Quantum Information and Computation, Vol. 13, No. 1&2 (2013) 0009–0020
© Rinton Press

FIDELITY AS A FIGURE OF MERIT IN QUANTUM ERROR CORRECTION

JONAS ALMLÖF[a]
Department of Applied Physics, Royal Institute of Technology (KTH), AlbaNova University Center
Stockholm, S-106 91, Sweden

GUNNAR BJÖRK
Department of Applied Physics, Royal Institute of Technology (KTH), AlbaNova University Center
Stockholm, S-106 91, Sweden

Received Oct 29 2011
Revised June 18, 2012

We discuss the fidelity as a figure of merit in quantum error correction schemes. We show that when identifiable but uncorrectable errors occur as a result of the action of the channel, a common strategy that improves the fidelity actually decreases the transmitted mutual information. The conclusion is that while the fidelity is simple to calculate and therefore often used, it is perhaps not always a recommendable figure of merit for quantum error correction. The reason is that while it roughly speaking encourages optimisation of the "mean probability of success", it gives no incentive for a protocol to indicate exactly where the errors lurk. For small error probabilities, the latter information is more important for the integrity of the information than optimising the mean probability of success.

Keywords: quantum error correction, mutual information, fidelity

Communicated by: R Cleve & B Terhal

1 Introduction

Quantum computing, i.e., physical operations on qubits and entangled qubits, has attracted interest since the '70s, see e.g. [1] and the references therein. Early on, it was realised that a major obstacle for its implementation is the undesired, but unavoidable, interaction with the environment, described as a quantum channel. Since this interaction typically is not unitary, the channel-qubit interaction can not be "undone" using a single unitary transformation. However, it was discovered in [2] that using a specific coding, a set of conditional, unitary transformations, each invoked by the result of a cleverly chosen measurement that locates and identifies the error, could correct one Pauli bit-flip error, phase error, or a combination of the two. In the footsteps of this discovery, many new quantum error correction codes (QECCs) were developed, including those based on the stabiliser formalism [3, 4, 5], Calderbank-Shor-Steane (CSS) codes [6, 7], decoherence-free subspace codes [8, 9], channel adapted codes such as amplitude damping channel codes [5, 10]. Such QECCs utilise redundant information for their function. These QECCs have often been compared using fidelity, i.e., essentially the probability that the state after being acted on by the channel and corrected by the so-called

[a] jalml@kth.se

recovery operations, is identical to the initial state. Therefore, this figure of merit (FOM) is a measure of "likeness", and it has the characteristic property that identical states have unit fidelity, while orthogonal states have zero fidelity. We have also presented a code with fidelity as its FOM [11]. However, in this work we point out some problems associated with the use of fidelity as a FOM for QECCs.

First we would like to comment on "similarity measures", i.e., measures that quantify how well one quantum state resembles another, and how they are fundamentally different from information integrity measures, such as Shannon's "rate of transmission" [12] – nowadays referred to as mutual information [13, 14]. E.g, assume that Alice has three binary, classical channels to use for communicating with Bob, one with fidelity $F = 1$, one with $F = 0.5$ and one with $F = 0$. At a first glance, the $F = 1$ channel appears to be the best choice, but in fact, the $F = 0$ channel is equally good, because Bob needs only to bit-flip the data stream to get the correct data sequence. In contrast, the $F = 0.5$ channel is utterly useless (if this fidelity is due to random noise), since it gives a vanishing value of the mutual information between Alice and Bob. Conversely, the maximum value of the mutual information for a given channel achieved by optimising the input alphabet, equals to the channel capacity.

The quantum counterpart of mutual information is the quantum mutual information [24, 25], and analogously to the situation above, for transmission of qubits, it vanishes for $F = 0.5$. It is also true that a QECC with zero overall fidelity could indeed be a very useful one!

Quantum mutual information and its classical counterpart, are relative measures that quantify how much information two parties can agree upon. Mutual information relates two sets of outcomes linked by a joint probability distribution, and quantum mutual information relates two subsystems described by a bipartite density operator. Also, a connection between quantum mutual information and perfect error correction exists, as shown in [26] and references therein.

Another illustrative example is given by Shannon [12]: Assume a transmission of 1000 (uncoded) classical bits per second - each bit taking the values 0 or 1 with probability 1/2, and also assume an error rate of 1%, where the errors result with equal probabilities in 0 and 1 regardless of the original bit value. A reader not familiar with information theory may be tempted to guess that the rate of transmission will decrease with 1%, becoming 990 bits per second. However, the loss of transmission rate (equivocation) in this case is significantly larger, about 8.1%. This is due to the lack of knowledge of *where* the errors are located. This knowledge suggests that in the case of detectable but uncorrectable errors, the faulty bits should be discarded and their locations "tagged". The error rate would then be minimised to about 10 bits per second. Note that had we wanted to optimise "similarity" between the sent and received bit-strings, we could in the case of detectable errors "improve" the erroneous bits by replacing them by random bits with equal probabilities. This would cause the number of identical bits between sender and receiver to become 995 per second on average. However, this optimisation of similarity will not increase the rate of information transmission at all. This illustrates the seemingly odd fact that the integrity of each bit is more important than how similar the input symbol strings are to the output ones.

Criteria for the construction of QECCs exist for the case where errors are reversible, i.e., expressed in terms of Pauli X- and Z operators [15, 16] where, given a maximum number of such errors e, perfect recovery of a coded logical qubit can be achieved. If, on the other hand, logical code words are allowed to interact with a reservoir system, such errors can approximately be corrected if criteria given in [10] are satisfied, reaching a fidelity of the same order (in the error rate) as the reversible case. Other codes take advantage of the fact that in addition to correcting e errors, any occurrence of $e+1$ errors can also be detected, so that such uncorrectable qubits can be discarded [16].

When more than e errors occur for a coded qubit, the ensuing state either fall into a space orthogonal to the code space, or overlap it. If such an erroneous state belongs to the orthogonal space, it can be identified but not corrected. There are then two conceivable strategies: One is to apply some recovery operator $\mathcal{R}_i$ that effectively replaces the erroneous state with a predefined state. The other strategy is to discard the state and take note of the location of the error (i.e., to "tag" it).

If the code is not perfect, i.e., if the code space does not completely occupy the extended space including all possible errors, one is left a possibility to recover additional errors. How these recovery operators should be constructed is not immediately given by the code criteria, but has to be established as the optimum recovery for a particular FOM, typically maximising fidelity for a fixed assumed statistical distribution of channel input states, and a given encoding and channel interaction [17, 18, 19]. If, in addition, the error rate is known, the FOM can be improved significantly [5]. In these contexts the optimisation FOMs used are fidelity, entanglement fidelity, minimum fidelity or average fidelity. When discussing QECC in the context of quantum computers, other measures of closeness has been introduced [16], particularly in connection to estimating errors from additional sources, such as quantum gate errors. It has been shown that for trace-preserving completely positive maps, optimising such closeness measures, e.g., trace distance, is akin to optimising fidelity [17, 20]. However, closeness measures such as fidelity can be deceitful as we have exemplified above.

Below we show that optimisation of the mutual information does not lead to optimisation of "probability of sameness", i.e., fidelity, and vice versa. In Sec. 2 we shall compare the two figures of merit for a simple classical channel, using two different strategies of handling errors. We shall then repeat the calculation for a quantum channel in Sec. 3. As for the channel, we will assume that for some of the error events, all information about the error type was lost in the channel interaction. We compare two strategies that deal with such errors, one devised in [21] where the scheme applies the identity operator to map the error back to the code space – in doing so optimising fidelity – and one where the qubit is discarded and its location is "tagged" and transmitted to the receiver. We shall show that the two strategies lead to opposite results, the former gives a higher fidelity than the latter, while the latter gives a higher mutual information than the former.

2 Classical treatment

To study the effects of the two strategies in a classical setting, we consider errors caused by bit flipping. A sequence of bits is sent through a channel which influences the individual bits independently, and causes them to flip, $0 \leftrightarrow 1$, with some probability p_f per unit time.

Assume we have a classical binary channel and that at a certain time $t = 0$ the probability that the bit is incorrect is $\gamma(0)$. This means that between the times t and $t + dt$, the change in probability of having the incorrect state will be

$$d\gamma = -\gamma p_f dt + (1 - \gamma) p_f dt. \tag{1}$$

Going to the limit $dt \to 0$ and solving the ensuing differential equation one finds that

$$\gamma(t) = \left(\gamma(0) - \frac{1}{2}\right) \exp(-2p_f t) + \frac{1}{2}. \tag{2}$$

Below we shall assume that after passing through the flip channel, each bit is in the incorrect state with probability γ, where $0 \leq \gamma \leq 1/2$. We shall also assume that the bits flip independently of each other, so that the probability of having, e.g., four bits all in the correct state after passing through the channel is $(1-\gamma)^4$, and having exactly three out of four bits in the correct state is $4\gamma(1-\gamma)^3$.

To demonstrate our thesis, we will first consider error correction in this classical channel model. Assume, e.g., that we code a logical zero onto the four bit string 0000, and the logical one onto the string 1111. This code can correct single bit flip errors, e.g., 0010 will be interpreted as a logical zero, and the code can identify two bit flip errors, e.g., 0011 can either be the result of flipping the last two bits of 0000 or the first two bits of 1111. If three errors occur, then the string will be incorrectly interpreted as correctable and the "correction" will result in an erroneous bit. We write the channel model for such a channel in Table 1.

Table 1. The channel matrix for a symmetric flip channel.

Input x	Output y		
	0	1	Uncorrectable
0	p_{00}	p_{01}	p_{02}
1	p_{01}	p_{00}	p_{02}

For simplicity we have assumed that the channel is symmetric, that is, both the bit flips and the coding is symmetric with respect of the permutation $0 \leftrightarrow 1$. The probability of obtaining the correct bit after decoding is denoted $p_{00} = (1-\gamma)^4 + 4\gamma(1-\gamma)^3$, the probability for detecting an uncorrectable string (e.g., 0101) is denoted $p_{02} = 6\gamma^2(1-\gamma)^2$, and the probability for erroneous correction is $p_{01} = \gamma^4 + 4\gamma^3(1-\gamma)$. An asymmetric channel will yield quantitatively different but qualitatively similar results.

Using this channel model we shall investigate two strategies. The first, called I, is to randomly assign the value zero and the value one when we find that the code word we get cannot be corrected. Hence, the effective new channel matrix is shown in Table 2.

Table 2. The effective channel matrix for a random assignment of bit-value when the error cannot be corrected.

Input x	Output y	
	0	1
0	$p_{00} + \frac{p_{02}}{2}$	$p_{01} + \frac{p_{02}}{2}$
1	$p_{01} + \frac{p_{02}}{2}$	$p_{00} + \frac{p_{02}}{2}$

We see from the table that this strategy will increase the probability of success from p_{00} to $p_{00} + p_{02}/2$. At the same time the probability of error will increase from p_{01} to $p_{01} + p_{02}/2$. If we were to optimise the probability for success, which is akin to maximise the fidelity between the received and corrected state in a quantum error correction scheme, then this would be a good strategy.

The transmitted mutual information I is defined

$$I(X,Y) = S(X) + S(Y) - S(X,Y), \tag{3}$$

where $S(X)$ is the entropy function

$$S(X) = -\sum_x p(x)\log_2[p(x)] \tag{4}$$

of the random variable X expressed in bits. Using the properties of logarithmic function and probability distributions, the mutual information can be expressed in the simple form

$$I(X,Y) = -\sum_{x,y} p(x,y)\log_2[\frac{p(x)p(y)}{p(x,y)}]. \tag{5}$$

We shall assume that we transmit bits, that is $p(x=0) = p(x=1) = 1/2$. The assumption of symmetric channel then leads to $p(y=0) = p(y=1) = 1/2$. From the table and the assumptions we also get $p(0,0) = p(1,1) = p_{00} + p_{02}/2$ and $p(0,1) = p(1,0) = p_{01} + p_{02}/2$. The mutual information in bits can now readily be computed as

$$I_I = 1 + \left(p_{00} + \frac{p_{02}}{2}\right)\log_2\left(p_{00} + \frac{p_{02}}{2}\right) + \left(p_{01} + \frac{p_{02}}{2}\right)\log_2\left(p_{01} + \frac{p_{02}}{2}\right). \tag{6}$$

Another strategy, called II, is to simply discard any string that is identified to be uncorrectable. If the information is to be transmitted further, this information is irrevocably lost, but any such string will be identified by appropriate tagging, e.g., by sending a message on a separate channel pointing out which bit strings should be ignored. The channel model *for the "successful" bits that are left after this "error correction"* is displayed in Table 3.

Table 3. The effective channel matrix for the bits that are corrected, after bits that have been determined as erroneous but uncorrectable have been tagged as such and discarded.

Input x	Output y	
	0	1
0	$\frac{p_{00}}{1-p_{02}}$	$\frac{p_{01}}{1-p_{02}}$
1	$\frac{p_{01}}{1-p_{02}}$	$\frac{p_{00}}{1-p_{02}}$

In this case, quite trivially, the overall probability for success is $p_{00} \leq p_{00} + p_{02}/2$. The marginal distributions will remain the same, but the mutual information of each of these "successful" bits is now

$$I_{\rm OK} = 1 + \left(\frac{p_{00}}{1-p_{02}}\right)\log_2\left(\frac{p_{00}}{1-p_{02}}\right) + \left(\frac{p_{01}}{1-p_{02}}\right)\log_2\left(\frac{p_{01}}{1-p_{02}}\right). \tag{7}$$

However, in a long string of bits, a fraction p_{02} of the transmitted and coded bits are going to be tagged to be discarded or ignored. Hence, the average transmitted mutual information is

$$I_{\rm II} = (1-p_{02})I_{\rm OK}\ . \tag{8}$$

A plot of $I_{\rm II}$ and $I_{\rm I}$ is shown in Fig. 1. It is seen that $I_{\rm II} \geq I_{\rm I}$, and that it is only equal for the totally deterministic and the totally random channel. In Fig. 2 the success probability is plotted, that is, the classical counterpart to the fidelity between the input and output bits. For strategy II we assign zero success if the bit is either incorrectly "corrected" or tagged as uncorrectable. It is clear that for protocol II, where the probability of success is lower than for protocol I, the transmitted mutual information is higher. This indicates that optimisation of the success probability of a error correction protocol does not lead to a (simultaneous) maximisation of the mutual information, and vice versa, in general.

Next we shall show that the same conclusion holds for a quantum channel.

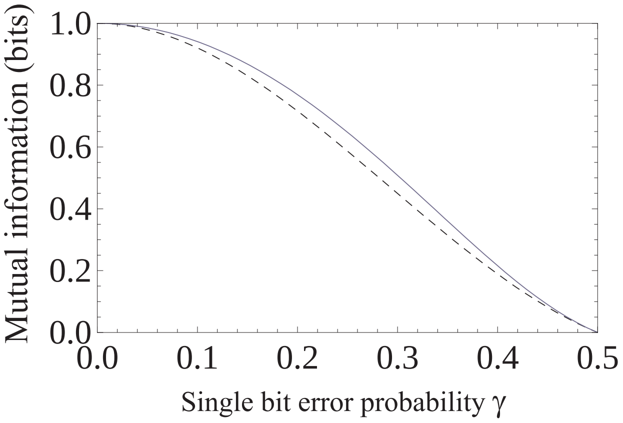


Fig. 1. The mutual information between a sent logical bit, and the coded, received, and corrected bit, as a function of the probability that a bit passing through the flip channel becomes incorrect. Strategy I is drawn dashed and strategy II drawn solid.

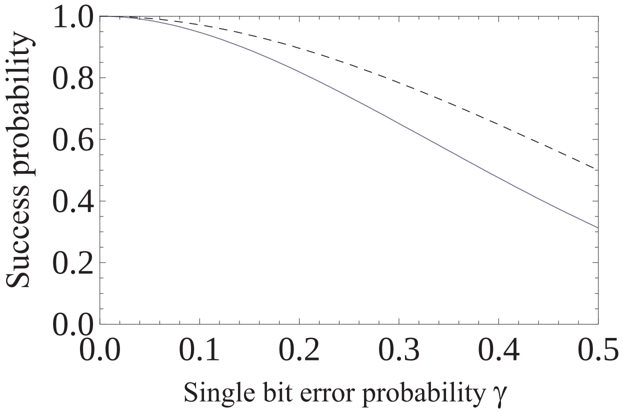


Fig. 2. The probability that a sent logical bit, and the coded, received, and corrected bit have the same values, as a function of the probability that a bit passing through the flip channel becomes incorrect. Strategy I is drawn dashed and strategy II drawn solid.

3 Quantum treatment

In the following, to exemplify our assertion, we shall look at a specific channel and a specific code. For simplicity we shall look at a qubit flip channel which is a special case of a Pauli channel. The code we shall employ is the $[[1,5,3]]$ perfect code discovered independently in [22] and [23]. This code uses five physical qubits to code one logical qubit and it can correct one Pauli error, that is a bit flip, a sign flip, or a combination of both. These errors are operationally described by the Pauli operators $\hat{\sigma}_x$, $\hat{\sigma}_z$, and $\hat{\sigma}_y$, hence the name.

One implementation of the $[[1,5,3]]$ code uses the following encoding [22]:

$$\begin{aligned}
|0_{\rm L}\rangle \quad \rightarrow \quad & |S_{000}\rangle = (-|00000\rangle + |01111\rangle - |10011\rangle + |11100\rangle \\
& + |00110\rangle + |01001\rangle + |10101\rangle + |11010\rangle)/\sqrt{8} \\
|1_{\rm L}\rangle \quad \rightarrow \quad & |S_{100}\rangle = (-|11111\rangle + |10000\rangle + |01100\rangle - |00011\rangle \\
& + |11001\rangle + |10110\rangle - |01010\rangle - |00101\rangle)/\sqrt{8} \qquad (9)
\end{aligned}$$

The code words are orthogonal and the code is designed such that that bit flipping ($|0\rangle \leftrightarrow |1\rangle$) any one of the five qubits will give an additional 2×5 state vectors that are mutually orthogonal, and also orthogonal to the code words. Likewise, sign flipping ($-|1\rangle \leftrightarrow |1\rangle$) or simultaneously flipping both the bit-value and the sign $\pm|0\rangle \rightarrow \pm|1\rangle$ and $\pm|1\rangle \rightarrow \mp|0\rangle$) of any of the five qubits will result in twenty vectors that are mutually orthogonal, and also orthogonal to the 12 previous vectors. Hence, these 32 vectors, called syndrome vectors, will form an orthonormal basis in the $2^5 = 32$-dimensional Hilbert code space. We will use the notation $|S_{jkl}\rangle$, where the index j denotes the logical qubit (0 or 1), $1 \leq k \leq 5$ indicates which qubit is erroneous, and $l = 1,2,3$ denotes that a $\hat{\sigma}_x$-, $\hat{\sigma}_z$-, or $\hat{\sigma}_y$-flip has occurred, respectively. The corresponding eigenvalue is denoted S_{jkl}.

In the following we will, for the sake of reasoning, assume the following model: A long string of states, randomly selected between two orthogonal qubit states $|Q\rangle$ and $|Q_\perp\rangle$ is generated in duplicate. (Should the states not be selected with equal probability our conclusions below would still hold qualitatively, but would be quantitatively somewhat different.) In the ensemble sense, the generated states are thus described by the following density matrix:

$$\hat{\rho} = \frac{1}{2}\left(|Q\rangle \otimes |Q\rangle \langle Q| \otimes \langle Q| + |Q_\perp\rangle \otimes |Q_\perp\rangle \langle Q_\perp| \otimes \langle Q_\perp|\right). \qquad (10)$$

One of the duplicate states is subsequently encoded onto a five-qubit state according to the encoding (9). The coded state is sent to a receiver through a flip channel where any qubit will be flipped with probability γ independently of any other qubit. On the receiver side, the five-qubit states are measured by a quantum non-demolition, von Neumann measurement having the code words and the syndrome vectors as eigenstates, with pairwise degenerate eigenvalues. The states $|S_{000}\rangle$ and $|S_{100}\rangle$ form one such pair, and the syndrome vectors $|S_{0kl}\rangle$ and $|S_{1kl}\rangle$ constitute the 15 other pairs. We shall then apply one of two different strategies with the measured five-qubit state:

I Decode it by a measurement-result specific unitary operation, unless the measurement collapsed the vector onto one of the vectors $|S_{jk2}\rangle$ and $|S_{jk3}\rangle$, $j = 0,1$, $k = 1,\ldots,5$, in which case the erroneous state is randomly replaced by either $|0_{\rm L}\rangle$ or $|1_{\rm L}\rangle$. In the ensemble sense it is replaced with the maximally mixed state $(|0_{\rm L}\rangle\langle 0_{\rm L}| + |1_{\rm L}\rangle\langle 1_{\rm L}|)/2$.

II Decode it by a measurement-result specific unitary operation, unless the measurement collapsed the vector onto one of the vectors $|S_{jk2}\rangle$ or $|S_{jk3}\rangle$, $j = 0, 1$, $k = 1, \ldots, 5$, in which case the qubit is flagged as uncorrectable and is discarded.

The motivation behind the two strategies is as follows: In strategy I we use the $[[1, 5, 3]]$ code, but since the code is designed for the Pauli channel, which is more general than the flip channel we have assumed, it is not optimal. We know from the code's construction, that a single bit flip in a code word will always result in a state that is orthogonal to the states $|S_{jk2}\rangle$ and $|S_{jk3}\rangle$, $j = 0, 1$, $k = 1, \ldots, 5$. Hence, if the measured vector collapses onto one of the syndrome vectors $|S_{jk2}\rangle$ or $|S_{jk3}\rangle$ we know that there has been more than one flip error. This implies that we cannot correct the vector, as the code is designed to correct only single errors. In order to increase the fidelity between the final vector and the sent vector we shall randomly replace the measured vector with $|0_\mathrm{L}\rangle\langle 0_\mathrm{L}|$ or $|1_\mathrm{L}\rangle\langle 1_\mathrm{L}|$. In an ensemble sense, the measured state is replaced by the density matrix $(|0_\mathrm{L}\rangle\langle 0_\mathrm{L}| + |1_\mathrm{L}\rangle\langle 1_\mathrm{L}|)/2$. A motivation for this strategy is, e.g., given in [21]: "As the initial state ... is known to be in the code space, it is clearly more beneficial to return the state ... to the code space than do otherwise: lacking any other information one could at least prepare the completely mixed state in the code space$(|0\rangle\langle 0| + |1\rangle\langle 1|)/2$, yielding an average fidelity of 1/2, rather than leaving the register outside the code space, yielding the fidelity of 0." We have used the same strategy to increase a code's fidelity in [11].

Strategy II does exactly that [21] argues against, namely when the five-qubit state is projected onto an uncorrectable syndrome vector, it is identified as uncorrectable and sent outside the code space. In this case, since the code and syndrome vectors span the whole space, only the null vector remains. The result is exactly the one Rahn *et al.* predicts, the average fidelity becomes lower than using strategy I, but as we shall see below, *the average quantum mutual information between the sent and error corrected qubit is higher than for strategy I.*

A general logical-qubit state can be written

$$|Q\rangle = \sin\alpha\,|0_\mathrm{L}\rangle + e^{i\phi}\cos\alpha\,|1_\mathrm{L}\rangle\,. \tag{11}$$

The state orthogonal to $|Q\rangle$ is

$$|Q_\perp\rangle = \cos\alpha\,|0_\mathrm{L}\rangle - e^{i\phi}\sin\alpha\,|1_\mathrm{L}\rangle\,. \tag{12}$$

Using (9), (10), (11), (12), and the channel model we can compute the joint density matrix after the coded state has been subjected to the flip channel and been corrected using either strategy I or II. The probability of having no flip will be $(1-\gamma)^5$, and the probability of having qubit 1, 2 and 3 flipped is $\gamma^3(1-\gamma)^2$. In the second case, the code word $|0_\mathrm{L}\rangle$ will become

$$(-|11100\rangle + |10011\rangle - |01111\rangle + |00000\rangle + |11010\rangle + |10101\rangle + |01001\rangle + |00110\rangle)/\sqrt{8}. \tag{13}$$

However, it is straightforward to compute that this is identically the same state as the syndrome $|S_{113}\rangle$, that is the syndrome for the case when the first qubit of code word$|1_\mathrm{L}\rangle$ has undergone a simultaneous bit and sign flip. Hence, the three-bit flipped state $|0_\mathrm{L}\rangle$ will be detected as erroneous but uncorrectable. Depending on strategy, this state will subsequently be “corrected” to become $(|0_\mathrm{L}\rangle\langle 0_\mathrm{L}| + |1_\mathrm{L}\rangle\langle 1_\mathrm{L}|)/2$ according to strategy I, or detected as erroneous and uncorrectable, and be discarded according to strategy II.

Assume that the state $|Q\rangle\langle Q| = \hat{\rho}(0)$ was sent. After being transmitted through the channel it becomes $\hat{\rho}(\gamma)$. Suppose this state is measured by a quantum non-demolition detector in the syndrome vector basis, and that the syndrome measurement resulted in the (degenerate) eigenvalue associated with the indices kl. The state then collapses into the unnormalised density matrix

$$\hat{\varrho}_{kl} = \begin{pmatrix} \langle S_{0kl}|\,\hat{\rho}(\gamma)\,|S_{0kl}\rangle & \langle S_{0kl}|\,\hat{\rho}(\gamma)\,|S_{1kl}\rangle \\ \langle S_{1kl}|\,\hat{\rho}(\gamma)\,|S_{0kl}\rangle & \langle S_{1kl}|\,\hat{\rho}(\gamma)\,|S_{1kl}\rangle \end{pmatrix} \tag{14}$$

when expressed in the $\{|0_\mathrm{L}\rangle, |1_\mathrm{L}\rangle\}$ basis. The trace of $\hat{\varrho}_{kl}$ is equal to the probability of obtaining the measurement results S_{0kl} or S_{1kl}. If the syndrome has $l = 1$ the state is correctable via the unitary transformation $\hat{U}(k) = |0_\mathrm{L}\rangle\langle S_{0k1}| + |1_\mathrm{L}\rangle\langle S_{1k1}|$. If the syndrome has $l = 2, 3$ it is not correctable, and one can show that the measured syndrome gives no clue as how to correct the state. Summing up the undisturbed, and the correctable contributions, one arrives at the density matrix

$$\hat{\varrho} = \hat{\varrho}_{00} + \sum_{k=1}^{5} \hat{\varrho}_{k1}. \tag{15}$$

The probability to receive an uncorrectable state is $P(\hat{\rho}(\gamma)) = \sum_{k=1}^{5}\sum_{l=2}^{3}(\langle S_{0kl}|\,\hat{\rho}(\gamma)\,|S_{0kl}\rangle + \langle S_{1kl}|\,\hat{\rho}(\gamma)\,|S_{1kl}\rangle)$. Replacing $\hat{\rho}(\gamma)$ in (14) and (15) by $\hat{\rho}_\perp(\gamma)$ gives the state matrix $\hat{\varrho}_\perp$ and the probability $P(\hat{\rho}_\perp(\gamma))$ of not being able to correct the state, given that $|Q_\perp\rangle\langle Q_\perp| = \hat{\varrho}_\perp(0)$ was sent. Then the symmetry of the code implies that $P(\hat{\rho}(\gamma)) = P(\hat{\rho}_\perp(\gamma))$ in this case. The overall matrix ensuing from strategy I, given that the sent matrix is $(|Q\rangle\langle Q| \otimes |Q\rangle\langle Q| + |Q_\perp\rangle\langle Q_\perp| \otimes |Q_\perp\rangle\langle Q_\perp|)/2$ is thus a block diagonal matrix, that we can symbolically express

$$\hat{\rho}_\mathrm{c} = \frac{1}{2}\begin{pmatrix} \hat{\rho} & 0 \\ 0 & \hat{\rho}_\perp \end{pmatrix} \tag{16}$$

in the basis $\{|0_\mathrm{L}\rangle \otimes |Q\rangle, |1_\mathrm{L}\rangle \otimes |Q\rangle, |0_\mathrm{L}\rangle \otimes |Q_\perp\rangle, |1_\mathrm{L}\rangle \otimes |Q_\perp\rangle\}$, where the diagonal elements can be identified as

$$\hat{\rho} = \hat{\varrho} + P(\hat{\rho}(\gamma))\mathbb{1}/2, \qquad \hat{\rho}_\perp = \hat{\varrho}_\perp + P(\hat{\rho}(\gamma))\mathbb{1}/2. \tag{17}$$

The identity operator in (17) represents the projection onto the $(|0_\mathrm{L}\rangle\langle 0_\mathrm{L}| + |1_\mathrm{L}\rangle\langle 1_\mathrm{L}|)/2$ state.

Strategy I now dictates that we have $\mathrm{Tr}(\hat{\rho}_\mathrm{c}) = 1$ due to the completeness of the code- and syndrome vectors. From the assumption of equal probability in the transmitted sequence of vectors $|Q\rangle$ and $|Q_\perp\rangle$ we have that $\rho_{\mathrm{c}11} + \rho_{\mathrm{c}22} = \rho_{\mathrm{c}33} + \rho_{\mathrm{c}44} = 1/2$. Due to the symmetry of the code under the action of bit flipping we also have $\rho_{\mathrm{c}11} + \rho_{\mathrm{c}33} = \rho_{\mathrm{c}22} + \rho_{\mathrm{c}44} = 1/2$. Hence, the qubits $|0_\mathrm{L}\rangle$ and $|1_\mathrm{L}\rangle$ are received with equal probability.

In the case of strategy II, *for all states that are not flagged as uncorrectable*, one instead gets $P(\hat{\rho}) = 0$ and using this P, one gets $\mathcal{N} = \rho_{\mathrm{c}11} + \rho_{\mathrm{c}22} + \rho_{\mathrm{c}33} + \rho_{\mathrm{c}44}$. The fraction of these correctable states in a large ensemble will be $\mathcal{N}$. Again, due to the symmetry of sent state, channel, and code, the mutual information between the sender's and the receiver's error corrected and decoded qubits is given by Eq. (18).

The mutual information between the sender's and the receiver's error corrected and decoded qubits (in bits) is

$$I(\gamma, \alpha, \phi) = S(\hat{\rho}_{\mathrm{s}}) + S(\hat{\rho}_{\mathrm{r}}) - S(\hat{\rho}_{\mathrm{c}}) = 1 + S(\hat{\rho}_{\mathrm{r}}) - S(\hat{\rho}_{\mathrm{c}}), \tag{18}$$

where

$$S(\hat{\rho}) = -\mathrm{Tr}[\hat{\rho}\log_2(\hat{\rho})] \tag{19}$$

is the von Neumann entropy (in bits) and $\hat{\rho}_{\mathrm{s}}$ and $\hat{\rho}_{\mathrm{r}}$ are obtained by tracing out the receiver and sender system, respectively, from $\hat{\rho}_{\mathrm{c}}$. The unity in the simplified, right-hand expression of (18) comes from the fact that the sent qubits have the average entropy 1 bit. A numerical evaluation of this expression yields Fig. 3(a), which shows the mutual information as a function of α and γ. (The function is quite naturally symmetric with respect to the line $\gamma = 1/2$). The function is very weakly dependent on ϕ and the figure is drawn for $\phi = 0$.

The average fidelity between the sender's qubit and the received and decoded qubit is

$$F = \mathrm{Tr}\left(\sqrt{\sqrt{\hat{\rho}_{\mathrm{s}}}\hat{\rho}_{\mathrm{r}}\sqrt{\hat{\rho}_{\mathrm{s}}}}\right), \tag{20}$$

where

$$\hat{\rho}_{\mathrm{s}} = \frac{1}{2}\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \tag{21}$$

and

$$\hat{\rho}_{\mathrm{r}} = \begin{pmatrix} \rho_{\mathrm{c}11} + \rho_{\mathrm{c}33} & \rho_{\mathrm{c}12} + \rho_{\mathrm{c}34} \\ \rho_{\mathrm{c}21} + \rho_{\mathrm{c}43} & \rho_{\mathrm{c}22} + \rho_{\mathrm{c}44} \end{pmatrix} \tag{22}$$

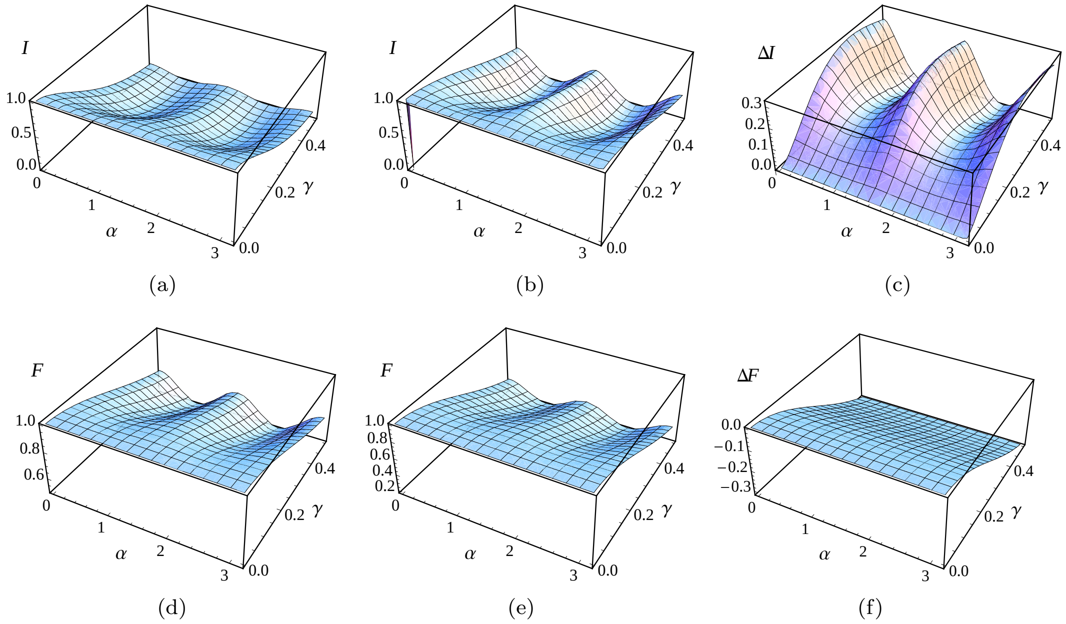
in the $\{|Q\rangle, |Q_{\perp}\rangle\}$ basis. Due to the simple form of $\hat{\rho}_{\mathrm{s}}$, one gets $F = (\sqrt{\lambda_1} + \sqrt{\lambda_2})/\sqrt{2}$, where λ_1 and λ_2 are the eigenvalues of $\hat{\rho}_{\mathrm{r}}$. A numerical evaluation of this expression yields Fig. 3(d).

However, the received information, averaged over all sent states, is given by $\mathcal{N}I(\gamma)$ since the uncorrectable states carry no information. A numerical evaluation of the expression $\mathcal{N}I(\gamma)$, valid for strategy II, yields Fig. 3(b).

The average fidelity in this case is obtained in the same manner as for strategy I, except for the fact that the so obtained fidelity must be multiplied with the factor $\mathcal{N}$ to take into account the instances when the states are uncorrectable and hence do not contribute to the fidelity. A numerical evaluation of the average fidelity, valid for strategy II, yields Fig. 3(e).

4 Conclusions

We have shown for both a classical and a quantum channel that if one wants to optimise the fidelity between a sent string of (qu)bits and the string after it has been error correction coded, transmitted through a noisy channel, and its error correction syndrome has been measured, then (qu)bits that can be identified as erroneous but that cannot be corrected should be mapped back onto the code space using the identity operation. We called this strategy I. If instead, one would like to optimise the (quantum) mutual information between the sender and receiver in an error correcting context, the best strategy is to discard (qu)bits that can be identified as erroneous but that cannot be corrected. This strategy was called II. Strategy I results in a lower (quantum) mutual information than strategy II, while strategy



(a) (b) (c)

(d) (e) (f)

Fig. 3. The mutual information between the sender's and the receiver's error corrected and decoded qubits (in bits), using strategy I in (a), strategy II in (b) and the difference, $I(\alpha,\gamma)_{\mathrm{II}} - I(\alpha,\gamma)_{\mathrm{I}}$ in (c). The average fidelity between the sender's qubit and the received and decoded qubit, using strategy I in (d), strategy II in (e) and the difference $F(\alpha,\gamma)_{\mathrm{II}} - F(\alpha,\gamma)_{\mathrm{I}}$ in (f).

II results in a lower fidelity than strategy I. This illustrates an important insight, namely that fidelity and mutual information are not necessarily positively correlated in quantum error correction schemes. Depending on what quantum information protocol one intends to implement, one may want to optimise one of these figure of merits, but this should be done knowing that, in general, it will be done at the expense of the other. Hence, while fidelity is quite straightforward to calculate in comparison to quantum mutual information, one should avoid drawing the conclusion that all modifications that result in a higher fidelity will also increase the mutual information (and vice versa). Our simple example shows that the two figures of merit in general require different strategies to optimise.

Acknowledgements

This work was supported by the Swedish Research Council (VR) through its Linnæus Center of Excellence ADOPT.

References

1. R. P. Feynman (1986), *Quantum mechanical computers*, Found. Phys., 16, p. 507.
2. P. W. Shor (1995), *Scheme for reducing decoherence in quantum computer memory*, Phys. Rev. A, 52, R2493.
3. D. Gottesman (1996), *Class of quantum error-correcting codes saturating the quantum Hamming bound*, Phys. Rev. A, 54, p. 1862.
4. R. Cleve (1997), *Quantum stabilizer codes and classical linear codes*, Phys. Rev. A, 55, pp. 4054-

4059.

5. A. S. Fletcher, P. W. Shor, M. Z. Win (2008), *Channel-Adapted Quantum Error Correction for the Amplitude Damping Channel*, IEEE Trans. Inf. Theory, 54, p. 5705.
6. A. R. Calderbank and P. W. Shor (1996), *Good quantum error-correcting codes exist*, Phys. Rev. A, 54, pp. 1098-1106.
7. A. M. Steane (1996), *Multiple particle interference and quantum error correction*, Proc. R. Soc. London Ser. A, 452, p. 2551.
8. G. M. Palma, K.-A. Suominen and A. K. Ekert (1996), *Quantum Computers and Dissipation*, Proc. R. Soc. London Ser. A, 452, p. 567.
9. D. A. Lidar, I. L. Chuang and K. B. Whaley (1998), *Decoherence-Free Subspaces for Quantum Computation*, Phys. Rev. Lett, 81, p. 2594.
10. D. W. Leung, M. A. Nielsen, I. L Chuang, and Y. Yamamoto (1997), *Approximate quantum error correction can lead to better codes*, Phys. Rev. A, 56, p. 2567.
11. J. Almlöf, G. Björk (2011), *A short and efficient error correcting code for polarization coded photonic qubits in a dissipative channel*, Opt. Comm., 284, p. 550.
12. C. E. Shannon (1948), *A Mathematical Theory of Communication*, Bell Syst. Tech. J., 27, pp. 379-423 and 623-656.
13. R. Hamming (1980), *Coding and Information Theory*, Prentice Hall, Englewood Cliffs.
14. T. M. Cover and J. A. Thomas (1991), *Elements of information theory*, John Wiley & Sons, New York, NY.
15. A. Ekert and C. Machiavello (1996), *Quantum Error Correction for Communication*, Phys. Rev. Lett., 77, p. 2585.
16. E. Knill and R. Laflamme (1997), *Theory of quantum error-correcting codes*, Phys. Rev. A, 55, p. 900.
17. S. Taghavi, R. L. Kosut, and D. A. Lidar (2010), *Channel-Optimized Quantum Error Correction*, IEEE Trans. Inf. Theory, 56, p. 1461.
18. N. Yamamoto, S. Hara, and K. Tsumura (2005), *Suboptimal quantum-error correcting procedure based on semidefinite programming*, Phys. Rev. A, 71, 022322.
19. A. S. Fletcher, P. W. Shor, and M. Z. Win (2007), *Channel-optimized quantum error correction*, Phys. Rev. A, 75, 012338.
20. B. Schumacher and M. D. Westmoreland (2002), *Approximate Quantum Error Correction*, Quantum Inf. Process., 1, p. 5.
21. B. Rahn, A. C. Doherty, and H. Mabuchi (2002), *Exact performance of concatenated quantum codes*, Phys. Rev. A, 66, 032304.
22. R. Laflamme, C. Miquel, J. P. Paz., and W. H. Zurek (1996), *Perfect Quantum Error Correcting Code*, Phys. Rev. Lett. 77, p. 198.
23. C. H. Bennett, D. P. DiVincenzo, J. A. Smolin, and W. K. Wootters (1996), *Mixed state entanglement and quantum error correction*, Phys. Rev. A, 54, p. 3824.
24. C. Adami, N. J. Cerf, (1997), *von Neumann capacity of noisy quantum channels*, Phys. Rev. A, 56, p. 3470.
25. N. J. Cerf, C. Adami (1997), *Negative Entropy and Information in Quantum Mechanics*, Phys. Rev. Lett., 79, p. 5194.
26. T. Ogawa (2005), *Perfect quantum error-correcting condition revisited*, arXiv:quant-ph/0505167.