of forgetting classical information. *In preparation*, 2015.

[95] Troy Lee, Adi Shraibman, and Robert Špalek. A direct product theorem for discrepancy. In *The 23rd Annual IEEE Conference on Computational Complexity*, pages 71–80, 2008.

[96] Debbie Leung and Graeme Smith. Communicating over adversarial quantum channels using quantum list codes. *IEEE Transactions on Information Theory*, 54(2):883–887, 2009.

[97] Ke Li and Andreas Winter. Relative entropy and squashed entanglement. *Commununications in Mathematical Physics*, 326(1):63–80, 2014.

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

[99] Seth Lloyd. Capacity of the noisy quantum channel. *Physical Review A*, 55(3): 1613–1622, 1997.

[100] Zhicheng Luo and Igor Devetak. Channel simulation with quantum side information. *IEEE Transactions on Information Theory*, 55(3):1331–1342, 2009.

[101] Laura Mančinska and Thomas Vidick. Unbounded entanglement can be needed to achieve the optimal success probability. In *Proceedings of the 41st International Colloquium on Automata, Languages and Programming*, pages 835–846, 2014.

[102] A. Nayak. Optimal lower bounds for quantum automata and random access codes. In *Proceedings of the 40th Annual IEEE Symposium on Foundations of Computer Science*, pages 369–376, 1999.

[103] Ashwin Nayak and Julia Salzman. Limits on the ability of quantum states to convey classical messages. *Journal of the ACM*, 53(1):184–206, 2006.

[104] Michael A. Nielsen. Quantum information theory. *PhD Dissertation, The University of New Mexico*, 1998.

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

[106] Noam Nisan and Avi Widgerson. Rounds in communication complexity revisited. In *Proceedings of the 23rd Annual ACM Symposium on Theory of Computing*, pages 419–429, 1991.

[107] Jonathan Oppenheim. State redistribution as merging: introducing the coherent relay. *arXiv: quant-ph/0805.1065*, 2008.

[108] Eric M. Rains. Nonbinary quantum codes. *IEEE Transactions on Information Theory*, 45(6):1827–1832, 1999.

[109] Ran Raz. Exponential separation of quantum and classical communication complexity. In *Proceedings of the 31st Annual ACM Symposium on Theory of Computing*, pages 358–367, 1999.

[110] Alexander Razborov. Quantum communication complexity of symmetric predicates. *Izvestiya of the Russian Academy of Science, mathematics*, 67(1):159–176 (145–159 in Engl. transl.), 2003.

[111] Renato Renner. Security of quantum key distribution. *Ph.D. thesis, ETH Zurich*, 2005.

[112] Leonard J. Schulman. Communication on noisy channels: A coding theorem for computation. In *Proceedings of the 33rd Annual IEEE Symposium on Foundations of Computer Science*, pages 724–733, 1992.

[113] Leonard J. Schulman. Deterministic coding for interactive communication. In *Proceedings of the 25th Annual ACM Symposium on Theory of Computing*, pages 747–756, 1993.

[114] Leonard J. Schulman. Coding for interactive communication. *IEEE Transactions on Information Theory*, 42(6):1745–1756, 1996.

[115] Benjamin Schumacher. Quantum coding. *Physical Review A*, 51(4):2738–2747, 1995.

[116] Benjamin Schumacher and Michael D. Westmoreland. Sending classical information via noisy quantum channels. *Physical Review A*, 56:131–138, 1997.

[117] Ronen Shaltiel. Towards proving strong direct product theorems. In *The 16th Annual IEEE Conference on Computational Complexity*, pages 71–80, 2001.

[118] Claude E. Shannon. A mathematical theory of communication. *Bell System Technical Journal*, 27:379–423, 1948.

[119] Alexander A. Sherstov. Strong direct product theorems for quantum communication and query complexity. In *Proceedings of the 43rd Annual ACM Symposium on Theory of Computing*, pages 41–50, 2011.

[120] Peter W. Shor. Algorithms for quantum computation: discrete logarithms and factoring. In *Proceedings of the 35th Annual IEEE Symposium on Foundations of Computer Science*, pages 124–134, 1994.

[121] Peter W. Shor. Scheme for reducing decoherence in quantum computer memory. *Physical Review A*, 52(4):2493–2496, 1995.

[122] Peter W. Shor. The quantum channel capacity and coherent information. *Lecture Notes, MSRI Workshop on Quantum Computation*, 2002.

[123] Graeme Smith and Jon Yard. Quantum communication with zero-capacity channels. *Science*, 321:1812–1815, 2008.

[124] Andrew Steane. Multiple particle interference and quantum error correction. *Proceedings of the Royal Society of London. Series A*, pages 2551–2577, 1996.

[125] Frode Terkelsen. Some minimax theorems. *Mathematica Scandinavica*, 31:405–413, 1972.

[126] Marco Tomamichel. A framework for non-asymptotic quantum information theory. *Ph.D. thesis, ETH Zurich*, 2012.

[127] Marco Tomamichel, Roger Colbeck, and Renato Renner. A fully quantum asymptotic equipartition property. *IEEE Transactions on Information Theory*, 55:5840–5847, 2009.

[128] Marco Tomamichel, Roger Colbeck, and Renato Renner. Duality between smooth min- and max-entropies. *IEEE Transactions on Information Theory*, 56:4674–4681, 2010.

[129] Dave Touchette. Quantum information complexity. In *Proceedings of the 47th Annual ACM Symposium on Theory of Computing*, pages 317–326, 2015.

[130] Wim van Dam and Patrick Hayden. Universal entanglement transformations without communication. *Physical Review A*, 67:060302(R), 2003.

[131] John Watrous. Theory of quantum information. *Lecture notes from Fall 2013, https://cs.uwaterloo.ca/ watrous/CS766/*, 2013.

[132] Reinhard F. Werner. Quantum states with Einstein–Podolsky–Rosen correlations admitting a hidden-variable model. *Physical Review A*, 40(8):4277–4281, 1989.

[133] Mark M. Wilde. *Quantum Information Theory*. Cambridge University Press, 2013.

[134] Mark M. Wilde, Nilanjana Datta, Min-Hsiu Hsieh, and Andreas Winter. Quantum rate distortion coding with auxiliary resources. *IEEE Transactions on Information Theory*, 59(10):6755–6773, 2013.

[135] Andreas Winter. Compression of sources of probability distributions and density operators. *arXiv:quant-ph/0208131*, 2002.

[136] William K. Wootters and Wojciech H. Zurek. A single quantum cannot be cloned. *Nature*, 299(5886):802–803, 1982.

[137] Aaron D. Wyner and Jacob Ziv. The rate-distortion function for source coding with side information at the decoder. *IEEE Transactions on Information Theory*, 22(1):1–10, 1976.

[138] Andrew C.-C. Yao. Some complexity questions related to distributive computing (preliminary report). In *Proceedings of the 11th Annual ACM Symposium on Theory of Computing*, pages 209–213, 1979.

[139] Andrew C.-C. Yao. Quantum circuit complexity. In *Proceedings of the 34th Annual IEEE Symposium on Foundations of Computer Science*, pages 352–361, 1993.

[140] Jon T. Yard and Igor Devetak. Optimal quantum source coding with quantum side information at the encoder and decoder. *IEEE Transactions on Information Theory*, 55(11):5339–5351, 2009.

[141] Ming-Yong Ye, Yan-Kui Bai, and Z. D. Wang. Quantum state redistribution based on a generalized decoupling. *Physical Review A*, 78:030302(R), 2008.