<table>
<thead>
<tr>
<th>Title and Subtitle</th>
<th>Optimization of Communication in Noisy Quantum Channels</th>
</tr>
</thead>
<tbody>
<tr>
<td>Author(s)</td>
<td>Mary Beth Ruskai</td>
</tr>
<tr>
<td>Performing Organization</td>
<td>UMass Lowell, 600 Suffolk St., Lowell, MA 01854</td>
</tr>
<tr>
<td>Sponsor/monitoring organization</td>
<td>U. S. Army Research Office, P.O. Box 12211, Research Triangle Park, NC 27709-2211</td>
</tr>
</tbody>
</table>

The views, opinions and/or findings contained in this report are those of the author(s) and should not be construed as an official Department of the Army position, policy or decision, unless so designated by other documentation.

Abstract

When quantum particles are used to transmit or process information, noise will affect the fidelity of the transmission. This project has been concerned with the analysis of mathematical models of noise for qubit channels, with the capacity of qubit channels used to transmit classical information, and with exchange errors in quantum computation. Some results were also obtained on adiabatic quantum computation.

Subject Terms

quantum information theory, channel capacity, error correction, adiabatic quantum computation

Security Classification

UNCLASSIFIED
A Overview

When quantum particles are used to transmit information, one can expect that, as with classical communication, noise in the channel will affect the fidelity of the transmission. This is true whether the particles are used to transmit quantum or classical information. Similar concerns arise in quantum computation when the noise due to interactions with the environment gives rise to errors.
This project has been concerned with the analysis of mathematical models of noise for qubit channels, with the capacity of qubit channels used to transmit classical information, and with exchange errors in quantum computation. Some related results on entropy inequalities were also obtained.

A large part of the project was devoted to questions about the capacity of channels used to transmit classical information, particularly the question of whether or not entangled inputs and/or measurements can increase the capacity. Questions have been raised recently as to whether or not quantum channels offer sufficient advantages to justify their use solely for transmission of classical information without prior entanglement. To some extent this skepticism is justified by results obtained in this project which showed no advantages for unital qubit channels and no advantages to entangled inputs when only product measurements are available. Such negative results are also valuable.

Moreover, results about situations when entangled inputs and/or measurements do increase the capacity of quantum channels contribute to our understanding of the role of entanglement in other situations. Finally, even when qubits are used to transmit quantum information within a system, such as a quantum computer, that information is not accessible until after a measurement has been performed. The entire process of input to output measurement in the computational basis can then be regarded as the transmission of classical information through a very complex composite channel. Thus the results of this project contribute to the much larger process of understanding the full spectrum of different types of quantum information systems.

In a somewhat different and unanticipated direction, the P.I. considered a proposed method for adiabatic computation. This work, which is described below, is the result of interest generated by hearing a report at the annual QCPR program review in August, 2001.

B Summary of Results

B.1 Analysis of Noise Maps

Noise in quantum information theory, whether the system is a communication channel, a quantum computer, or some other type of quantum information processor is described by a linear map on density matrices which has two mathematical properties known as completely positive and trace preserving. Such maps have been called “super-operators,” “quantum operations,” “stochastic maps” and often, simply, “channels” in the literature. Completely positive maps can always be written
[34, 46] (albeit non-uniquely) in the form

\[ \Phi(\rho) = \sum_k A_k \rho A_k^\dagger \]  \hspace{1cm} (1)

where \( \rho \) is a density matrix and the requirement that \( \Phi \) be trace-preserving is equivalent to the condition \( \sum_k A_k^\dagger A_k = I \).

Early work on channel capacity and related questions used a relatively small set of examples, defined via their Kraus operators \( A_k \) using (1). King and the P.I. [32] noted that (up to rotations) any qubit channel could be written as

\[ \Phi : \rho = \frac{1}{2} [I + \mathbf{w} \cdot \sigma] \mapsto \frac{1}{2} \left[ I + \sum_k (t_k + \lambda_k w_k) \sigma_k \right] = \Phi(\rho) \]  \hspace{1cm} (2)

where \( \sigma_j, \ (j = 1, 2, 3) \) denote the Pauli matrices and \( \mathbf{w} \) is a vector in \( \mathbb{R}^3 \). Thus, a stochastic map contracts the Bloch sphere to an ellipsoid whose axes have lengths \( \lambda_k \) and for which the vector \( \mathbf{t} \) (with components \( t_k \)) describes the translation of the center from the origin. However, not every possible ellipsoid arises as the image of a stochastic map. For unital maps (\( \mathbf{t} = 0 \)) Algoet and Fujiwara [19] showed that the \( \lambda_k \) must satisfy \( (\lambda_1 \pm \lambda_2)^2 \leq (1 \pm \lambda_3)^2 \) and also obtained some results when two of the \( t_k \) are zero. In [58], the Szarek, Werner and the P.I. obtained a complete set of necessary and sufficient inequalities on the parameters \( t_k \) and \( \lambda_k \) using Choi’s result [10] that a map is completely positive if and only if its action on a maximally entangled Bell state is positive semi-definite. In addition, we gave an explicit form for extreme channels, which are equivalent to the optimal cloning channels discovered independently by Niu and Griffiths [47] and by Rieffel and Zalka [59]. This work has given us a much richer class of examples with which to study channel capacity, leading to a number of new results some of which are discussed below.

In [24] Holevo introduced a special class of channels of the form

\[ \Phi(\rho) = \sum_k R_k \text{Tr} F_k \rho \]  \hspace{1cm} (3)

where each \( R_k \) is a density matrix and the \( \{F_k\} \) form a POVM. Holevo also introduced two subclasses and called a channel

- **classical-quantum (CQ)** if each \( F_k = |k \rangle \langle k | \) in the POVM is a one-dimensional projection,
- **quantum-classical (QC)** if \( \sum_k R_k = I \) and each density matrix \( R_k = |k \rangle \langle k | \) is a one-dimensional projection.
Holevo [23] gave examples of particular CQ maps for which entangled measurements can increase the capacity of the channel. Horodecki and Shor [26] recently observed that channels of the form (3) are entanglement breaking in the sense (B) below, and Shor [68] used this to show that entangled inputs cannot further increase the capacity of such channels.

Because it is important to understand the distinction between channels which break entanglement, those which preserve certain types of entanglement, and those which may be enhanced by entanglement (in the sense that entangled inputs can increase capacity), the P.I. undertook a more detailed study of entanglement breaking channels. In [56] it was shown that, for qubit channels, the following are equivalent

(A) $\Phi$ is a Holevo channel, i.e., $\Phi$ can be written in the form (3),

(B) $\Phi$ is entanglement breaking, i.e., $(I \otimes \Phi)(\Gamma)$ is always separable (for $\Gamma$ an arbitrary, possibly entangled, state on the tensor product space).

(C) $\Phi \circ T$ is completely positive, where $T(\rho) = \rho^T$ is the transpose.

(D) $\Phi$ has a sign change property (i.e., when $\Phi$ is written in the canonical form (2) and any one of the $\lambda_k \rightarrow -\lambda_k$ the resulting map is also completely positive).

(E) $\Phi$ is in the convex hull of CQ channels.

Thus, the properties of qubit entanglement breaking channels are now well understood. The P.I. also showed that, roughly speaking, entanglement breaking channels are those which are extremely noisy. For example, any qubit channel which maps the Bloch sphere into a plane is entanglement breaking.

For channels in $d$-dimensions most of the equivalences above extend to implications in only one direction. This is because the so-called partial transpose operation does not completely distinguish between entangled and product states of bipartite systems in higher dimension. Thus, a map may break some types of entanglement but preserve others. The results obtained thus far are described in Section 3 of [56]. (Recently P. Shor [67] informed the P.I. that he had found a counter-example to the conjecture that (A) implies (E) for $d = 3$.) An understanding of entanglement breaking maps for $d > 2$ seems to be closely related to understanding the different types of bipartite entanglement in higher dimensions.

### B.2 Channel Capacity

Shannon introduced the notion of capacity, which is roughly the maximum rate of reliable transmission of information per bit. In quantum information theory this concept can be generalized in several ways [6], depending on whether quantum or
classical information is transmitted and on the resources available. When quantum particles are used to transmit information, the noise is represented by the action of a stochastic map, $\Phi$. The capacity of a memoryless channel, i.e., one for which the noise on multi-bit signals is $\Phi \otimes \Phi \ldots \otimes \Phi$, need not necessarily be additive, because of the possibility of using entangled inputs and/or entangled measurements on the outputs.

A simple expression for the “classical capacity”, i.e., the optimal capacity of a memoryless quantum channel to transmit classical information has not yet been found. However, such expressions are known if the inputs are restricted to product states; we denote these by $C_{PP}(\Phi) = C_{Shan}(\Phi)$ or $C_{PE}(\Phi) = C_{Holv}(\Phi)$ depending on whether the measurements are required to be products or permitted to be entangled. $C_{EP}(\Phi)$ denotes the capacity when inputs are unrestricted, but measurements are required to be products, and $C_{EE}(\Phi)$ is the unrestricted capacity for transmitting classical information, i.e., the “classical capacity” of a quantum channel. The bound $C_{PP}(\Phi) \leq C_{Holv}(\Phi)$ is a consequence of Holevo’s 1973 bound [21] on the accessible information. Much later, Holevo [23] and (independently) Schumacher and Westmoreland [60, 61] showed that $C_{PE}(\Phi) = C_{Holv}(\Phi)$. Although, one can easily see that $C_{PP}(\Phi) \leq \{C_{PE}(\Phi), C_{EP}(\Phi)\} \leq C_{EE}(\Phi)$, little was known about $C_{PE}(\Phi)$ or its relation to $C_{EP}(\Phi)$ until the spring of 1999, when the C. King and the P.I. [33], proved that $C_{PP}(\Phi) = C_{PE}(\Phi)$. This result, which means that when only product measurements are used entangled inputs cannot increase the capacity, can be extended to a more general class of “sequential product measurements”. (This result was also obtained independently by P. Shor [66] and by A. Holevo [25], both of whom proved the extended result and used the P.I.’s observation that one could regard the noise as acting on the measurement through the adjoint of $\Phi$.) Thus, one now knows that

$$C_{PP}(\Phi) = C_{PE}(\Phi) \leq C_{EP}(\Phi) \leq C_{EE}(\Phi).$$

Moreover, one knows that the first inequality can be strict and that this is the generic situation for non-unital channels, i.e., maps for which $\Phi(I) \neq I$. The question of whether or not equality always holds in $C_{EP}(\Phi) \leq C_{EE}(\Phi)$ is one of the major open questions and is closely related to the question of whether or not the Holevo capacity is always additive, i.e., does equality always hold in $C_{Holv}(\Phi \otimes \Omega) \geq C_{Holv}(\Phi) + C_{Holv}(\Omega)$ or is superadditivity possible? Both questions can be rephrased as asking if entangled inputs can ever enhance the classical capacity of a quantum channel when the noise acts independently on subsystems.

In the special case of unital qubit maps, King and the P.I. [32] observed that the additivity of $C_{Holv}(\Phi)$ is equivalent to the additivity of minimal entropy, a long-standing conjecture of P. Shor [67]. In [32] King and the P.I. gave strong evidence for the latter conjecture by showing that for unital qubit maps entangling states of minimal entropy could not decrease the entropy below that of the optimal product.
We also gave the first result on additivity of another measure of purity and additional evidence in support of the entropy conjecture in other situations. Building on this work, and related conjectures [2] about other measures of purity, King [28, 29] succeeded in proving the additivity of both minimal entropy and Holevo capacity in a number of special cases, including situations in which one of the maps is a unital qubit channel and the other is completely arbitrary. (The P.I. played a minor role by contributing to the lemma in the Appendix of [28].) Thus, at least for transmission of classical information, qubit channels with unital stochastic maps have little advantage over classical channels because the capacity is optimized when a particular pair of orthogonal inputs is used to represent 0,1.

For practical applications, it is important to know the actual inputs which maximize the capacity. Fuchs [18] was the first to observe that there were situations in which maximizing the Holevo capacity required the use of non-orthogonal inputs, and it was subsequently realized [32, 62] that this behavior is generic for non-unital channels. Recently, the P.I. (with C. King and M. Nathanson) [31] showed that there are also situations in which qubit channels require three (non-orthogonal) inputs to maximize the Holevo capacity, i.e., even for tensor products of two-level quantum systems, entangled measurements can best distinguish between alphabets of product inputs when these products are formed from three non-orthogonal states \{\rho_0, \rho_1, \rho_2\}.

Although the increase in capacity associated with 3- and 4-input channels and with superadditivity, is likely to be small, there are still potential practical applications. The need for three inputs arises from a competition between a QC capacity with an asymmetric probability distribution, and a CQ capacity with a 50/50 distribution. It may be possible to exploit the flatness near the optimum to use a variety of alphabet distributions with only a small sacrifice in capacity.

It is noteworthy that 3-input channels were not found in the very extensive numerical searches for channels with other properties, such as superadditivity of the Holevo capacity. The construction in [31] exploits a symmetry which results in an essentially planar image which cannot require more than 3 inputs. One would expect that 4-input qubit channels also exist and that these would be good candidates for superadditivity. However, they seem to be rare events which are difficult to find because one must break the symmetry used to construct 3-input examples.

Moreover, in retrospect, many of the numerical searches for superadditivity were futile because they studied examples which are now known to satisfy general additivity theorems [29, 30, 68]. Using the results of [29, 56, 58, 68] it is now possible to conduct a much more focused search. In work this summer [12] many additional 3-input channels were found with quite different characteristics than those of [31], but 4-input qubit channels remain elusive. (By fundamental convexity theorems, qubit channels never require more than 4 inputs.) The P.I. suspects that 4-input qubit channels do exist, but are rare.
In any case, one should not draw any conclusions about superadditivity until such 4-input qubit channels are found and tested. Recently, two groups [27, 68] found connections between the additivity of the Holevo capacity and that of the entanglement of formation. Hence, this question has implications well beyond any small increase in capacity which may exist. Finding channels for which entangled inputs enhance capacity, may lead to the discovery of additional situations in which entanglement enhances information processing.

B.3 Equality Conditions for Entropy Inequalities

An important tool in quantum information theory is the strong-subadditivity (SSA) property of quantum mechanical entropy, i.e.,

\[ S(\rho_{123}) + S(\rho_2) \leq S(\rho_{12}) + S(\rho_{23}) \] (5)

where \( S(\rho) = -\text{Tr} \rho \log \rho \) denotes the von Neumann entropy of the density matrix \( \rho \) and the subscripts refer to subsystems. This property, which was proved by Lieb and the P.I. [42] is closely related to the joint convexity of relative entropy

\[ H(\rho, \gamma) = \text{Tr} \rho \log \rho - \text{Tr} \rho \log \gamma \] (6)

where \( \rho, \gamma \) are density matrices with \( \ker(\gamma) \subset \ker(\rho) \). SSA can also be used to derive the monotonicity of relative entropy under stochastic maps which, in turn, can be used to prove the Holevo bound [21] for accessible information. Hence, there has been some interest in knowing the conditions for equality in SSA and related inequalities. The P.I. showed [55] that equality holds in (5) if and only if

\[ \log \rho_{123} - \log \rho_{12} = \log \rho_{23} - \log \rho_2 \] (7)

which can be regarded as a kind of quantum Markov condition. (A different form of this result was obtained independently by Petz [49, 51].) The P.I. also obtained conditions for equality in the joint convexity and monotonicity of relative entropy under stochastic maps. Because the monotonicity of relative entropy can be used to give a simple proof of the Holevo bound on accessible information, the results of [55] yield a new short proof that the this bound can be achieved only when the density matrices in the upper bound all commute.

This paper also gives a simple, self-contained proof of SSA, following the original strategy in [42] but using Epstein’s elegant proof [14] that the map \( A \to \text{Tr} e^{K+\log A} \) is concave in \( A \) rather than Lieb’s [41] original proof of this concavity. The paper is intended to make the proofs of important entropy inequalities accessible to a wide audience without requiring a high level of mathematical sophistication.
**B.4 Pauli Exchange Errors**

Most discussions of quantum error correction assume, at least implicitly, that errors result from interactions with the environment and that single qubit errors are much more likely than two qubit errors. Most of the quantum computing literature also does not make explicit the spatial component of the qubit wave function and thus seems to ignore the Pauli exclusion principle and permutational symmetry of states describing multi-qubit systems. In [53] the P.I. showed, by considering the full wave function, that interactions between qubits of identical particles can give rise to a type of error, not seen classically, in which a single exchange error can affect two qubits. An explicit 9-qubit code was constructed [53] which can handle both Pauli exchange errors and all one-bit errors.

Subsequently, Lidar, et al [36] noted that their DFS (Decoherence Free Subspace) codes were resistant to this type of error, because exchange errors on physical qubits acted as single Pauli errors on their logical bits. Turning this idea around, they [37, 38] observed that exchange interactions could be used for universal computation in the DFS scheme. Then a larger group [40] observed that exchange interactions could also be used as to implement a set of universal gates in other situations.

There is some question as to whether exchange errors will be important in physical realizations of practical quantum computers, or whether, instead, such computers will be designed to be sufficiently robust to use exchange interactions to implement gates. In either case, it appears that [53] attracted attention to a topic which (at least in some small way) affected subsequent developments.

Moreover, the Berkeley group [37] observed that some DFS codes could be regarded as a different type of stabilizer code and the P.I. proposed [53] the construction of more powerful permutationally invariant codes involving higher dimensional representations of the symmetric group. Recently, the P.I. realized that the proposal in [53] was too strong, but that both the DFS and permutationally invariant codes can be regarded as stabilizers for certain non-Abelian groups. The P.I.’s 9-bit code in [53] is “non-additive” in the sense of not arising as the stabilizer of an Abelian subgroup of the Pauli group. The P.I. has now begun to examine the more general question of stabilizer codes associated with non-Abelian groups (not necessarily contained within the Pauli group). Such codes may be useful for adaptive error correction in quantum computers which are robust in some ways, but vulnerable to certain types of correlated errors.

**B.5 Adiabatic Quantum Computation**

Recently Fahri, et al [15, 16] proposed using adiabatic evolution to construct a type of analogue quantum computer and suggested that such computers might be able to
solve NP-complete problems efficiently. However, the only evidence for their claim is based upon extrapolation from small $n$ using numerical simulations of adiabatic quantum evolution. A proof (or refutation) of their claim would require rigorous analysis of the behavior of the lowest eigenvalue gap as $n$ increases. But this has only been done for very simple models.

The P.I. has shown [57] that for a large class of Hamiltonians, including those proposed for adiabatic quantum computation, the ground state is unique. This explains the eigenvalue gaps that have been observed numerically, but says nothing at all about their behavior as $n$ increases.

Several groups [1, 13, 69, 15] have studied adiabatic searches using the Grover oracle to construct the final Hamiltonian. They have observed that, although the gap decreases exponentially, knowledge of its location does permit some speed-up of the algorithm. In this analysis, the reduction to an essentially two-dimensional problem, which plays a critical role in Grover’s original algorithm is essential.

In a completely different approach, the P.I. [57] has observed that because only the ground state plays a role, an adiabatic search can be performed using a more general Hamiltonian than that associated with the Grover oracle $G$ (which has a very high degeneracy because of the two-dimensional structure). In particular, replacing any final $H_1$ by $GH_1$, where $G$ multiplies the “target” state by $-1$, permitting adiabatic quantum computation to perform an unsorted search. There are three possible situations:

(A) Both the problem encoded in $H_1$ and an unsorted search can be performed in polynomial time, contradicting the conventional wisdom [4, 46] that one can not improve on the $O(2^{n/2}) = O(\sqrt{N})$ behavior of Grover’s algorithm.

(B) Both the problem encoded in $H_1$ and an unsorted search require exponential time, refuting Fahri, et al’s claim of efficient solution of hard problems.

(C) The simple process of moving a single excited state below the ground state, equivalent to subtracting a multiple of a one-dimensional projection from $H_1$, drastically alters the computational complexity of finding the ground state by adiabatic quantum computation.

Fahri and Guttman [17] have given a simple model problem in which situation (C) holds. However, this problem also has a high level of symmetry which prevents crossings and gives an eigenvalue gap that is $O(1)$ precisely because it permits a high level of persistent degeneracy (with only $n + 1$ distinct eigenvalues). Moving just one excited state breaks much of this symmetry. However, a preliminary analysis suggests that the existence of a slowly decreasing eigenvalue gap requires a symmetry (or other structure) which allows sufficient degeneracy (or squeezing) for $2^n$ states to have eigenvalues in the range of $[0, n]$ with at least one gap that decreases slowly.
Vazirani [69, 70] has pointed out that situation (A) is not necessarily precluded by
the standard arguments because the construction of the final Hamiltonian in adiabatic
quantum computation seems to involve a more complex type of oracle. Indeed, it has
recently been observed [11, 45] that improvements on Grover’s algorithm may be
possible in other situations.

The resolution of the efficiency of adiabatic quantum computation for hard prob-
lems is likely to require a different approach; one possibility, building on ideas from
random Schrödinger operators is discussed briefly in [57].

C List of Publications


b) C. King and M.B. Ruskai, “Minimal Entropy of States Emerging from Noisy

c) C. King and M.B. Ruskai, “Capacity of Quantum Channels Using Product

d) M.B. Ruskai, S. Szarek and E. Werner, “An analysis of completely positive

e) C. King, M. Nathanson and M.B. Ruskai “Qubit Channels Can Require More

f) M.B. Ruskai, “Inequalities for Quantum Entropy: A Review with Conditions

g) M.B. Ruskai, “Pauli-Exchange Errors and Quantum Error correction” in press
for AMS conference proceedings. quant-ph/0006008

h) “Comments on Adiabatic Quantum Computation” to appear in QMath-8
conference proceedings in AMS Contemporary Mathematics series.

D Scientific personnel

In the summer of 2001, this grants provided some support for an undergraduate student (Stephen Boyer) and a graduate student (Michael Nathanson) both at Northeastern University and working jointly with the P.I. and Christopher King, who is on the faculty at Northeastern. The work of Michael Nathanson contributed to the construction of a 3-input example which was discussed in Section B.2 and published [31].

In the summer of 2002, John Cortese, a graduate student with John Preskill at Caltech, spent 8 weeks working with the P.I. This work [12] is discussed briefly at the end of Section B.2. Travel and subsistence for his visit were provided by this grant.

The P.I. also collaborated with Christopher King (Northeastern University), Harriet Pollatsek (Mt. Holyoke College), Stanislaw Szarek (University of Paris VI) and Elisbathe Werner (Case Western Reserve University). Some travel expenses for H. Pollatsek and E. Werner were provided by this grant.

References


[26] M. Horodecki and P. Shor, private communications cited in [56] and [68]


[59] E. Rieffel and C. Zalka, private communication, independently obtained the extreme points in [58].


[67] P. Shor, private communication


