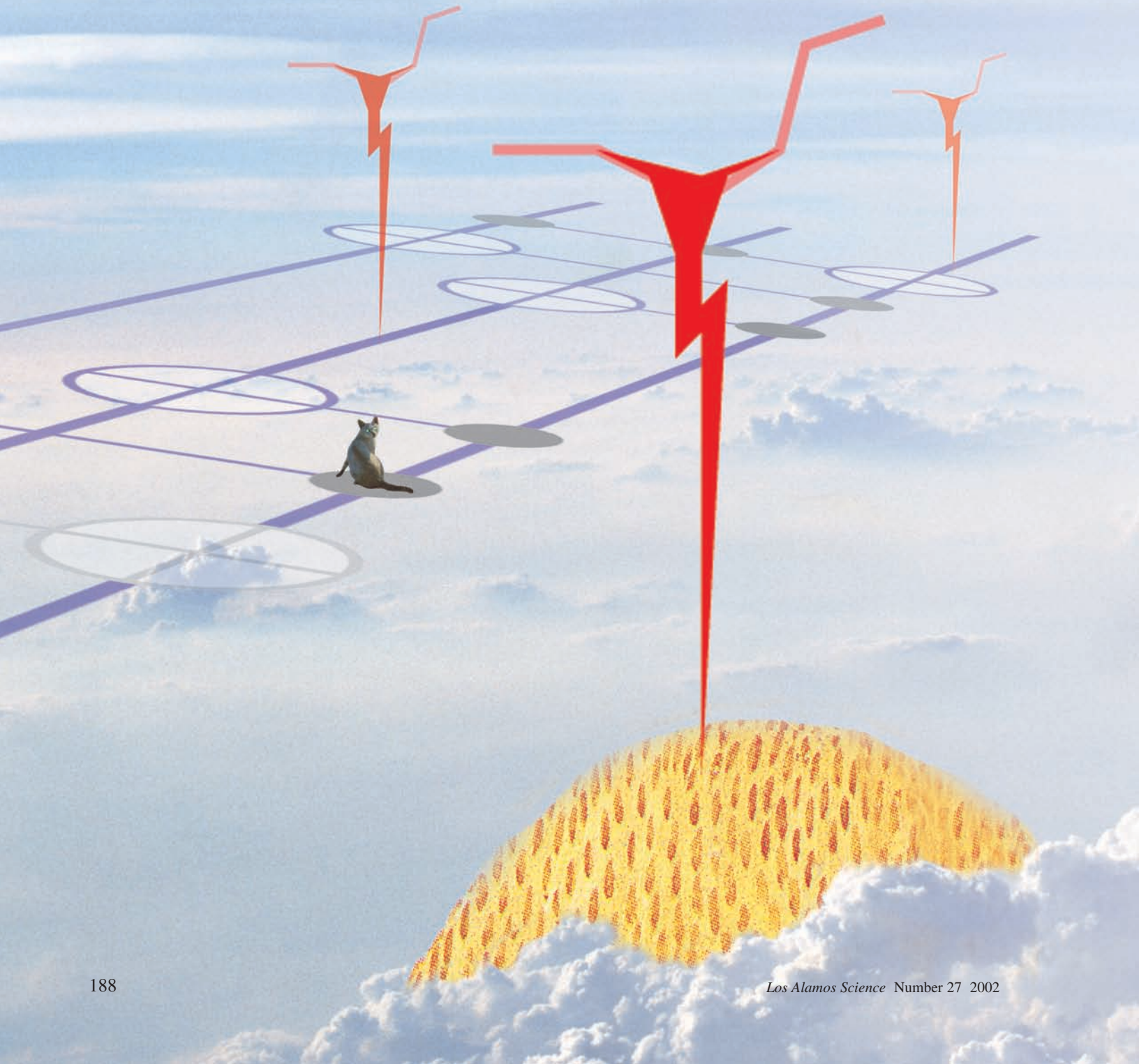
Introduction to Quantum Error Correction

Emanuel Knill, Raymond Laflamme, Alexei Ashikhmin, Howard N. Barnum, Lorenza Viola, and Wojciech H. Zurek



When physically realized, quantum information processing (QIP) can be used to solve problems in physics simulation, cryptanalysis, and secure communication for which there are no known efficient solutions based on classical information processing. Numerous proposals exist for building the devices required for QIP by using systems that exhibit quantum properties. Examples include nuclear spins in molecules, electron spins or charge in quantum dots, collective states of superconductors, and photons (Braunstein and Lo 2000). In all these cases, there are well-established physical models that, under ideal conditions, allow for exact realizations of quantum information and its manipulation. However, real physical systems never behave exactly like the ideal models. The main problems are environmental noise, which is due to incomplete isolation of the system from the rest of the world, and control errors, which are caused by calibration errors and random fluctuations in control parameters. Attempts to reduce the effects of these errors are confronted by the conflicting needs of being able to control and reliably measure the quantum systems. These needs require strong interactions with control devices and systems that are sufficiently well isolated to maintain coherence, the subtle relationship between the phases in a quantum superposition. The fact that quantum effects rarely persist on macroscopic scales suggests that meeting these needs requires considerable outside intervention.

Soon after Peter Shor published the efficient quantum factoring algorithm with its applications to breaking commonly used public-key cryptosystems, Andrew Steane (1996) and Shor (1995) gave the first constructions of quantum error-correcting codes. These codes make it possible to store quantum information so that one can reverse the effects of the most likely errors. By demonstrating that quantum information can exist in protected parts of the state space, they showed that, in principle, it is possible to protect against environmental noise when storing or transmitting information. Stimulated by these results and in order to solve errors happening during computation with quantum information, researchers initiated a series of investigations to determine whether it was possible to quantum-compute in a fault-tolerant manner. The outcome of these investigations was positive and culminated in what are now known as accuracy threshold theorems (Gottesman 1996, Calderbank et al. 1997, Calderbank et al. 1998, Shor 1996, Kitaev 1997, Knill and Laflamme 1996, Aharonov and Ben-Or 1996, Aharonov and Ben-Or 1999, Knill et al. 1998a, Knill et al. 1998b, Gottesman 1998, Preskill 1998). According to these theorems, if the effects of all errors are sufficiently small per quantum bit (qubit) and computation step, then it is possible to process quantum information arbitrarily accurately with reasonable resource overheads. The requirement on errors is quantified by a maximum tolerable error rate called the threshold. The threshold value depends strongly on the details of the assumed error model. All threshold theorems require that errors at different times and locations be independent and that the basic computational operations can be applied in parallel. Although the proven thresholds are well out of the range of today's devices, there are signs that, in practice, fault-tolerant quantum computation may be realizable.

In retrospect, advances in quantum error correction and fault-tolerant computation were made possible by the realization that accurate computation does not require the state of the physical devices supporting the computation to be perfect. In classical information processing, this observation is so obvious that it is often forgotten: No two letters "e" on a written page are physically identical, and the number of electrons used to store a bit in the computer's memory varies substantially. Nevertheless, we have no difficulty in accurately identifying the desired letter or state. A crucial conceptual difficulty with quantum information is that, by its very nature, it cannot be identified by being "looked" at. As a result, the sense in which quantum information can be

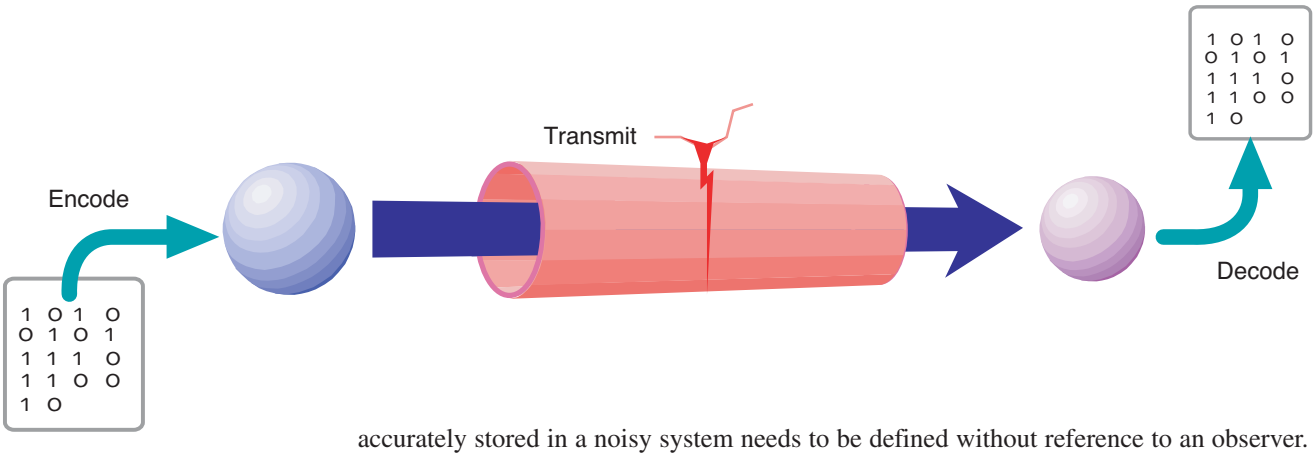


Figure 1. Typical Application of Error Correction Methods
The three main steps required for communication are shown in this figure: Information is first encoded in a physical system, then transmitted over the noisy communication channel, and finally decoded. The combination of encoding and decoding is chosen so that errors have no effect on the transmitted information.

accurately stored in a noisy system needs to be defined without reference to an observer. There are two ways to accomplish this task. The first is to define stored information to be the information that can, in principle, be extracted by a quantum decoding procedure. The second is to explicitly define "subsystems" (particle-like aspects of the quantum device) that contain the desired information. The first approach is a natural generalization of the usual interpretations of classical error-correction methods, whereas the second is motivated by a way of characterizing quantum particles.

In this article, we motivate and explain the decoding and subsystems view of quantum error correction. We explain how quantum noise in QIP can be described and classified and summarize the requirements that need to be satisfied for fault tolerance. Considering the capabilities of currently available quantum technology, the requirements appear daunting. But the idea of subsystems shows that these requirements can be met in many different, and often unexpected, ways.

Our article is structured as follows: The basic concepts are introduced by example, first for classical and then for quantum codes. We then show how the concepts are defined in general. Following a discussion of error models and analysis, we state and explain the necessary and sufficient conditions for detectability of errors and correctability of error sets. That section is followed by a brief introduction to two of the most important methods for constructing error-correcting codes and subsystems. For a basic overview, it suffices to read the beginnings of these more-technical sections. The principles of fault-tolerant quantum computation are outlined in the last section.

Concepts and Examples

Communication is the prototypical application of error correction methods. To communicate, a sender needs to convey information to a receiver over a noisy communication channel. Such a channel can be thought of as a means of transmitting an information-carrying physical system from one place to another. During transmission, the physical system is subject to disturbances that can affect the information carried. To use a communication channel, the sender needs to encode the information to be transmitted in the physical system. After transmission, the receiver decodes the information. The procedure is shown in Figure 1.

Protecting stored information is another important application of error correction methods. In this case, the user encodes the information in a storage system and retrieves it later. Provided that there is no communication from the receiver to the sender, any error correction method applicable to communication is also applicable to storage and vice versa. In a later section ("Fault-Tolerant Quantum Communication and Computation" on page 217), we discuss the problem of fault-tolerant computation,

which requires enhancing error correction methods in order to enable applying operations to encoded information without losing protection against errors.

To illustrate the different features of error correction methods, we consider three examples. We begin by describing them for classical information, but in each case, there is a quantum analogue that will be introduced later.

Trivial Two-Bit Example. Consider a physical system consisting of two bits with state space {00, 01, 10, 11}. We use the convention that state symbols for physical systems subject to errors are in gray. States changed by errors are shown in red.[1] In this example, the system is subject to errors that flip (apply the **not** operator to) the first bit with probability .5. We wish to safely store one bit of information. To this end, we store the information in the second physical bit because this bit is unaffected by the errors (see Figure 2).

As suggested by the usage examples in Figure 1, one can encode one bit of information in the physical system by the map that takes 0 → 00 and 1 → 01. This means that the states 0 and 1 of an ideal bit are represented by the states 00 and 01 of the noisy physical system, respectively.

To decode the information, one can extract the second bit by the following map:

$$\begin{aligned} 00 &\rightarrow 0 \\ 10 &\rightarrow 0 \\ 01 &\rightarrow 1 \\ 11 &\rightarrow 1 \end{aligned} \qquad (1)$$

This procedure ensures that the encoded bit is recovered by the decoding regardless of the error. There are other combinations of encoding and decoding that work. For example, in the encoding, we could swap the meaning of 0 and 1 by using the map 0 → 01 and 1 → 00. The new decoding procedure adds a bit flip to the one shown above. The only difference between this combination of encoding/decoding and the previous one lies in the way in which the information is represented in the range of the encoding. This range consists of the two states 00 and 01 and is called the code. The states in the code are called code words.

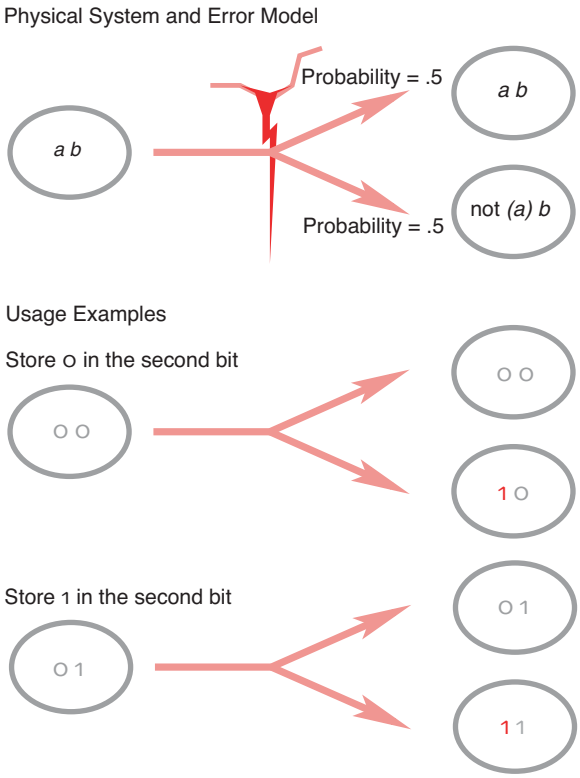


Figure 2. A Simple Error Model
Errors affect only the first bit of a physical two-bit system. All joint states of the two bits are affected by errors. For example, the joint state 00 is changed by the error to 10. Nevertheless, the value of the information represented in the second physical bit is unchanged.

Although trivial, the example just given is typical of ways for dealing with errors. That is, there is always a way of viewing the physical system as a pair of abstract systems: The first member of the pair experiences the errors, and the second carries the information to be protected. The two abstract systems are called subsystems of the physical system and are usually not identifiable with any of the system's physical components. The first is the syndrome subsystem, and the second is the information-carrying subsystem. Encoding consists of initializing the first system and storing the information in the second. Decoding is accomplished by extraction of the second system. In the example, the two subsystems are readily identified as the two physical bits that make up the physical system. The first is the syndrome subsystem and is initialized to 0 by the encoding. The second carries the encoded information.

[1] These graphical conventions are not crucial for understanding what the symbols mean and are intended for emphasis only.

The Repetition Code. The next example is a special case of the main problem of classical error correction and occurs in typical communication settings and in computer memories. Let the physical system consist of three bits. The effect of the errors is to independently flip each bit with probability p, which we take to be $p = .25$. The repetition code results from triplicating the information to be protected. An encoding is given by the map $0 \to 000$, $1 \to 111$. The repetition code is the set $\{000, 111\}$, which is the range of the encoding. The information can be decoded with majority logic: If two out of three bits are 0, the output is 0; otherwise, the output is 1.

How well does this encoding/decoding combination work for protecting one bit of information against the errors? The decoding fails to extract the bit of information correctly if two or three of the bits were flipped by the error. We can calculate the probability of incorrect decoding as follows: The probability of a given pair of bits having flipped is $.25^2 * .75$. There are three different pairs. The probability of three bits having flipped is $.25^3$. Thus, the probability of error in the encoded bit is $3 \cdot .25^2 * .75 + .25^3 = 0.15625$. This is an improvement over .25, which is the probability that the information represented in one of the three physical bits is corrupted by error.

To see that one can interpret this example by viewing the physical system as a pair of subsystems, it suffices to identify the physical system's states with the states of a suitable pair. The following shows such a subsystem identification:

$$\begin{aligned}
000 &\leftrightarrow 00 \cdot 0 \\
001 &\leftrightarrow 11 \cdot 0 \\
010 &\leftrightarrow 01 \cdot 0 \\
100 &\leftrightarrow 10 \cdot 0 \\
011 &\leftrightarrow 10 \cdot 1 \\
101 &\leftrightarrow 01 \cdot 1 \\
110 &\leftrightarrow 11 \cdot 1 \\
111 &\leftrightarrow 00 \cdot 1
\end{aligned} \tag{2}$$

The left side consists of the 8 states of the physical system, which are the possible states for the three physical bits making up the system. The right side shows the corresponding states for the subsystem pair. The syndrome subsystem is a two-bit subsystem, whose states are shown first. The syndrome subsystem's states are called syndromes. After the "·" symbol are the states of the information-carrying one-bit subsystem.

In the subsystem identification above, the repetition code consists of the two states for which the syndrome is 00. That is, the code states 000 and 111 correspond to the states $00 \cdot 0$ and $00 \cdot 1$ of the subsystem pair. For a state in this code, single-bit flips do not change the information-carrying bit, only the syndrome. For example, a bit flip of the second bit changes 000 to 010, which is identified with $01 \cdot 0$. The syndrome has changed from 00 to 01. Similarly, this error changes 111 to $101 \leftrightarrow 01 \cdot 1$. The following diagram shows these effects :

$$\begin{array}{ccc}
000 & \leftrightarrow & 00 \cdot 0 \\
\downarrow & & \\
010 & \leftrightarrow & 01 \cdot 0
\end{array}
\qquad
\begin{array}{ccc}
111 & \leftrightarrow & 00 \cdot 1 \\
\downarrow & & \\
101 & \leftrightarrow & 01 \cdot 1
\end{array} \tag{3}$$

Note that the syndrome change is the same. In general, with this subsystem identification, we can infer from the syndrome which single bit was flipped on an encoded state.

Errors usually act cumulatively over time. For the repetition code, this is a problem in the sense that it takes only a few actions of the above error model for the two- and three-bit errors to overwhelm the encoded information. One way to delay the loss of information is to decode and reencode sufficiently often. Instead of explicitly decoding and reencoding, the subsystem identification can be used directly for the same effect, namely, that of resetting the syndrome subsystem's state to 00. For example, if the state is $10 \cdot 1$, it needs to be reset to $00 \cdot 1$. Therefore, using the subsystem identification, resetting requires changing the state 011 to 111. It can be checked that, in every case, what is required is to set all bits of the physical system to the majority of the bits. After the syndrome subsystem has been reset, the information is again protected against the next one-bit error.

A Code for a Cyclic System. We next consider a physical system that does not consist of bits. This system has seven states symbolized by 0, 1, 2, 3, 4, 5, and 6. Let s_1 be the right-circular shift operator defined by $s_1(l) = l + 1$ for $0 \le l \le 5$ and $s_1(6) = 0$. Define $s_0 = \mathbb{1}$ (the identity operator),

$$s_k = \underbrace{s_1 \ldots s_1}_{k \text{ times}}, \tag{4}$$

and $s_{-k} = s_k^{-1}$ (left-circular shift by k). The model can be visualized as a pointer on a dial with seven positions, as shown in Figure 3. Suppose that the errors consist of applying s_k with probability qe^{-k^2}, where $q = 0.5641$ is chosen so that the probabilities sum to 1, that is $\sum_{k=-\infty}^{\infty} qe^{-k^2} = 1$. Thus, s_0 has probability 0.5641, and each of s_{-1} and s_1 has probability 0.2075. These are the main errors that we need to protect against. Continuous versions of this error model in the context of communication channels are known as Gaussian channels.

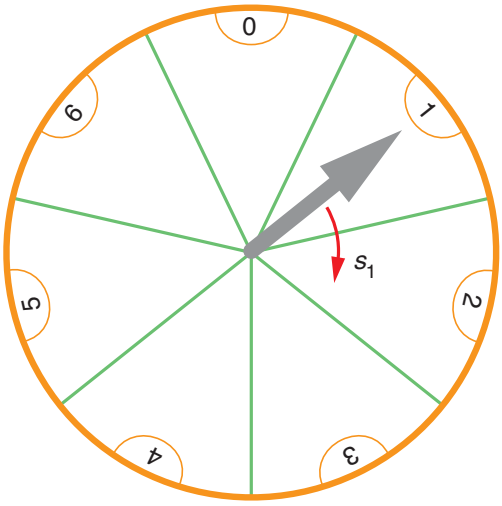


Figure 3. A Seven-State Cyclic System
The position of the pointer on the seven-position dial determines the state of the system. With the pointer in the position shown, the state is 1. Errors have the effect of rotating the pointer clockwise or counterclockwise. The effect of s_1 is to rotate the pointer clockwise, as shown by the red arrow.

One bit can be encoded in this physical system by the map $0 \to 1$, $1 \to 4$. To decode with protection against s_0, s_{-1}, and s_1, use the mapping

$$\begin{aligned} 0 &\to 0 \\ 1 &\to 0 \\ 2 &\to 0 \\ 3 &\to 1 \\ 4 &\to 1 \\ 5 &\to 1 \\ 6 &\to \text{fail} \end{aligned} \tag{5}$$

If state 6 is encountered, we know that an error involving a shift of at least 2 (left or right) occurred, but there is no reasonable way of decoding it to the state of a bit. This means that the error is detected, but we cannot correct it. Error detection can be used by the receiver to ask for information to be sent again. The probability of correctly decoding with this code is at least 0.9792, which is the probability that the error caused a shift of at most 1.

As before, a pair of syndrome and information-carrying subsystems can be identified as being used by the encoding and decoding procedures. It suffices to correctly identify the syndrome states, which we name -1, 0, and 1, because they indicate which of the likeliest shifts happened. The resulting subsystem identification is

$$\begin{aligned} 0 &\leftrightarrow -1 \cdot 0 \\ 1 &\leftrightarrow 0 \cdot 0 \\ 2 &\leftrightarrow 1 \cdot 0 \\ 3 &\leftrightarrow -1 \cdot 1 \\ 4 &\leftrightarrow 0 \cdot 1 \\ 5 &\leftrightarrow 1 \cdot 1 \end{aligned} \tag{6}$$

A new feature of this subsystem identification is that it is incomplete: Only a subset of the state space is identified. In this case, the complement can be used for error detection.

Like the repetition code, this code can be used in a setting where the errors happen repeatedly. Again, it suffices to reset the syndrome subsystem, in this case to 0, to keep the encoded information protected. After the syndrome subsystem has been reset, a subsequent s_1 or s_{-1} error affects only the syndrome.

Principles of Error Correction

When considering the problem of limiting the effects of errors in information processing, the first task is to establish the properties of the physical systems that are available for representing and computing with information. Thus, it is necessary to learn the following: the physical system to be used, in particular the structure of its state space; the available means for controlling this system; the type of information to be processed; and the nature of the errors, that is, the error model. With this information, the approaches used to correct errors in the three examples provided in the previous section involve the following:

1. Determine a code, which is a subspace of the physical system, that can represent the information to be processed.
2. (a) Identify a decoding procedure that can restore the information represented in the code after any one of the most likely errors occurred or (b) determine a pair of syndrome and information-carrying subsystems such that the code corresponds to a "base" state of the syndrome subsystem and the primary errors act only on the syndrome.
3. Analyze the error behavior of the code and subsystem.

The tasks of determining a code and identifying decoding procedures or subsystems are closely related. As a result, the following questions are at the foundation of the theory of error correction: What properties must a code satisfy so that it can be used to protect well against a given error model? How does one obtain the decoding or subsystem identification that achieves this protection? In many cases, the answers can be based on choosing a fixed set of error operators that represents well the most likely errors and then determining whether these errors can be protected against without any loss of information. Once an error set is fixed, determining whether it is correctable can be cast in terms of the idea of detectable errors. This idea works equally well for both classical and quantum information. We introduce it using classical information concepts.

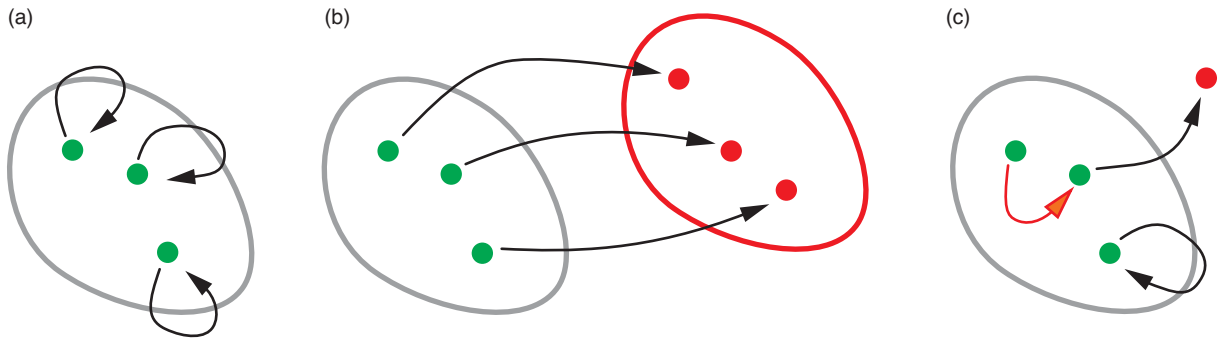


Figure 4. Typical Detectable and Undetectable Code Errors
Three examples are shown. In each, the code is represented by a brown oval containing three code words (green points). The effect of the error operator is shown as arrows. (a) The error does not change the code words and is therefore considered detectable. (b) The error maps the code words outside the code so that it is detected. (c) One code word is mapped to another, as shown by the red arrow. Finding that a received word is still in the code does not guarantee that it was the originally encoded word. The error is therefore not detectable.

Error Detection. Error detection was used in the cyclic-system example to reject a state that could not be properly decoded. In the communication setting, error control methods based on error detection alone work as follows: The encoded information is transmitted. The receiver checks whether the state is still in the code, that is, whether it could have been obtained by encoding. If not, the result is rejected. The sender can be informed of the failure so that the information can be retransmitted. Given a set of error operators that need to be protected against, the scheme is successful if, for each error operator, either the information is unchanged or the error is detected. Thus, we can say that an operator E is detectable by a code if, for each state x in the code, either $Ex = x$ or Ex is not in the code (see Figure 4).

What errors are detectable by the codes in the examples? The code in the first example consists of 00 and 01. Every operator that affects only the first bit is therefore detectable. In particular, all the operators in the error model are detectable. In the second example, the code consists of the states 000 and 111. The identity operator has no effect and is therefore detectable. Any flips of exactly one or two bits are detectable because the states in the code are changed to states outside the code. The error that flips all bits is not detectable because it preserves the code but changes the states in the code. With the code for the cyclic system, shifts by –2, –1, 0, 1, and 2 are detectable but not shifts by 3.

To conclude the section, we state a characterization of detectability, which has a natural generalization to the case of quantum information.

Theorem 1. E is detectable by a code if and only if for all $x \neq y$ in the code, $Ex \neq y$.

From Error Detection to Error Correction. Given a code C and a set of error operators $\mathcal{E} = \{\mathbb{1} = E_0, E_1, E_2 \ldots\}$, is it possible to determine whether a decoding procedure or subsystem exists such that $\mathcal{E}$ is correctable (by C), that is, such that the errors in $\mathcal{E}$ do not affect the encoded information? As explained below, the answer is yes, and the solution is to check the condition in the following theorem:

Theorem 2. $\mathcal{E}$ is correctable by C if and only if, for all $x \neq y$ in the code and all i and j, it is true that $E_i x \neq E_j y$.

Observe that the notion of correctability depends on all the errors in the set under consideration and, unlike detectability, cannot be applied to individual errors.

To see that the condition for correctability in Theorem 2 is necessary, suppose that for some $x \neq y$ in the code and some i and j, we have $z = E_i x = E_j y$. If the state z is obtained after an unknown error in $\mathcal{E}$, then it is not possible to determine whether the original code word was x or y because we cannot tell whether E_i or E_j occurred.

To see that the condition for correctability in Theorem 2 is sufficient, we assume it and construct a decoding method $z \rightarrow \mathrm{dec}(z)$. Suppose that after an unknown error

occurred, the state z is obtained. There can be one and only one x in the code for which some $E_{i(z)} \in \mathcal{E}$ satisfies the condition that $E_{i(z)}x = z$. Thus, x must be the original code word, and we can decode z by defining $x = \mathrm{dec}(z)$. Note that it is possible for two errors to have the same effect on some code words. A subsystem identification for this decoding is given by $z \leftrightarrow i(z) \cdot \mathrm{dec}(z)$, where the syndrome subsystem's state space consists of error operator indices $i(z)$ and the information-carrying system's consists of the code words $\mathrm{dec}(z)$ returned by the decoding. The subsystem identification thus constructed is not necessarily onto the state space of the subsystem pair. That is, for different code words x, the set of $i(z)$ such that $\mathrm{dec}(z) = x$ can vary and need not be all the error indices. As we will show, the subsystem identification is onto the state space of the subsystem pair in the case of quantum information. It is instructive to check that, when applied to the examples, this subsystem construction does give a version of the subsystem identifications provided earlier.

It is possible to relate the condition for correctability of an error set to detectability. For simplicity, assume that each E_i is invertible. (This assumption is satisfied by our examples but not by error operators such as "reset bit one to 0.") In this case, the correctability condition is equivalent to the statement that all products $E_j^{-1} E_i$ are detectable. To see the equivalence, first suppose that some $E_j^{-1} E_i$ is not detectable. Then, there are $x \neq y$ in the code such that $E_j^{-1} E_i\, x = y$. Consequently, $E_i x = E_j y$, and the error set is not correctable. This argument can be reversed to complete the proof of equivalence.

If the assumption that the errors are invertible does not hold, the relationship between detectability and correctability becomes more complicated, requiring a generalization of the inverse operation. This generalization is simpler in the quantum setting.

Quantum Error Correction

The principles of error correction outlined before apply to the quantum setting as readily as to the classical setting. The main difference is that the physical system to be used for representing and processing information behaves quantum mechanically and the type of information is quantum. The question of how classical information can be protected in quantum systems is also interesting but will not be discussed here. We illustrate the principles of quantum error correction by considering quantum versions of the three examples given in "Concepts and Examples" and then add a uniquely quantum example with potentially practical applications in, for example, quantum dot technologies. For an explanation of the basic quantum-information concepts and conventions, see the article "Quantum Information Processing" on page 2.

Trivial Two-Qubit Example. A quantum version of the two-bit example from the previous section consists of two physical qubits, where the errors randomly apply the identity or one of the Pauli operators to the first qubit. The Pauli operators are defined by

$$\mathbb{1} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}, \sigma_x = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \sigma_y = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}, \text{ and } \sigma_z = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}. \quad (7)$$

Explicitly, the errors have the effect

$$|\psi\rangle_{12}\begin{cases} \mathbb{1}|\psi\rangle_{12} & \text{Probability } .25 \\ \sigma_x^{(1)}|\psi\rangle_{12} & \text{Probability } .25 \\ \sigma_y^{(1)}|\psi\rangle_{12} & \text{Probability } .25 \\ \sigma_z^{(1)}|\psi\rangle_{12} & \text{Probability } .25 \end{cases}, \tag{8}$$

where the superscripts in parentheses specify the qubit that an operator acts on. This error model is called completely depolarizing on qubit 1. Obviously, a one-qubit state can be stored in the second physical qubit without being affected by the errors. An encoding operation that implements this observation is

$$|\psi\rangle \rightarrow |0\rangle_1 |\psi\rangle_2 , \tag{9}$$

which realizes an ideal qubit as a two-dimensional subspace of the physical qubits. This subspace is the quantum code for this encoding. To decode, one can discard physical qubit 1 and return qubit 2, which is considered a natural subsystem of the physical system. In this case, the identification of syndrome and information-carrying subsystems is the obvious one associated with the two physical qubits.

Quantum Repetition Code. The repetition code can be used to protect quantum information in the presence of a restricted error model. Let the physical system consist of three qubits. Errors act by independently applying, to each qubit, the flip operator σ_x with probability .25. The classical code can be made into a quantum code by the superposition principle. Encoding one qubit is accomplished by

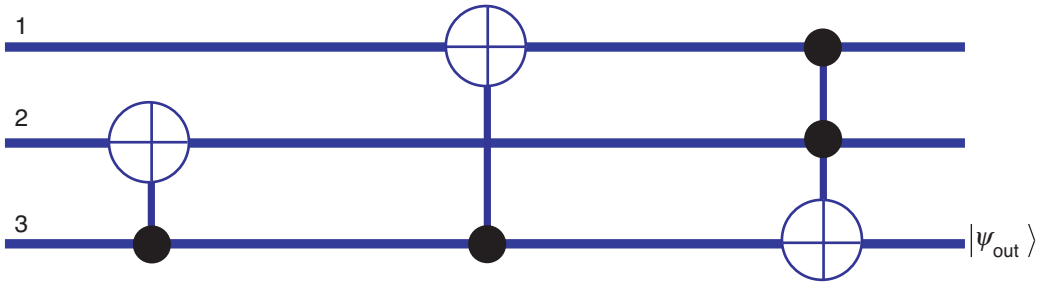
$$\alpha|0\rangle + \beta|1\rangle \rightarrow \alpha|000\rangle + \beta|111\rangle \quad . \tag{10}$$

The associated quantum code is the range of the encoding, that is, the two-dimensional subspace spanned by the encoded states $|000\rangle$ and $|111\rangle$.

As in the classical case, decoding is accomplished by majority logic. However, it must be implemented carefully to avoid destroying quantum coherence in the stored information. One way to do that is to use only unitary operations to transfer the stored information to the output qubit. Figure 5 shows a quantum network that accomplishes this task.

As shown, the decoding network establishes an identification between the three physical qubits and a pair of subsystems consisting of two qubits representing the syndrome subsystem and one qubit for the information-carrying subsystem. On the left side of the correspondence, the information-carrying subsystem is not identifiable with any one (or two) of the physical qubits. Nevertheless, it exists there through the identification.

To obtain a network for encoding, we reverse the decoding network and initialize qubits 2 and 3 in the state $|00\rangle$. The initialization renders the Toffoli gate unnecessary. The complete system with a typical error is shown in Figure 6.



$\|000\rangle$ $\|001\rangle$ $\|010\rangle$ $\|100\rangle$	$\|000\rangle$ $\|011\rangle$ $\|010\rangle$ $\|100\rangle$	$\|000\rangle$ $\|111\rangle$ $\|010\rangle$ $\|100\rangle$	$\|00\rangle\|0\rangle$ $\|11\rangle\|0\rangle$ $\|01\rangle\|0\rangle$ $\|10\rangle\|0\rangle$
$\|111\rangle$ $\|110\rangle$ $\|101\rangle$ $\|011\rangle$	$\|101\rangle$ $\|110\rangle$ $\|111\rangle$ $\|001\rangle$	$\|001\rangle$ $\|110\rangle$ $\|011\rangle$ $\|101\rangle$	$\|00\rangle\|1\rangle$ $\|11\rangle\|1\rangle$ $\|01\rangle\|1\rangle$ $\|10\rangle\|1\rangle$

Figure 5. Majority Logic Decoding into the Output Qubit 3
The effect of the quantum network on the basis states is shown. The top half shows the states with majority 0. The decoded qubit is separated in the last step. The conventions for illustrating quantum networks are explained in the article "Quantum Information Processing" on page 2.

As in the case of the classical repetition code, we can protect against cumulative errors without explicitly decoding and then reencoding, which would cause a temporary loss of protection. Instead, one can find a means for directly resetting the syndrome subsystem to $|00\rangle$ (thus returning the information to the code) before the errors happen again. After resetting in this way, the errors in the correctable set have no effect on the encoded information because they act only on the syndrome subsystem.

Part of the task of designing error-correcting systems is to determine how well the system performs. An important performance measure is the probability of error. In quantum systems, the probability of error is intuitively interpreted as the maximum probability with which we can see a result different from the expected one in any measurement. Specifically, to determine the error, one compares the output $|\psi_0\rangle$ of the system with the input $|\psi\rangle$. An upper bound is obtained if the output is written as a combination of the input state and an error state. For quantum information, combinations are linear combinations (that is, superpositions). Thus $|\psi_0\rangle = \gamma\,|\psi\rangle + |e\rangle$ (see Figure 7). The probability of error is bounded by $\varepsilon = \||e\rangle\|^2$ (which we call an error estimate). In general, there are many different ways of writing the output as a combination of an acceptable state and an error term. One attempts to choose the combination that minimizes the error estimate. This choice yields the number ε for which $1 - \varepsilon$ is called fidelity. A fidelity of 1 means that the output is the same (up to a phase factor) as the input.

To illustrate error analysis, we calculate the error for the repetition code example for the two initial states $|0\rangle$ and $(1/\sqrt{2})(|0\rangle + |1\rangle)$.

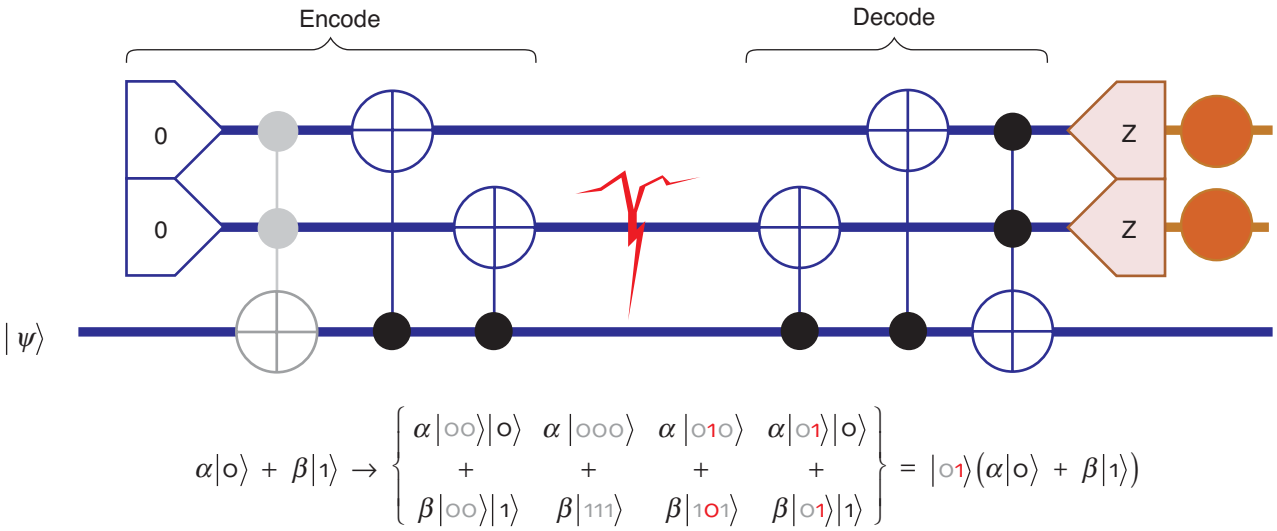


Figure 6. Networks for the Quantum Repetition Code with a Typical Error
The error that occurred can be determined from the state of the syndrome subsystem, which consists of the top two qubits. The encoding is shown as the reverse of the decoding, starting with an initialized syndrome subsystem. When the decoding is reversed to yield the encoding, there is an initial Toffoli gate (shown in gray). Because of the initialization, this gate has no effect and is therefore omitted in an implementation.

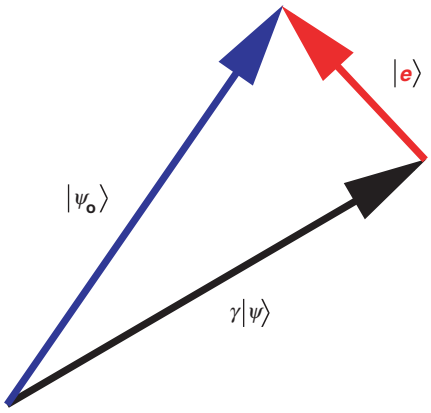


Figure 7. Error Estimate
Any decomposition of the output state $|\psi_o\rangle$ into a "good" state $\gamma|\psi\rangle$ and an (unnormalized) error term $|e\rangle$ gives an estimate $\varepsilon = \||e\rangle\|^2$. For pure states, the optimum estimate is obtained when the error term is orthogonal to the input state. To obtain an error estimate for mixtures, one can use any representation of the state as a probabilistic combination of pure states and calculate the probabilistic sum of the pure-state errors.

$$|0\rangle \xrightarrow{\text{encode}} |000\rangle \tag{11}$$

$$\longrightarrow \begin{cases} .75^3 : |000\rangle , \\ .25*.75^2 : |100\rangle , \\ .25*.75^2 : |010\rangle , \\ .25*.75^2 : |001\rangle , \\ .25^2*.75 : |110\rangle , \\ .25^2*.75 : |101\rangle , \\ .25^2*.75 : |011\rangle , \\ .25^3 : |111\rangle . \end{cases} \tag{12}$$

$$\xrightarrow{\text{decode}} \begin{cases} .4219 : |00\rangle\cdot|0\rangle , \\ .1406 : |10\rangle\cdot|0\rangle , \\ .1406 : |01\rangle\cdot|0\rangle , \\ .1406 : |11\rangle\cdot|0\rangle , \\ .0469 : |11\rangle\cdot|1\rangle , \\ .0469 : |01\rangle\cdot|1\rangle , \\ .0469 : |10\rangle\cdot|1\rangle , \\ .0156 : |00\rangle\cdot|1\rangle . \end{cases} \tag{13}$$

The final state is a mixture consisting of four correctly decoded components and four incorrectly decoded ones. The probability of each state in the mixture is shown before the colon. The incorrectly decoded information is orthogonal to the encoded information, and its probability is 0.1563, an improvement over the one-qubit error probability of 0.25. The second state behaves quite differently:

$$\frac{1}{\sqrt{2}}\left(|0\rangle+|1\rangle\right) \xrightarrow{\text{encode}} \frac{1}{\sqrt{2}}\left(|000\rangle+|111\rangle\right) \tag{14}$$

$$\longrightarrow \begin{cases} \vdots \\ .25^2*.75 : \dfrac{1}{\sqrt{2}}\left(|110\rangle+|001\rangle\right) \\ \vdots \end{cases} \tag{15}$$

$$\xrightarrow{\text{decode}} \begin{cases} \vdots \\ .0469 : \dfrac{1}{\sqrt{2}}|11\rangle\cdot\left(|1\rangle+|0\rangle\right) \\ \vdots \end{cases} . \tag{16}$$

Not all error events have been shown, but in each case it can be seen that the state is decoded correctly, so the error is 0. This shows that the error probability can depend

significantly on the initial state. To remove this dependence and give a state independent error quantity, one can use the worst-case, the average, or the entanglement error. See the section "Quantum Error Analysis" on page 209.

Quantum Code for a Cyclic System. The shift operators introduced earlier act as permutations of the seven states of the cyclic system. They can therefore be extended to unitary operators on a seven-state cyclic quantum system with logical basis $|0\rangle$, $|1\rangle$, $|2\rangle$, $|3\rangle$, $|4\rangle$, $|5\rangle$, and $|6\rangle$. The error model introduced earlier makes sense here without modification, as does the encoding. The subsystem identification now takes the six-dimensional subspace spanned by $|0\rangle$,.... $|5\rangle$ to a pair consisting of a three-state system with basis $|-1\rangle$, $|0\rangle$, $|1\rangle$ and a qubit. The identification of Equation (6) extends linearly to a unitary subsystem identification. The procedure for decoding is modified as follows: First, a measurement determines whether the state is in the six-dimensional subspace or not. If it is, the identification is used to extract the qubit. Here is an outline of what happens when the state $(1/\sqrt{2})(|0\rangle + |1\rangle)$ is encoded:

$$\frac{1}{\sqrt{2}}\left(|0\rangle+|1\rangle\right) \xrightarrow{\text{encode}} \frac{1}{\sqrt{2}}\left(|1\rangle+|4\rangle\right) \tag{17}$$

$$\longrightarrow \begin{cases} \vdots \\ .05641e^{-4} : \dfrac{1}{\sqrt{2}}\left(|3\rangle+|6\rangle\right) \\ \vdots \end{cases} \tag{18}$$

$$\xrightarrow{\text{detect}} \begin{cases} \vdots \\ .001 : \begin{cases} .5 : & \text{fail} \\ .5 : & |3\rangle \end{cases} \\ \vdots \end{cases} \tag{19}$$

$$\xrightarrow{\text{decode}} \begin{cases} \vdots \\ .0005 : \text{fail} \\ .0005 : |-1\rangle\cdot|1\rangle \\ \vdots \end{cases} \tag{20}$$

$$= \begin{cases} \vdots \\ .0005 : \text{fail} \\ .0005 : |-1\rangle\cdot\left(\dfrac{1}{2}\left(|0\rangle+|1\rangle\right)+\dfrac{1}{2}\left(-|0\rangle+|1\rangle\right)\right) . \\ \vdots \end{cases} \tag{21}$$

A "good" state was separated from the output in the case that is shown. The leftover error term has probability amplitude $.0005 * ((1/2)^2 + (1/2)^2) = .00025$, which contributes to the total error (not shown).

Three Quantum Spin-1/2 Particles. Quantum physics provides a rich source of systems with many opportunities for representing and protecting quantum information. Sometimes, it is possible to encode information in such a way that it is protected from the errors indefinitely, without intervention. An example is the trivial two-qubit system discussed before. Whenever error protection without intervention is possible, there is an information-carrying subsystem such that errors act only on the associated syndrome subsystem regardless of the current state. An information-carrying subsystem with this property is called "noiseless." A physically motivated example of a one-qubit noiseless

subsystem can be found in three spin-1/2 particles with errors due to random fluctuations in an external field.

A spin-1/2 particle's state space is spanned by two states, $|\uparrow\rangle$ and $|\downarrow\rangle$. Intuitively, these states correspond to the spin pointing "up" ($|\uparrow\rangle$) or "down" ($|\downarrow\rangle$) in some chosen reference frame. The state space is therefore the same as that of a qubit, and we can make the identifications $|\uparrow\rangle \leftrightarrow |0\rangle$ and $|\downarrow\rangle \leftrightarrow |1\rangle$. An external field causes the spin to rotate according to an evolution of the form

$$|\psi_t\rangle = e^{-i\left(u_x\sigma_x + u_y\sigma_y + u_z\sigma_z\right)t/2}|\psi\rangle \ . \tag{22}$$

The vector $\mathbf{u} = (u_x, u_y, u_z)$ characterizes the direction of the field and the strength of the spin's interaction with the field. This situation arises, for example, in nuclear magnetic resonance with spin-1/2 nuclei, where the fields are magnetic fields (see the article "NMR and Quantum Information Processing" on page 226).

Now consider the physical system composed of three spin-1/2 particles with errors acting as identical rotations of the three particles. Such errors occur if they are due to a uniform external field that fluctuates randomly in direction and strength. The evolution caused by a uniform field is given by

$$\begin{aligned}
|\psi_t\rangle_{123} &= e^{-i\left(u_x\sigma_x^{(1)} + u_y\sigma_y^{(1)} + u_z\sigma_z^{(1)}\right)t/2} e^{-i\left(u_x\sigma_x^{(2)} + u_y\sigma_y^{(2)} + u_z\sigma_z^{(2)}\right)t/2} e^{-i\left(u_x\sigma_x^{(3)} + u_y\sigma_y^{(3)} + u_z\sigma_z^{(3)}\right)t/2}|\psi\rangle_{123} \\
&= e^{-i\left(u_x\left(\sigma_x^{(1)} + \sigma_x^{(2)} + \sigma_x^{(3)}\right) + u_y\left(\sigma_y^{(1)} + \sigma_y^{(2)} + \sigma_y^{(3)}\right) + u_z\left(\sigma_z^{(1)} + \sigma_z^{(2)} + \sigma_z^{(3)}\right)\right)t/2}|\psi\rangle_{123} \\
&= e^{-i\left(u_xJ_x + u_yJ_y + u_zJ_z\right)t}|\psi\rangle_{123} \ ,
\end{aligned} \tag{23}$$

with $J_u = (\sigma_u^{(1)} + \sigma_u^{(2)} + \sigma_u^{(3)})/2$ for $u = x$, y, and z. We can exhibit the error operators arising from a uniform field in a compact form by defining $\boldsymbol{J} = (J_x, J_y, J_z)$ and $\boldsymbol{v} = (u_x, u_y, u_z)t$. Then the error operators are given by $E(\boldsymbol{v}) = e^{-i\boldsymbol{v}\cdot\boldsymbol{J}}$, where the dot product in the exponent is calculated like the standard vector dot product.

For a one-qubit noiseless subsystem, the key property of the error model is that the errors are symmetric under any permutation of the three particles. A permutation of the particles acts on the particles' state space by permuting the labels in the logical states. For example, the permutation π that swaps the first two particles acts on logical states as

$$\pi|a\rangle_1|b\rangle_2|c\rangle_3 = |a\rangle_2|b\rangle_1|c\rangle_3 = |b\rangle_1|a\rangle_2|c\rangle_3 \ . \tag{24}$$

To say that the errors are symmetric under particle permutations means that each error E satisfies $\pi^{-1}E\pi = E$, or equivalently, $E\pi = \pi E$ (E commutes with π). To see that this condition is satisfied, write

$$\begin{aligned}\pi^{-1}E(\boldsymbol{\nu})\pi &= \pi^{-1}e^{-i\boldsymbol{\nu}\cdot\boldsymbol{J}}\pi \\ &= e^{-i\pi^{-1}(\boldsymbol{\nu}\cdot\boldsymbol{J})\pi} \\ &= e^{-i\boldsymbol{\nu}\left(\pi^{-1}\boldsymbol{J}\pi\right)} \quad . \end{aligned} \tag{25}$$

If π permutes particle a to particle b, then $\pi^{-1}\sigma_u^{(a)}\pi = \sigma_u^{(b)}$. It follows that $\pi^{-1}\boldsymbol{J}\pi = \boldsymbol{J}$. This expression shows that the errors commute with the particle permutations and therefore cannot distinguish between the particles. An error model satisfying this property is called a collective error model.

If a noiseless subsystem exists, then learning the symmetries of the error model suffices for constructing the subsystem. This procedure is explained later, in "Conserved Quantities, Symmetries, and Noiseless Subsystems." For the three spin-1/2 system, the procedure results in a one-qubit noiseless subsystem protected from all collective errors. We first exhibit the subsystem identification and then discuss its properties to explain why it is noiseless. As in the case of the seven-state cyclic system, the identification involves a proper subspace of the physical system's state space. The subsystem identification involves a four-dimensional subspace and is defined by the following correspondence:

$$\begin{aligned}
&\frac{1}{\sqrt{3}}\left(|\downarrow\rangle_1|\uparrow\rangle_2|\uparrow\rangle_3 + e^{-i2\pi/3}|\uparrow\rangle_1|\downarrow\rangle_2|\uparrow\rangle_3 + e^{i2\pi/3}|\uparrow\rangle_1|\uparrow\rangle_2|\downarrow\rangle_3\right) \leftrightarrow |\uparrow\rangle\cdot|0\rangle \\
&\frac{1}{\sqrt{3}}\left(|\downarrow\rangle_1|\uparrow\rangle_2|\uparrow\rangle_3 + e^{i2\pi/3}|\uparrow\rangle_1|\downarrow\rangle_2|\uparrow\rangle_3 + e^{-i2\pi/3}|\uparrow\rangle_1|\uparrow\rangle_2|\downarrow\rangle_3\right) \leftrightarrow |\uparrow\rangle\cdot|1\rangle \\
&-\frac{1}{\sqrt{3}}\left(|\uparrow\rangle_1|\downarrow\rangle_2|\downarrow\rangle_3 + e^{-i2\pi/3}|\downarrow\rangle_1|\uparrow\rangle_2|\downarrow\rangle_3 + e^{i2\pi/3}|\downarrow\rangle_1|\downarrow\rangle_2|\uparrow\rangle_3\right) \leftrightarrow |\downarrow\rangle\cdot|0\rangle \\
&-\frac{1}{\sqrt{3}}\left(|\uparrow\rangle_1|\downarrow\rangle_2|\downarrow\rangle_3 + e^{i2\pi/3}|\downarrow\rangle_1|\uparrow\rangle_2|\downarrow\rangle_3 + e^{-i2\pi/3}|\downarrow\rangle_1|\downarrow\rangle_2|\uparrow\rangle_3\right) \leftrightarrow |\downarrow\rangle\cdot|1\rangle \quad .
\end{aligned} \tag{26}$$

The state labels for the syndrome subsystem (before the dot in the expressions on the right side) identify it as a spin-1/2 subsystem. In particular, it responds to the errors caused by uniform fields in the same way as the physical spin-1/2 particles. This behavior is caused by $2J_u$ acting as the u-Pauli operator on the syndrome subsystem. To confirm this property, we apply $2J_u$ to the logical states of Equation (26) for $u = z, x$. The property for $u = y$ then follows because $i\sigma_y = \sigma_z\sigma_x$. Consider $2J_z$. Each of the four states shown in Equation (26) is an eigenstate of $2J_z$. For example, the physical state for $|\uparrow\rangle \cdot |0\rangle$ is a superposition of states with two spins up ($\uparrow$) and one spin down ($\downarrow$). The eigenvalue of such a state with respect to $2J_z$ is the difference Δ between the number of spins that are up and down. Thus, $2J_z|\uparrow\rangle \cdot |0\rangle = |\uparrow\rangle \cdot |0\rangle$. The difference is also $\Delta = 1$ for $|\uparrow\rangle \cdot |1\rangle$ and $\Delta = -1$ for $|\downarrow\rangle \cdot |0\rangle$ and $|\downarrow\rangle \cdot |1\rangle$. Therefore, $2J_z$ acts as the z-Pauli operator on the syndrome subsystem. To confirm this behavior for $2J_x$, we compute $2J_x|\uparrow\rangle \cdot |0\rangle$.

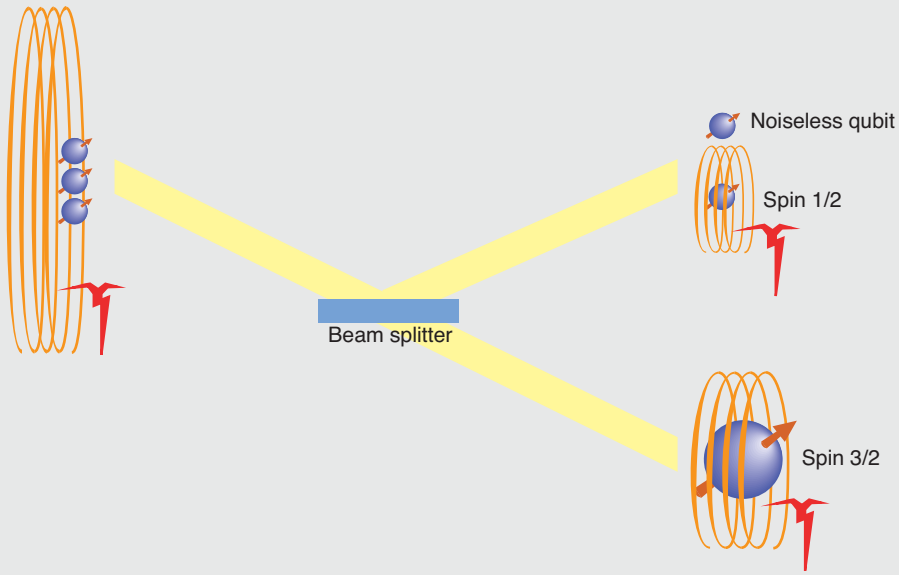
$$
\begin{aligned}
2J_x\left|\uparrow\right\rangle\cdot|o\rangle &= 2J_x\,\frac{1}{\sqrt{3}}\left(\left|\downarrow\right\rangle_1\left|\uparrow\right\rangle_2\left|\uparrow\right\rangle_3 + e^{-i2\pi/3}\left|\uparrow\right\rangle_1\left|\downarrow\right\rangle_2\left|\uparrow\right\rangle_3 + e^{i2\pi/3}\left|\uparrow\right\rangle_1\left|\uparrow\right\rangle_2\left|\downarrow\right\rangle_3\right)\\
&= \frac{1}{\sqrt{3}}\left(\sigma_x^{(1)}+\sigma_x^{(2)}+\sigma_x^{(3)}\right)\left|\downarrow\right\rangle_1\left|\uparrow\right\rangle_2\left|\uparrow\right\rangle_3\\
&\quad + e^{-i2\pi/3}\,\frac{1}{\sqrt{3}}\left(\sigma_x^{(1)}+\sigma_x^{(2)}+\sigma_x^{(3)}\right)\left|\uparrow\right\rangle_1\left|\downarrow\right\rangle_2\left|\uparrow\right\rangle_3\\
&\quad + e^{i2\pi/3}\,\frac{1}{\sqrt{3}}\left(\sigma_x^{(1)}+\sigma_x^{(2)}+\sigma_x^{(3)}\right)\left|\uparrow\right\rangle_1\left|\uparrow\right\rangle_2\left|\downarrow\right\rangle_3\\
&= \frac{1}{\sqrt{3}}\left(\left|\uparrow\right\rangle_1\left|\uparrow\right\rangle_2\left|\uparrow\right\rangle_3+\left|\downarrow\right\rangle_1\left|\downarrow\right\rangle_2\left|\uparrow\right\rangle_3+\left|\downarrow\right\rangle_1\left|\uparrow\right\rangle_2\left|\downarrow\right\rangle_3\right)\\
&\quad + e^{-i2\pi/3}\,\frac{1}{\sqrt{3}}\left(\left|\downarrow\right\rangle_1\left|\downarrow\right\rangle_2\left|\uparrow\right\rangle_3+\left|\uparrow\right\rangle_1\left|\uparrow\right\rangle_2\left|\uparrow\right\rangle_3+\left|\uparrow\right\rangle_1\left|\downarrow\right\rangle_2\left|\downarrow\right\rangle_3\right)\\
&\quad + e^{i2\pi/3}\,\frac{1}{\sqrt{3}}\left(\left|\downarrow\right\rangle_1\left|\uparrow\right\rangle_2\left|\downarrow\right\rangle_3+\left|\uparrow\right\rangle_1\left|\downarrow\right\rangle_2\left|\downarrow\right\rangle_3+\left|\uparrow\right\rangle_1\left|\uparrow\right\rangle_2\left|\uparrow\right\rangle_3\right)\\
&= \frac{1}{\sqrt{3}}\left(1+e^{-i2\pi/3}+e^{i2\pi/3}\right)\left|\uparrow\right\rangle_1\left|\uparrow\right\rangle_2\left|\uparrow\right\rangle_3\\
&\quad + \frac{1}{\sqrt{3}}\left(e^{-i2\pi/3}+e^{i2\pi/3}\right)\left|\uparrow\right\rangle_1\left|\downarrow\right\rangle_2\left|\downarrow\right\rangle_3\\
&\quad + \frac{1}{\sqrt{3}}\left(1+e^{i2\pi/3}\right)\left|\downarrow\right\rangle_1\left|\uparrow\right\rangle_2\left|\downarrow\right\rangle_3\\
&\quad + \frac{1}{\sqrt{3}}\left(1+e^{-i2\pi/3}\right)\left|\downarrow\right\rangle_1\left|\downarrow\right\rangle_2\left|\uparrow\right\rangle_3\\
&= -\frac{1}{\sqrt{3}}\left(\left|\uparrow\right\rangle_1\left|\downarrow\right\rangle_2\left|\downarrow\right\rangle_3 + e^{-i2\pi/3}\left|\downarrow\right\rangle_1\left|\uparrow\right\rangle_2\left|\downarrow\right\rangle_3 + e^{i2\pi/3}\left|\downarrow\right\rangle_1\left|\downarrow\right\rangle_2\left|\uparrow\right\rangle_3\right)\\
&= \left|\downarrow\right\rangle\cdot|o\rangle \quad .
\end{aligned}
\tag{27}
$$

Similarly, one can check that, for the other logical states, the effect of $2J_x$ is to flip the orientation of the syndrome spin. That the subsystem identified in Equation (26) is noiseless now follows from the fact that the errors $E(\boldsymbol{\nu})$ are exponentials of sums of the syndrome spin operators J_u. The errors therefore act as the identity on the information-carrying subsystem.

The noiseless qubit supported by three spin-1/2 particles with collective errors is another example in which the subsystem identification does not involve the whole state space of the system. In this case, the errors of the error model cannot remove amplitude from the subspace. As a result, if we detect an error, that is, if we find that the system's state is in the orthogonal complement of the subspace of the subsystem identification, we can deduce that either the error model is inadequate or we introduced errors in the manipulations required for transferring information to the noiseless qubit.

The noiseless subsystem of three spin-1/2 particles can be physically motivated by an analysis of quantum spin numbers. This analysis is outlined in the box on the opposite page.

Creating a Noiseless Subsystem from Three Spin-1/2 Particles



The left side shows the three particles with errors caused by fluctuations in a uniform magnetic field depicted by a noisy coil. The spin along direction u ($u = x, y, z$) can be measured, and its expectation is given by $\langle\psi|J_u|\psi\rangle$, where $|\psi\rangle$ is the quantum state of the particles and J_u is the total spin observable along the u-axis given by the half sum of the u-Pauli matrices of the particles as defined in the text. The squared magnitude of the total spin is given by the expectation of the observable $J^2 = \boldsymbol{J} \cdot \boldsymbol{J} = J_x^2 + J_y^2 + J_z^2$. The observable J^2 commutes with the J_u and therefore also with the errors $E(\boldsymbol{v}) = e^{-i\boldsymbol{v}\cdot\boldsymbol{J}}$ caused by uniform field fluctuations. This statement can be verified directly, or one can note that $E(\boldsymbol{v})$ acts on $\boldsymbol{J}$ as a rotation in three dimensions, and as one would expect, such rotations preserve the squared length J^2 of $\boldsymbol{J}$. It now follows that the eigenspaces of J^2 are invariant under the errors and, therefore, that the eigenspaces are good places to look for noiseless subsystems. The eigenvalues of J^2 are of the form $j\,(j + 1)$, where j is the spin quantum number of the corresponding eigenspace. There are two eigenspaces, one with spin $j = 1/2$ and the other with spin $j = 3/2$.

The figure shows a thought experiment that involves passing the three-particle system through a type of beam splitter or Stern-Gerlach apparatus sensitive to J^2. Using such a beam splitter, the system of particles can be made to go in one of two directions, depending on j. In the figure, if the system's state is in the spin-3/2 subspace, it passes through the beam splitter; if it is in the spin-1/2 subspace, the system is reflected up. It can be shown that the subspace with $j = 3/2$ is four dimensional and spanned by the states that are symmetric under particle permutations. Unfortunately, there is no noiseless subsystem in this subspace (refer to the section "Conserved Quantities, Symmetries, and Noiseless Subsystems"). The spin-1/2 subspace is also four dimensional and spanned by the states in Equation (26). The spin-1/2 property of the subspace implies that the spin operators J_u act in a way that is algebraically identical to the way $\sigma_u/2$ acts on a single spin-1/2 particle. This property implies the existence of the syndrome subsystem introduced in the text. Conventionally, the spin-1/2 subspace is thought of as consisting of two orthogonal two-dimensional subspaces, each behaving like a spin-1/2 with respect to the J_u. This choice of subspaces is not unique, but by associating them with two logical states of a noiseless qubit, one can obtain the subsystem identification of Equation (26). Some care needs to be taken to ensure that the noiseless qubit operators commute with the J_u, as they should. In the thought experiment shown in the figure, one can imagine unitarily rotating the system emerging in the upper path to make explicit the syndrome spin-1/2 subsystem and the noiseless qubit with which it must be paired. The result of this rotation is shown.

Error Models

We have seen several models of physical systems and errors in the examples of the previous sections. Most physical systems under consideration for QIP consist of particles or degrees of freedom that are spatially localized, a feature reflected in the error models that are usually investigated. Because we also expect the physically realized qubits to be localized, the standard error models deal with quantum errors that act independently on different qubits. Logically realized qubits, such as those implemented by subsystems different from the physically obvious ones, may have more complicated residual-error behaviors.

The Standard Error Models for Qubits. The most investigated error model for qubits consists of independent, depolarizing errors. This model has the effect of completely depolarizing each qubit independently with probability p—see Equation (8). For one qubit, the model is the least biased in the sense that it is symmetric under rotations. As a result, every state of the qubit is equally affected. Independent depolarizing errors are considered to be the quantum analogue of the classical independent bit-flip error model.

Depolarizing errors are not typical for physically realized qubits. However, given the ability to control individual qubits, it is possible to enforce the depolarizing model (see below). Consequently, error correction methods designed to control depolarizing errors apply to all independent error models. Nevertheless, it is worth keeping in mind that given detailed knowledge of the physical errors, a special purpose method is usually better than one designed for depolarizing errors. We therefore begin by showing how one can think about arbitrary error models.

There are several different ways of describing errors affecting a physical system (or "sys" for short) of interest. For most situations, in particular if the initial state of the system is pure, errors can be thought of as being the result of coupling to an initially independent environment for some time. Because of this coupling, the effect of error can always be represented by the process of adjoining an environment (or "env" for short) in some initial state $|0\rangle_{\text{env}}$ to the arbitrary state $|\psi\rangle_{\text{sys}}$, followed by a unitary coupling evolution $U^{(\text{env, sys})}$ acting jointly on the environment and the system. Symbolically, the process can be written as the map

$$|\psi\rangle_{\text{sys}} \rightarrow U^{(\text{env, sys})}|0\rangle_{\text{env}}|\psi\rangle_{\text{sys}} \ . \tag{28}$$

Choosing an arbitrary orthonormal basis consisting of the states $|e\rangle_{\text{env}}$ for the state space of the environment, the process can be rewritten in the form

$$\begin{aligned} |\psi\rangle_{\text{sys}} \rightarrow \ & \mathbb{1}^{(\text{env})} U^{(\text{env,sys})}|0\rangle_{\text{env}}|\psi\rangle_{\text{sys}} \\ = & \left(\sum_e |e\rangle_{\text{env}}\,{}^{\text{env}}\langle e|\right) U^{(\text{env,sys})}|0\rangle_{\text{env}}|\psi\rangle_{\text{sys}} \\ = & \sum_e |e\rangle_{\text{env}} \left({}^{\text{env}}\langle e|U^{(\text{env,sys})}|0\rangle_{\text{env}}\right)|\psi\rangle_{\text{sys}} \\ = & \sum_e |e\rangle_{\text{env}} A_e^{(\text{sys})}|\psi\rangle_{\text{sys}} \ , \end{aligned} \tag{29}$$

where the last step defines operators $A_e^{(\mathrm{sys})}$ acting on the physical system by $A_e^{(\mathrm{sys})} = {}^{\mathrm{env}}\langle e|U^{(\mathrm{env,\,sys})}|0\rangle_{\mathrm{env}}$. The expression $\sum_e |e\rangle_{\mathrm{env}} A_e^{(\mathrm{sys})}$ is called an environment-labeled operator. The unitarity condition implies that $\sum_e A_e^\dagger A_e = \mathbb{1}$ (with system labels omitted). The environment basis $|e\rangle_{\mathrm{env}}$ need not represent any physically meaningful choice of basis of a real environment. For error analysis, the states $|e\rangle_{\mathrm{env}}$ are formal states that label the error operators A_e. One can use an expression of the form shown in Equation (29) even when the $|e\rangle$ are not normalized or orthogonal, keeping in mind that, as a result, the identity implied by the unitarity condition changes.

Note that the state on the right side of Equation (29), representing the effect of the errors, is correlated with the environment. This means that after removing (or "tracing over") the environment, the state of the physical system is usually mixed. Instead of introducing an artificial environment, we can also describe the errors by using the density operator formalism for mixed states. Define $\rho = |\psi\rangle_{\mathrm{sys}}{}^{\mathrm{sys}}\langle\psi|$. The effect of the errors on the density matrix ρ is given by the transformation

$$\rho \to \sum_e A_e \rho A_e^\dagger \quad . \tag{30}$$

This is the "operator sum" formalism (Kraus 1983).

The two ways of writing the effects of errors can be applied to the depolarizing-error model for one qubit. As an environment-labeled operator, depolarization with probability p can be written as

$$\sqrt{1-p}\,|0\rangle_{\mathrm{env}}\mathbb{1} + \frac{\sqrt{p}}{2}\left(|1\rangle_{\mathrm{env}}\mathbb{1} + |x\rangle_{\mathrm{env}}\sigma_x + |y\rangle_{\mathrm{env}}\sigma_y + |z\rangle_{\mathrm{env}}\sigma_z\right) \quad , \tag{31}$$

where we introduced five abstract, orthonormal environment states to label the different events. In this case, one can think of the model as applying no error with probability $1 - p$ or completely depolarizing the qubit with probability p. The latter event is represented by applying one of $\mathbb{1}$, σ_x, σ_y, or σ_z with equal probability $p/4$. To be able to think of the model as randomly applied Pauli matrices, it is crucial that the environment states labeling the different Pauli matrices be orthogonal. The square roots of the probabilities appear in the operator because, in an environment-labeled operator, it is necessary to give quantum amplitudes. Environment-labeled operators are useful primarily because of their great flexibility and redundancy.

In the operator sum formalism, depolarization with probability p transforms the input density matrix ρ as

$$\begin{aligned} \rho &\to (1-p)\rho + \frac{p}{4}\left(\mathbb{1}\rho\mathbb{1} + \sigma_x\rho\sigma_x + \sigma_y\rho\sigma_y + \sigma_z\rho\sigma_z\right) \\ &= (1-3p/4)\rho + \frac{p}{4}\left(\sigma_x\rho\sigma_x + \sigma_y\rho\sigma_y + \sigma_z\rho\sigma_z\right) . \end{aligned} \tag{32}$$

Because the operator sum formalism has less redundancy, it is easier to tell when two error effects are equivalent.

In the remainder of this section, we discuss how one can use active intervention to simplify the error model. To realize this simplification, we intentionally randomize the

qubit so that the environment cannot distinguish between the different axes defined by the Pauli spin matrices. Here is a simple randomization that actively converts an arbitrary error model for a qubit into one that consists of randomly applying Pauli operators according to some distribution. The distribution is not necessarily uniform, so the new error model is not yet depolarizing. Before the errors act, apply a random Pauli operator σ_u ($u = 0, x, y, z$, $\sigma_0 = \mathbb{1}$). After the errors act, apply the inverse of that operator, $\sigma_u^{-1} = \sigma_u$; then "forget" which operator was applied. This randomization method is called twirling (Bennett et al. 1996). To understand twirling, we use environment-labeled operators to demonstrate some of the techniques useful in this context. The sequence of actions implementing twirling can be written as follows (omitting labels for the physical system):

$$
\begin{aligned}
|\psi\rangle &\rightarrow \frac{1}{2}\sum_u |u\rangle_C \sigma_u |\psi\rangle && \text{Apply a random } \sigma_u \text{ remembering } u \text{ with the help of the system } C.\\
&\rightarrow \sum_e |e\rangle_{env} \frac{1}{2}\sum_u |u\rangle_C A_e \sigma_u |\psi\rangle && \text{Errors act.}\\
&\rightarrow \sum_e |e\rangle_{env} \frac{1}{2}\sum_u |u\rangle_C \sigma_u A_e \sigma_u |\psi\rangle && \text{Apply } \sigma_u = \sigma_u^{-1}.\\
&\rightarrow \sum_{eu} |eu\rangle_{env,C} \frac{1}{2}\sigma_u A_e \sigma_u |\psi\rangle \quad . && \text{Forget which } u \text{ was used by absorbing its memory in the environment.}
\end{aligned}
\tag{33}
$$

The system C that was artificially introduced to carry the memory of u may be a classical memory because there is no need for coherence between different $|u\rangle_C$.

To determine the equivalent random Pauli operator error model, it is necessary to rewrite the total effect of the procedure using an environment-labeled sum involving orthogonal environment states and Pauli operators. To do so, express A_e as a sum of the Pauli operators, $A_e = \sum_v \alpha_{ev} \sigma_v$, using the fact that the σ_v are a linear basis for the space of one-qubit operators. Recall that σ_u anticommutes with σ_v if $0 \neq u \neq v \neq 0$. Thus, $\sigma_u \sigma_v \sigma_u = (-1)^{\langle v,u\rangle}\sigma_v$, where $\langle v, u\rangle = 1$ if $0 \neq u \neq v \neq 0$ and $\langle v, u\rangle = 0$ otherwise. We can now rewrite the last expression of Equation (33) as follows:

$$
\begin{aligned}
\sum_{eu} |eu\rangle_{env,C} \frac{1}{2}\sigma_u A_e \sigma_u |\psi\rangle &= \sum_{eu} |eu\rangle_{env,C} \frac{1}{2}\sigma_u \sum_v \alpha_{ev}\sigma_v \sigma_u |\psi\rangle\\
&= \sum_v \left(\sum_{eu} \frac{1}{2}\alpha_{ev}(-1)^{\langle v,u\rangle} |eu\rangle_{env,C} \right) \sigma_v |\psi\rangle \quad .
\end{aligned}
\tag{34}
$$

It can be checked that the states $(1/2)\sum_u (-1)^{\langle v,u\rangle}|eu\rangle_{env,C}$ are orthonormal for different e and v. As a result, the states $\sum_{eu}(1/2)\alpha_{ev}(-1)^{\langle v,u\rangle}|eu\rangle_{env,C}$ are orthogonal for different v and have probability (square norm) given by $p_v = \sum_e |\alpha_{ev}|^2$. Introducing $\sqrt{p_v}|\tilde{v}\rangle_{env,C} = \sum_{eu}(1/2)\alpha_{ev}(-1)^{\langle v,u\rangle}|eu\rangle_{env,C}$, we can write the sum of Equation (34) as

$$
\sum_v \left(\sum_{eu} \frac{1}{2}\alpha_{ev}(-1)^{\langle v,u\rangle} |eu\rangle_{env,C} \right) \sigma_v |\psi\rangle = \sum_v \sqrt{p_v} |\tilde{v}\rangle_{env,C} \sigma_v |\psi\rangle \quad , \tag{35}
$$

showing that the twirled error model behaves like randomly applied Pauli matrices with σ_v applied with probability p_v. It is a recommended exercise to reproduce the above argument using the operator sum formalism.

To obtain the standard depolarizing error model with equal probabilities for the Pauli matrices, it is necessary to strengthen the randomization procedure by applying a random member U of the group generated by the 90° rotations around the x-, y-, and z-axis before the error and then undoing U by applying U^{-1}.

Randomization can be used to transform any one-qubit error model into the depolarizing error model. This explains why the depolarizing model is so useful for analyzing error correction techniques in situations in which errors act independently on different qubits. However, in many physical situations, the independence assumptions are not satisfied. For example, errors from common internal couplings between qubits are generally pairwise correlated to first order. In addition, the operations required to manipulate the qubits and to control the encoded information act on pairs at a time, which tends to spread even single-qubit errors. Still, in all these cases, the primary error processes are local. This means that there usually exists an environment-labeled sum expression for the total error process in which the amplitudes associated with errors acting simultaneously at k locations in time and space decrease exponentially with k. In such cases, error correction methods that handle all or most errors involving sufficiently few qubits are still applicable.

Quantum Error Analysis. One of the most important consequences of the subsystems interpretation of encoding quantum information in a physical system is that the encoded quantum information can be error-free even though errors have severely changed the state of the physical system. Almost trivially, any error operator acting only on the syndrome subsystem has no effect on the quantum information. The goal of error correction is to actively intervene and maintain the syndrome subsystem in states where the dominant error operators continue to have little effect on the information of interest. An important issue in analyzing error correction methods is to estimate the residual error in the encoded information. A simple example of how that can be done was discussed for the quantum repetition code. The same ideas can be applied in general. Let sys be the physical system in which the information is encoded, and $|\psi\rangle_{\text{sys}}$ an initial state containing such information with the syndrome subsystem appropriately prepared. Errors and error-correcting operations modify the state. The new state can be expressed with environment labeling as $\sum_e |e\rangle_{\text{env}} A_e^{(\text{sys})} |\psi\rangle_{\text{sys}}$. In view of the partitioning into information-carrying and syndrome subsystems, good states $|e\rangle_{\text{env}}$ are those states for which $A_e^{(\text{sys})}$ acts only on the syndrome subsystem, given that the syndrome has been prepared. The remaining states $|e\rangle$ form the set of bad states, B. The error probability p_e can be bounded from above by

$$
\begin{aligned}
p_e &\le \left| \sum_{e \in B} |e\rangle_{\text{env}} A_e^{(\text{sys})} |\psi\rangle_{\text{sys}} \right|^2 \\
&\le \left(\sum_{e \in B} \left| |e\rangle_{\text{env}} \left| A_e^{(\text{sys})} \right|_1 \right| \right)^2 ,
\end{aligned}
\tag{36}
$$

where $|A|_1 = \max_\phi \langle\phi|A|\phi\rangle$, the maximum being taken over normalized states. The second inequality usually leads to a gross overestimate but is independent of the encoded information and often suffices for obtaining good results. Because the environment-labeled

sum is not unique, a goal of the representation of the errors acting on the system is to use "good" operators to the largest extent possible. The flexibility of these error expansions makes them very useful for analyzing error models in conjunction with error correction methods.

In principle, we can obtain better expressions for p_e by calculating the density matrix ρ of the state of the subsystem containing the desired quantum information. This calculation involves tracing over the syndrome subsystem. The matrix ρ can then be compared to the intended state. If the intended state is pure, given by $|\phi\rangle$, the probability of error is given by $1 - \langle\phi|\rho|\phi\rangle$, which is the probability that a measurement that distinguishes between $|\phi\rangle$ and its orthogonal complement fails to detect $|\phi\rangle$. The quantity $\langle\phi|\rho|\phi\rangle$ is called the fidelity of the state ρ.

For applications to communication, the goal is to be able to reliably transmit arbitrary states through a communication channel, which may be physical or realized via an encoding/decoding scheme. It is therefore important to characterize the reliability of the channel independent of the information transmitted. Equation (36) can be used to obtain state-independent bounds on the error probability but does not readily provide a single measure of reliability. One way to quantify the reliability is to identify the error of the channel with the average error ε_a over all possible input states. The reliability is then given by the average fidelity $1 - \varepsilon_a$. Another elegant way appropriate for QIP is to use the entanglement fidelity (Schumacher 1996). Entanglement fidelity measures the error when the input is maximally entangled with an identical reference system. In this process, the reference system is imagined to be untouched, so that the state of the reference system, together with the output state, can be compared with the original entangled state. For a one-qubit channel labeled sys, the reference system is a qubit, which we label "ref." An initial, maximally entangled state is

$$|B\rangle = \frac{1}{\sqrt{2}}\left(|0\rangle_{\text{ref}}|0\rangle_{\text{sys}} + |1\rangle_{\text{ref}}|1\rangle_{\text{sys}}\right) \quad . \tag{37}$$

The reference qubit is assumed to be perfectly isolated and not affected by any errors. The final state $\rho^{(\text{ref,sys})}$ is compared with $|B\rangle$, which gives the entanglement fidelity according to the formula $f_e = \langle B|\rho^{(\text{ref,sys})}|B\rangle$. The entanglement error is $\varepsilon_e = 1 - f_e$. It turns out that this definition does not depend on the choice of maximally entangled state. Fortunately, the entanglement error and the average error ε_a are related by a linear expression:

$$\varepsilon_a = \frac{2}{3}\varepsilon_e \quad . \tag{38}$$

For k-qubit channels, the constant 2/3 is replaced by $2^k/(2^k + 1)$. Experimental measurements of these fidelities do not require the reference system. There are simple averaging formulas to express them in terms of the fidelities for transmitting each of a sufficiently large set of pure states. An example of the experimental determination of the entanglement fidelity when the channel is realized by error correction is provided in Knill et al. (2001).

From Quantum Error Detection to Error Correction

In the independent depolarizing error model with small probability p of depolarization, the most likely errors are those that affect a small number of qubits. That is, if we define the weight of a product of Pauli operators to be the number of qubits affected, the dominant errors are those of small weight. Because the probability of a nonidentity Pauli operator is $3p/4$—see Equation (31)—one expects about $(3p/4)n$ of n qubits to be changed. As a result, good error-correcting codes are considered to be those for which all errors of weight $\leq e \cong (3p/4)n$ can be corrected. It is desirable that e have a high rate, which means that it is a large fraction of the total number of qubits n (the length of the code). Combinatorially, good codes are characterized by a high minimum distance, a concept that arises naturally in the context of error detection.

Quantum Error Detection. Let C be a quantum code, that is, a subspace of the state space of a quantum system. Let P be the operator that projects onto C, and $P^{\perp} = \mathbb{1} - P$ the one that projects onto the orthogonal complement. Then the pair P, $P^{\perp}$ is associated with a measurement that can be used to determine whether a state is in the code or not. If the given state is $|\psi\rangle$, the result of the measurement is $P|\psi\rangle$ with probability $|P|\psi\rangle|^2$ and $P^{\perp}|\psi\rangle$ otherwise. As in the classical case, an error-detection scheme consists of preparing the desired state $|\psi_i\rangle \in C$, transmitting it through, say, a quantum channel, then measuring whether the state is still in the code, accepting the state if it is, and rejecting it otherwise. We say that C detects error operator E if states accepted after E had acted are unchanged except for an overall scale. Using the projection operators, this is the statement that for every state $|\psi_i\rangle \in C$, $PE|\psi_i\rangle = \lambda_E |\psi_i\rangle$. Because $P|\psi\rangle$ is in the code for every $|\psi\rangle$, it follows that $PEP|\psi\rangle = \lambda_E P|\psi\rangle$. It follows that a characterization of detectability is given by Theorem 3.

Theorem 3. E is detectable by C if and only if $PEP = \lambda_E P$ for some λ_E.

A second characterization is given by Theorem 4.

Theorem 4. E is detectable by C if and only if for all $|\psi\rangle, |\phi\rangle \in C$, $\langle\psi|E|\phi\rangle = \lambda_E \langle\psi|\phi\rangle$ for some λ_E.

A third characterization is obtained by taking the condition for classical detectability in Theorem 1 and replacing $\neq$ by orthogonal to:

Theorem 5. E is detectable by C if and only if for all $|\phi\rangle, |\psi\rangle$ in the code with $|\phi\rangle$ orthogonal to $|\psi\rangle$, $E|\phi\rangle$ is orthogonal to $|\psi\rangle$.

For a given code C, the set of detectable errors is closed under linear combinations. That is, if E_1 and E_2 are both detectable, then so is $\alpha E_1 + \alpha E_2$. This useful property implies that, to check detectability, one has to consider only the elements of a linear basis for the space of errors of interest.

Consider n-qubits with independent depolarizing errors. A robust error-detecting code should detect as many of the small-weight errors as possible. This requirement motivates the definition of minimum distance: The code C has minimum distance d if the smallest-weight product of Pauli operators E for which C does not detect E is d. The notion comes from classical codes for bits, where a set of code words C' has minimum distance d if the

smallest number of flips required to change one code word in C' into another one in C' is d. For example, the repetition code for three bits has minimum distance 3. Note that the minimum distance for the quantum repetition code is 1: Applying $\sigma_z^{(1)}$ preserves the code and changes the sign of $|111\rangle$ but not of $|000\rangle$. As a result, $\sigma_z^{(1)}$ is not detectable. The notion of minimum distance can be generalized for error models with specified first-order error operators (Knill et al. 2000). In the case of depolarizing errors, the first-order error operators are single-qubit Pauli matrices, which are the errors of weight 1.

Quantum Error Correction. Let $\mathcal{E} = \{E_0 = \mathbb{1}, E_1, \ldots\}$ be the set of errors that we wish to be able to correct. When a decoding procedure for the code C exists such that all errors in $\mathcal{E}$ are corrected, we say that $\mathcal{E}$ is correctable (by C). A situation in which correctability of $\mathcal{E}$ is apparent occurs when the errors E_i are unitary operators satisfying the condition that $E_i C$ are mutually orthogonal subspaces. The repetition code has this property for the set of errors consisting of the identity and Pauli operators acting on a single qubit. In this situation, the procedure for decoding is to first make a projective measurement and determine which of the subspaces $E_i C$ the state is in and then to apply the inverse of the error operator, that is, $E_i^\dagger$. This situation is not far from the generic one. One characterization of correctability is described in Theorem 6.

Theorem 6. $\mathcal{E}$ is correctable if and only if there is a linear transformation of the set $\mathcal{E}$ such that the operators E_i' in the new set satisfy the following properties: (1) The $E_i' C$ are mutually orthogonal, and (2) E_i' restricted to C is proportional to a restriction to C of a unitary operator.

To relate this characterization to detectability, note that the two properties imply that (*E*$_i'$) $E_j' C$ is orthogonal to C if $i \neq j$ and $(E_i')^\dagger E_i'$ restricted to C is proportional to the identity on C. In other words, the $(E_i')^\dagger E_j'$ are detectable. This detectability condition applied to the original error set constitutes a second characterization of correctability, as given in Theorem 7.

Theorem 7. $\mathcal{E}$ is correctable if and only if the operators in the set $\mathcal{E}^\dagger \mathcal{E} = \{E_1^\dagger E_2 : E_i \in \mathcal{E}\}$ are detectable.

Before explaining the characterizations of correctability, we consider the situation of n qubits, where the characterization by detectability (Theorem 7) leads to a useful relationship between minimum distance and correctability of low-weight errors.

Theorem 8. If a code on n qubits has a minimum distance of at least $2e + 1$, then the set of errors of weight at most e is correctable.

This theorem follows by observing that the weight of $E_1^\dagger E_2$ is at most the sum of the weights of the E_i. As a result of this observation, the problem of finding good ways to correct all errors up to a maximum weight reduces to that of constructing codes with sufficiently high minimum distance. Thus, questions such as "what is the maximum dimension of a code of minimum distance d on n qubits?" are of great interest. As in the case of classical coding theory, this problem appears to be very difficult in general. Answers are known for small n (Calderbank et al. 1998), and there are asymptotic bounds (Ashikhmin and Litsyn 1999). Of course, for achieving low error probabilities, it is not necessary to correct all errors of weight $\leq e$, just almost all such errors. For example, the concatenated codes used for fault-tolerant quantum computation achieve this goal (see "Fault-Tolerant Quantum Communication and Computation" later in this article).

For the remainder of this section, we explain the characterizations of correctability. Using the conditions for detectability from the previous section, the condition for correctability in Theorem 7 is equivalent to

$$PE_i^\dagger E_j P = \lambda_{ij} P \ . \tag{39}$$

This condition is preserved under a linear change of basis for $\mathcal{E}$. That is, if A is any invertible matrix with coefficients a_{ij}, we can define new error operators $D_k = \sum_i E_i a_{ik}$. For the D_k, the left side of Equation (39) is

$$\begin{aligned} PD_k^\dagger D_l P &= P\left(\sum_{ij} \bar{a}_{ik} E_i^\dagger E_j a_{jl}\right) P \\ &= \sum_{ij} \bar{a}_{ik} a_{jl} \lambda_{ij} P \\ &= \left(A^\dagger \Lambda A\right)_{kl} P \ , \end{aligned} \tag{40}$$

where Λ is the matrix formed from the λ_{ij}. Using the fact that Λ is a positive semidefinite matrix (that is, for all x, $x^\dagger \Lambda x \geq 0$, and $\Lambda^\dagger = \Lambda$), we can choose A such that $A^\dagger \Lambda A$ is of the form $\begin{pmatrix} \mathbb{1} & 0 \\ 0 & 0 \end{pmatrix}$. In this matrix, the upper left block is the identity operator for some dimension.

An important consequence of invariance under a change of basis of error operators is that the set of errors correctable by a particular code and decoding procedure is linearly closed. Thus, if E and D are corrected by the decoding procedure, then so is $\alpha E + \beta D$. This observation also follows from the linearity of quantum mechanically implementable operations.

We explain the condition for correctability by using the subsystems interpretation of decoding procedures. For simplicity, assume that $\mathbb{1} \in \mathcal{E}$. To show that correctability of $\mathcal{E}$ implies detectability of all $E \in \mathcal{E}^\dagger \mathcal{E}$, suppose that we have a decoding procedure that recovers the information encoded in C after any of the errors in the set $\mathcal{E}$ have occurred. Every physically realizable decoding procedure can be implemented by first adding ancilla quantum systems in a prepared pure state to form a total system labeled T, then applying a unitary map U to the state of T, and finally separating T into a pair of systems (syn, Q), where "syn" corresponds to the syndrome subsystem and Q is a quantum system with the same dimension as the code that carries the quantum information after decoding. Denote the state space of the physical system containing C as $\mathcal{H}$ and the state space of system X by $\mathcal{H}_X$, where X is any one of the other systems. Let V be the unitary operator that encodes information by mapping $\mathcal{H}_Q$ onto $C \subseteq \mathcal{H}$. We have the following relationships:

$$\mathcal{H}_Q \overset{V}{\leftrightarrow} C \subseteq \mathcal{H} \subseteq \mathcal{H}_T \overset{U}{\leftrightarrow} \mathcal{H}_{syn} \otimes \mathcal{H}_Q \ . \tag{41}$$

Here, we used bidirectional arrows to emphasize that the operators V and U can be inverted on their range and therefore identify the states in their domains with the states in their ranges. The inclusion $\mathcal{H} \subseteq \mathcal{H}_T$ implicitly identifies $\mathcal{H}$ with the subspace determined by the prepared pure state on the ancillas. The last state space of Equation (41) is expressed as a tensor product, which is the state space of the combined system (syn, Q).

For states of $\mathcal{H}_Q$, we will write $|\psi\rangle = |\psi\rangle_Q \overset{V}{\leftrightarrow} |\psi\rangle_L \in C$. Because $\mathbb{1}$ is a correctable error, it must be the case that $|\psi\rangle_L \overset{U}{\leftrightarrow} |0\rangle_{syn}|\psi\rangle \in \mathcal{H}_{syn} \otimes \mathcal{H}_Q$ for some state $|0\rangle_{syn}$. To establish this fact, use the linearity of the maps. In general,

$$\begin{aligned} |\psi\rangle_L &\to E_i|\psi\rangle_L \\ &\overset{U}{\leftrightarrow} |i\rangle_{syn}|\psi\rangle \quad . \end{aligned} \tag{42}$$

The $|i\rangle_{syn}$ need not be normalized or orthogonal. Let F be the subspace spanned by the $|i\rangle_{syn}$. Then U induces an identification of $F \otimes \mathcal{H}_Q$ with a subspace $\overline{C} \subseteq \mathcal{H}$. This is the desired subsystem identification. We can then see how the errors act in this identification.

$$\begin{aligned} |\psi\rangle_L \quad &\leftrightarrow |0\rangle_{syn}|\psi\rangle \\ \downarrow \quad & \\ E_i|\psi\rangle_L &\leftrightarrow |i\rangle_{syn}|\psi\rangle \quad . \end{aligned} \tag{43}$$

This means that for all $|\psi\rangle$ and $|\phi\rangle$,

$${}^{L}\langle\psi|E_j^\dagger E_i|\phi\rangle_L = {}^{syn}\langle j|i\rangle_{syn}\langle\psi|\phi\rangle \quad , \tag{44}$$

that is, all errors in $\mathcal{E}^\dagger\mathcal{E}$ are detectable.

Now, suppose that all errors in $\mathcal{E}^\dagger\mathcal{E}$ are detectable. To see that correctability of $\mathcal{E}$ follows, choose a basis for the errors so that $\lambda_{ij} = \delta_{ij}\lambda_i$ with $\lambda_i = 1$ for $i < s$ and $\lambda_i = 0$ otherwise. Define a subsystem identification by

$$|i\rangle_{sys}|\psi\rangle \underrightarrow{W} E_i|\psi\rangle_L \quad , \tag{45}$$

for $0 \le i < s$. By assumption and construction, ${}^{L}\langle\psi|E_j^\dagger E_i|\psi\rangle_L = \delta_{ij}$, which implies that W is unitary (after linear extension), and so this is a proper identification. For $i \ge s$, $E_i|\psi\rangle_L = 0$, which implies that for states in the code, these errors have probability 0. Therefore, the identification can be used to successfully correct $\mathcal{E}$.

Constructing Codes

Stabilizer Codes. Most useful quantum codes are based on stabilizer constructions (Gottesman 1996, Calderbank et al. 1997). Stabilizer codes are useful because they make it easy to determine which Pauli-product errors are detectable and because they can be interpreted as special types of classical, linear codes. The latter feature makes it possible to use well-established techniques from the theory of classical error-correcting codes to construct good quantum codes.

A stabilizer code of length n for k-qubits (abbreviated as an $[[n, k]]$ code), is a 2^k-dimensional subspace of the state space of n-qubits that is characterized by the set of

products of Pauli operators that leave each state in the code invariant. Such Pauli operators are said to stabilize the code. A simple example of a stabilizer code is the quantum repetition code introduced earlier. The code's states $\alpha|000\rangle + \beta|111\rangle$ are exactly the states that are unchanged after applying $\sigma_z^{(1)}\,\sigma_z^{(2)}$ or $\sigma_z^{(1)}\,\sigma_z^{(3)}$. To simplify the notation, we write $I = \mathbb{1}$, $X = \sigma_x$, $Y = \sigma_y$, and $Z = \sigma_z$. A product of Pauli operators can then be written as $ZIXI = \sigma_z^{(1)}\,\sigma_x^{(3)}$ (as an example of length 4) with the ordering determining which qubit is being acted upon by the operators in the product.

We can understand the properties of stabilizer codes by working out the example of the quantum repetition code with the stabilizer formalism. A stabilizer of the code is $S = \{ZZI, ZIZ\}$. Let $\overline{S}$ be the set of Pauli products that are expressible up to a phase as products of elements of S. For the repetition code, $\overline{S} = \{III, ZZI, ZIZ, IZZ\}$. $\overline{S}$ consists of all Pauli products that stabilize the code. The crucial property of S is that its operators commute, that is, for $A, B \in S$, $AB = BA$. According to results from linear algebra, it follows that the state space $\mathcal{H}$ can be decomposed into orthogonal subspaces $\mathcal{H}_\lambda$ such that for $A \in S$ and $|\psi\rangle \in \mathcal{H}_\lambda$, $A|\psi\rangle = \lambda(A)|\psi\rangle$. The $\mathcal{H}_\lambda$ are the common eigenspaces of S. The stabilizer code C defined by S is the subspace stabilized by the operators in S, which means that it is given by $\mathcal{H}_\lambda$ with $\lambda(A) = 1$. The subspaces for other $\lambda(A)$ have equivalent properties and are often included in the set of stabilizer codes. For the repetition code, the stabilized subspace is spanned by the logical basis $|000\rangle$ and $|111\rangle$. From the point of view of stabilizers, there are two ways in which a Pauli product B can be detectable: (1) if $B \in \overline{S}$ because, in this case, B acts as the identity on the code and (2) if B anticommutes with at least one member (say A) of S. To see that this statement is correct, let $|\psi\rangle$ be in the code. Then $A(B|\psi\rangle) = (AB)|\psi\rangle = -(BA)|\psi\rangle = -B(A|\psi\rangle) = -B|\psi\rangle$. Thus, $B|\psi\rangle$ belongs to $\mathcal{H}_\lambda$ with $\lambda(A) = -1$. Because this subspace is orthogonal to $C = \mathcal{H}_1$, B is detectable. We define the set of Pauli products that commute with all members of S as $\overline{S}^\perp$. Thus, B is detectable if either $B \notin \overline{S}^\perp$ or $B \in \overline{S}$. Note that because $\overline{S}$ consists of commuting operators, $\overline{S} \subseteq \overline{S}^\perp$.

To construct a stabilizer code that can correct all errors of weight at most one (a quantum one-error-correcting code), it suffices to find S with the minimum weight of nonidentity members of $\overline{S}^\perp$ being at least three ($3 = 2 \cdot 1 + 1$)—also refer to Theorem 8. In this case, we say that $\overline{S}^\perp$ has minimum distance 3. As an example, we can exhibit a stabilizer for the famous length-five one-error-correcting code for one qubit (Bennett et al. 1996, Laflamme et al. 1996):

$$S = \{X\,Z\,Z\,X\,I,\ I\,X\,Z\,Z\,X,\ X\,I\,X\,Z\,Z,\ Z\,X\,I\,X\,Z\} \quad . \tag{46}$$

As a general rule, it is desirable to exhibit the stabilizer minimally, which means that no member is the product up to a phase of some of the other members. In this case, the number of qubits encoded is $n - |S|$, where n is the length of the code and $|S|$ is the number of elements of S.

To obtain the correspondence between stabilizer codes and classical binary codes, we replace the symbols I, X, Y, and Z in a Pauli product by 00, 01, 10, and 11, respectively. Thus, the members of the stabilizer can be thought of as binary vectors of length $2n$. We use arithmetic modulo 2 for sums, inner products, and application of a binary matrix. Because the numbers modulo 2 ($\mathbb{Z}_2$) form a mathematical field, the basic properties of vector spaces and linear algebra apply to binary vectors and matrices. Thus, the stabilizer is minimal in the sense introduced above if the corresponding binary vectors are independent over $\mathbb{Z}_2$. Given two binary (column) vectors x and y of length 2 associated with Pauli products, the property of anticommuting is equivalent to $x^T B y = 1$, where B is the block diagonal $2n \times 2n$ matrix with 2×2 blocks given by $\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$.

This means that $\overline{S}^{\perp}$ can be identified with the set of vectors x such that $x^T By = 0$ for all binary vectors y associated with the members of S. It turns out that the inner product $\langle x, y\rangle = x^T By$ arises in the study of classical codes over the four-element mathematical field $GF(4)$, which can be represented by the vectors 00, 01, 10, and 11 with addition modulo 2 and a new multiplication operation. This relationship leads to the construction of many good stabilizer codes (Calderbank et al. 1998).

Conserved Quantities, Symmetries, and Noiseless Subsystems. Even though a physical system may be exposed to error, some of its properties are often not affected by the errors. If these conserved quantities can be identified with the defining quantities of qubits or other information units, error-free storage of information can be ensured without active intervention. This is the idea behind noiseless subsystems.

When do noiseless subsystems exist and how can they be constructed? The examples discussed in the previous sections show that a noiseless subsystem may be a subset of physical qubits, as in the trivial two-qubit example, or it may require a more abstract subsystem identification, as in the example of the three spin-1/2 particles. As will be explained, in both cases, there are quantities conserved by the errors that can be used to identify the noiseless subsystem.

A simple classical example for the use of conserved quantities consists of two physical bits subject to errors that either flip both bits or leave them alone. A quantity invariant under this noise model is the parity $P(s)$ of a state s of the two bits. The parity $P(s)$ is defined as the number of 1s in the bit string s reduced modulo 2: $P(00) = P(11) = 0$, and $P(01) = P(10) = 1$. Flipping both bits does not change the value of P. Consequently, the two values of P can be used to identify the two states of a noiseless bit. The syndrome subsystem can be associated with the value (nonconserved) of the first physical bit using the function defined by $F(0b) = 0$, $F(1b) = 1$. The corresponding subsystem identification is obtained by using the values of P and F as the states of the syndrome (left) and the noiseless information-carrying subsystem (right) according to $ab \leftrightarrow F(ab) \cdot P(ab)$.

In quantum systems, conserved quantities are associated with the presence of symmetries, that is, with operators that commute with all possible errors. In the trivial two-qubit example, operators acting only on qubit 2 commute with the error operators. In particular, if E is any one of the errors, $E\sigma_u^{(2)} = \sigma_u^{(2)}E$ for $u = x, y, z$. It follows that the expectations of $\sigma_u^{(2)}$ are conserved. That is, if ρ is the initial state (density matrix) of the two physical qubits and ρ' is the state after the errors acted, then $\mathrm{tr}\,\sigma_u^{(2)}\rho' = \mathrm{tr}\,\sigma_u^{(2)}\rho$. Because the state of qubit 2 is completely characterized by these expectations, it follows immediately that it is unaffected by the noise.

The trivial two-qubit example suggests a general strategy for finding a noiseless qubit: First, determine the commutant of the errors, which is the set of operators that commute with all errors. Then, find a subset of the commutant that is algebraically equivalent to the operators characterizing a qubit. The equivalence can be formulated as a one-to-one map f from qubit operators to operators in the commutant. For the range of f to be algebraically equivalent, f must be linear and satisfy $f(A^\dagger) = f(A)^\dagger$ and $f(AB) = f(A)f(B)$. Once such an equivalence is found, a fundamental theorem from the representation theory of finite dimensional operator algebras implies that a subsystem identification for a noiseless qubit exists (Knill et al. 2000, Viola et al. 2001).

The strategy can be applied to the example of three spin-1/2 particles subject to collective errors. One can determine the commutant by using the physical properties of spin to find the conserved quantities associated with operators in the commutant, as suggested in the box "Creating a Noiseless Subsystem from Three Spin-1/2 Particles" on page 205. Alternatively, observe that, by definition, this error model is symmetric under permutations of the particles. Therefore, the actions of these permutations on the state

space form a group Π of unitary operators commuting with the errors. It is a fact that the commutant of the set of collective errors consists of the linear combinations of operators in Π. With respect to the group Π, one can immediately determine the space $V_{3/2}$ of symmetric states, that is, those that are invariant under the permutations. It is spanned by

$$\left|\uparrow\uparrow\uparrow\right\rangle,\ \frac{1}{\sqrt{3}}\left(\left|\uparrow\uparrow\downarrow\right\rangle+\left|\uparrow\downarrow\uparrow\right\rangle+\left|\downarrow\uparrow\uparrow\right\rangle\right),\ \frac{1}{\sqrt{3}}\left(\left|\uparrow\downarrow\downarrow\right\rangle+\left|\downarrow\uparrow\downarrow\right\rangle+\left|\downarrow\downarrow\uparrow\right\rangle\right),\ \left|\downarrow\downarrow\downarrow\right\rangle\ . \tag{47}$$

A basic result from the representation theory of groups implies that the projection onto $V_{3/2}$ is given by $P_{3/2} = (1/6)\sum_{g\in\Pi} g$. The orthogonal complement $V_{1/2}$ of $V_{3/2}$ is invariant under Π and can be analyzed separately. With the subsystem identification of Equation (26) already in hand, one can see that the permutation π_1, which permutes the spins according to $1 \to 2 \to 3 \to 1$, acts on the noiseless qubit, by applying $Z_{240^\circ} = e^{-i\sigma_z 2\pi/3}$, a 240° rotation around the z-axis. Similarly, the permutation π_2, which exchanges the last two spins, acts as σ_x on the qubit. To make them algebraically equivalent to the corresponding qubit operators, it is necessary to eliminate their action on $V_{3/2}$ by projecting onto $V_{1/2}$: $\pi_1' = (1 - P_{3/2})\pi_1$ and $\pi_2' = (1 - P_{3/2})\pi_2$. Sums of products of π_1' and π_2' are equivalent to the corresponding sums of products of Z_{240° and σ_x, which generate all qubit operators. To get the subsystem identification of Equation (26), one can start with a common eigenstate $|\psi\rangle$ of π_1' (a z-rotation on the noiseless qubit) and $2J_z$ (the syndrome subsystem's σ_z) with eigenvalues $e^{-i2\pi/3}$ and 1, respectively. The choice of eigenvalues implies that $|\psi\rangle \leftrightarrow |\uparrow\rangle \cdot |0\rangle$ in the desired identification. We can obtain the other logical states of the syndrome spin 1/2 and the noiseless qubit by applying π_2', $2J_x$, and $\pi_2' 2J_x$ to $|\psi\rangle$, which act by flipping the states of the qubit or the syndrome spin. This method for obtaining the subsystem identification generalizes to other operator equivalences and error operators.

Fault-Tolerant Quantum Communication and Computation

The utility of information and information processing depends on the ability to implement large numbers of information units and information-processing operations. We say that an implementation of information processing is scalable if the implementation can realize arbitrarily many information units and operations without loss of accuracy and with physical resource overheads that are polynomial (or efficient) in the number of information units and operations. Scalable information processing is achieved by implementing information fault-tolerantly.

One of the most important results of the work in quantum error-correction and fault-tolerant computation is the accuracy threshold theorem, according to which scalability is possible, in principle, for quantum information.

Theorem 9. Assume the requirements for scalable QIP (see below). If the error per gate is less than a threshold, then it is possible to efficiently quantum-compute to arbitrary accuracy.

Requirements for Scalable QIP. The value of the threshold accuracy (or error) depends strongly on which set of requirements is used—in particular, the error model that is assumed. The requirements are closely related to the basic requirements for constructing a quantum information processor (DiVincenzo 2000) but have to include

explicit assumptions on the error model and on the temporal and spatial aspects of the available quantum control:

Scalable physical systems. It is necessary to access physical systems that are able to support qubits or other basic units of quantum information. The systems must be scalable; that is, they must be able to support any number of independent qubits.

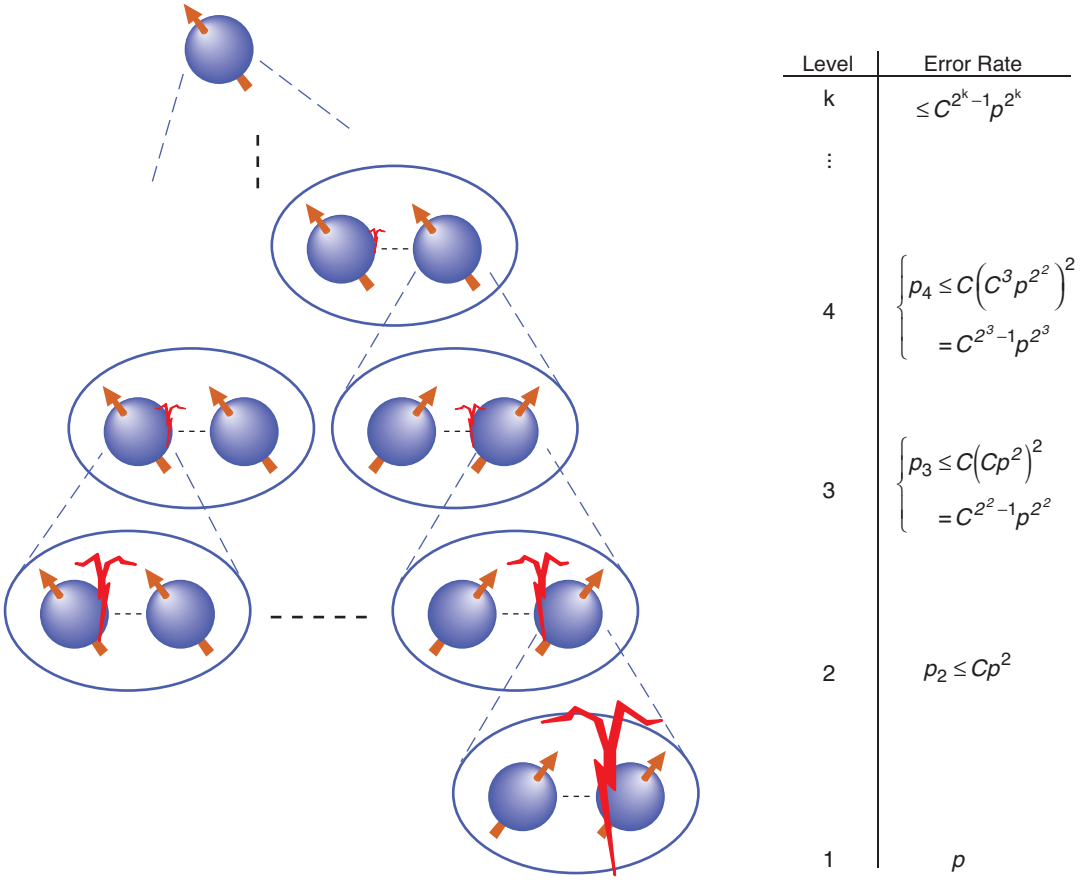
State preparation. One must be able to prepare any qubit in the standard initial state $|0\rangle$. Any preexisting content is assumed to be lost, as would happen if, for example, the qubit is first discarded and then replaced by a prepared one. The condition can be weakened; that is, it is sufficient that a large fraction of the qubits can be prepared in this way.

Measurement. Being able to measure any qubit in the logical basis is a requirement. Again, it is sufficient that a large enough fraction of the qubits are measurable. For solving computational problems with deterministic answers, the standard projective measurement can be replaced by weak measurements that return a noisy number whose expectation is the probability that a qubit is in the state $|1\rangle$ (Laflamme et al. 2001).

Quantum control. One must be able to implement a universal set of unitary quantum gates acting on a small number (usually, at most, two at a time) of qubits. For most accuracy thresholds, it is necessary to be able to apply the quantum control in parallel to any number of disjoint pairs of qubits. This parallelism requirement can be weakened if a nearly noiseless quantum memory is available. The requirement that it be possible to apply two-qubit gates to any pair of qubits is unrealistic given the constraints of three-dimensional space. Work on how to deal with this problem is ongoing (Aharonov and Ben-Or 1999). The universality assumption can be substantially weakened by replacement of some or all unitary quantum gates with operations to prepare special states or by additional measurement capabilities. See, for example, Michael Nielsen (2001) and the references therein.

Errors. The error probability per gate must be below a threshold and satisfy independence and locality properties (refer to the section "Error Models"). The definition of gate includes the "no-op," which is the identity operation implemented over the time required for a computational step. For the most pessimistic, independent, local error models, the error threshold is above $\sim 10^{-6}$; for the independent depolarizing errors, it is believed to be better than 10^{-4} (Gottesman and Preskill 1999). For some special error models, the threshold is substantially higher. For example, for the independent "erasure" error model, where error events are always detected, the threshold is above .01, and for an error model whose errors are specific, unintentional measurements in the standard basis of a qubit, the threshold is 1 (Knill et al. 2000). The threshold is also well above .01 when the goal is only to transmit quantum information through noisy quantum channels (Briegel et al. 1998).

Realizing Fault Tolerance. The existing proofs of the accuracy threshold theorems consist of explicit instructions for building a scalable quantum information processor and analyses of its robustness against the assumed error model. The instructions for realizing scalable computation are based on the following simple idea. Suppose that the error rate per operation for some way of realizing qubits is p. We can use these qubits and a quantum error-correcting code to encode logical qubits for which the storage error rate is reduced. For example, if a one-error correcting code is used, the error rate per storage interval for the logical qubits is expected to be $\leq cp^2$ for some constant c. Suppose that we can show how to implement encoded operations, preparations, measurement, and the subroutines required for error correction such that this inequality is now valid for each basic encoded step, perhaps for a larger constant C. Suppose furthermore that the errors for the encoded information still satisfy the assumed error model. The newly defined logical qubits then have an error rate of $\leq Cp^2$, which is less than p



for $p < 1/C$. We can use the newly realized qubits as a foundation for making higher-level logical qubits. The result is multiple levels of encodings. In the next level (level 2), the error rate is $\le C^3p^4$, and after k iterations, it is $\le C^{2^k-1}p^{2^k}$, a doubly exponentially decreasing function of k. This procedure is called concatenation (refer to Figure 8). Because the complexity, particularly the number of physical qubits needed for each final logical qubit, grows only singly exponentially in k, the procedure is efficient. Specifically, to achieve a logical error of ε per operation requires of the order of $|\log(\varepsilon)|^r$ resources per logical qubit for some finite r. In practice, this simple idea is still dauntingly complex, but there is hope that, for realistic errors in physical systems and by cleverly trading off different variations of these techniques, much of the theoretical complexity can be avoided (Steane 1999).

Many important developments and ideas of quantum information were ultimately needed to realize encoded operations, preparations, measurements, and error correction subroutines that behave well with respect to concatenation. Stabilizer codes provide a particularly nice setting for implementing many of these techniques. One reason is that good stabilizer codes are readily constructed. Another is that they enable encoding operations in a way that avoids spreading errors between the qubits of a single code word (Gottesman 1998). In addition, there are many tricks based on teleportation that can be used to maintain the syndrome subsystems in acceptably low error states and to implement general operations systematically (Gottesman and Chuang 1999). To learn more about all these techniques, see the textbook by Nielsen and Isaac Chuang (2001) and the works of Daniel Gottesman (1998) and John Preskill (1998).

Figure 8. Schematic Representation of Concatenation
The bottom level represents qubits realized more or less directly in a physical system. Each next level represents logical qubits defined by means of subsystems in terms of the previous level's qubits. More efficient subsystems might represent multiple qubits in one code block rather than the one qubit per code block shown here.

Concluding Remarks

The advancements in quantum error-correction and fault-tolerant QIP have shown that, in principle, scalable quantum computation is achievable. This is a crucial result because it suggests that experimental efforts in QIP will eventually lead to more than a few small-scale applications of quantum information to communication and problems with few qubits. However, the general techniques for achieving scalability that are known are difficult to realize. Existing technologies are far from achieving sufficient accuracy even for just two qubits—at least in terms of the demands of the usual accuracy-threshold theorems. There is hope that more optimistic thresholds can be shown to apply if one takes into consideration the specific constraints of a physical device, better understands the dominant sources of errors, and exploits tailor-made ways of embedding quantum information into subsystems. Current work in this area is focused on finding such methods of quantum error control. These methods include approaches to error control not covered in this article—for example, techniques for actively turning off the error-inducing environmental interactions (Viola and Lloyd 1998, Viola et al. 1999) and modifications to controlling quantum systems that eliminate systematic and calibration errors (Levitt 1982, Cummins and Jones 1999). Further work is also needed to improve the thresholds for the more pessimistic error models and for developing more-efficient scalability schemes. ■

Contact Information

E. Knill: knill@lanl.gov

R. Laflamme: laflamme@iqc. ca

A. Ashikhmin: aea@research.bell-labs.com

H. Barnum: barnum@lanl.gov

L. Viola: viola@lanl.gov

W. H. Zurek: whz@lanl.gov

Further Reading

Aharonov, D., and M. Ben-Or. 1996. Fault-Tolerant Quantum Computation with Constant Error Rate. In *Proceedings of the 29th Annual ACM Symposium on the Theory of Computation* (*STOC*). New York: ACM Press.

———. 1999. Fault-Tolerant Quantum Computation with Constant Error. [Online]: http://eprints.lanl.gov. (quant-ph/9906129).

Ashikhmin, A., and S. Litsyn. 1999. Upper Bounds on the Size of Quantum Codes. *IEEE Trans. Inf. Theory* **45**: 1206.

Bennett, C. H., D. P. DiVincenzo, J. A. Smolin, and W. K. Wootters. 1996. Mixed State Entanglement and Quantum Error-Correcting Codes. *Phys. Rev. A* **54**: 3824.

Braunstein, S. L., and H.-K. Lo, eds. 2000. Special Focus Issue on Experimental Proposals for Quantum Computation: Forward. *Fortschr. Phys*. **48** (9–11): 76

Briegel, H.-J., W. Dür, J. I. Cirac, and P. Zoller. 1998. Quantum Repeaters for Communication. [Online]: http://eprints.lanl.gov. (quant-ph/9803056).

Calderbank, A. R., E. M. Rains, P. W. Shor, and N. J. A. Sloane. 1997. Quantum Error Correction and Orthogonal Geometry. *Phys. Rev. A* **78**: 405.

———. 1998. Quantum Error Correction via Codes over gf(4). *IEEE Trans. Inf. Theory* **44**: 1369.

Cummins, H. K., and J. A. Jones. 2000. Use of Composite Rotations to Correct Systematic Errors in NMR Quantum Computation. *New J. Phys*. **2**: 6.

DiVincenzo, D. P. 2000. The Physical Implementation of Quantum Computation. *Fortschr. Phys*. **48**: 771.

Gottesman, D. 1996. A Class of Quantum Error-Correcting Codes Saturating the Quantum Hamming Bound. *Phys. Rev. A* **54**: 1862.

———. 1998. A Theory of Fault-Tolerant Quantum Computation. *Phys. Rev. A* **57**: 127.

Gottesman, D., and I. L. Chuang. 1999. Demonstrating the Viability of Universal Quantum Computation Using Teleportation and Single-Qubit Operations. *Nature* **402**: 390.

Kitaev, A. Yu. 1997. Quantum Error Correction with Imperfect Gates. In *Quantum Communication and Computing and Measurement*. Edited by 0. Hirota et al. New York: Plenum.

Knill, E., and R. Laflamme. 1996. "Concatenated Quantum Codes." [Online]: http://eprints.lanl.gov. (quant-ph/9608012).

Knill, E., R. Laflamme, and G. Milburn. 2000. "Thresholds for Linear Optics Quantum Computation." [Online]: http://eprints.lanl.gov. (quant-ph/0006120).

Knill, E., R. Laflamme, and W. H. Zurek. 1998a. Resilient Quantum Computation. *Science* **279**: 342.

———. 1998b. Resilient Quantum Computation: Error Models and Thresholds. *Proc. R. Soc. London, Ser. A* **454**: 365.

Knill, E., R. Laflamme, and L. Viola. 2000. Theory of Quantum Error Correction for General Noise. *Phys. Rev. Lett.* **84**: 2525.

Kraus, K. 1983. States, Effects and Operations: Fundamental Notions of Quantum Theory. *Lecture Notes in Physics*. Vol. 190. Berlin: Springer-Verlag.

Laflamme, R., C. Miquel, J-P. Paz, and W. H. Zurek. 1996. Perfect Quantum Error-Correcting Code. *Phys. Rev. Lett.* **77**: 198.

Levitt, M. H. 1982. Symmetrical Composite Pulse Sequences for NMR Population-Inversion. I. Compensation for Radiofrequency Field Inhomogeneity. *J Mag. Res*. **48**:234.

Nielsen, M. A. 2001. Universal Quantum Computation Using Only Projective Measurement, Quantum Memory, and Preparation of the $|0\rangle$ State. [Online]: http://eprints.lanl.gov. (quant-ph/0108020).

Nielsen, M. A., and I. L. Chuang. 2001. *Quantum Computation and Quantum Information*. Cambridge, U.K.: Cambridge University Press.

Preskill, J. 1998. Reliable Quantum Computers. *Proc. R. Soc. Lond*on*, Ser. A* **454**: 385.

Schumacher, B. 1996. Sending Entanglement through Noisy Quantum Channels. *Phys. Rev. A* **54**: 2614.

Shor, P. W. 1995. Scheme for Reducing Decoherence in Quantum Computer Memory. *Phys. Rev. A* **2**: 2493.

———. 1996. Fault-Tolerant Quantum Computation. In *Proceedings of the 37th Symposium on the Foundations of Computer Science (FOCS)*. p. 56. Los Alamitos, CA: IEEE Press.

Steane, A., 1996. Multiple Particle Interference and Quantum Error Correction. *Proc. R. Soc. London, Ser. A* **452**: 2551.

———. 1999. Efficient Fault-Tolerant Quantum Computing. *Nature* **399**: 124.

Viola, L., and S. Lloyd. Dynamical Suppression of Decoherence in Two-State Quantum Systems. *Phys. Rev. A* **58**: 2733.

Viola, L., E. Knill, R. Laflamme. 2001. Constructing Qubits in Physical Systems. *J. Phys. A* **34** (35): 7067.

Viola, L., E. Knill, and S. Lloyd. 1999. Dynamical Decoupling of Open Quantum Systems. *Phys. Rev. Lett*. **82**: 2417.

Glossary

Bit. The basic unit of deterministic information. It is a system that can be in one of two possible states, 0 and 1.

Bit string. A sequence of 0s and 1s that represents a state of a sequence of bits. The bit strings are words in the binary alphabet.

Classical information. The type of information based on bits and bit strings and, more generally, on words formed from finite alphabets. This is the information used for communication between people. Classical information can refer to deterministic or probabilistic information, depending on the context.

Code. A set of states that can be used to represent information. The set of states needs to have the properties of the type of information to be represented. The code is usually a subset of the states of a given system Q. It is then a Q-code or a code on Q. If information is represented by a state in the code, Q is said to carry the information.

Code word. A state in a code. The term is primarily used for classical codes defined on bits or systems with nonbinary alphabets.

Concatenation. An iterative procedure in which higher-level logical information units are implemented in terms of lower-level units.

Control error. An error due to nonideal control in applying operations or gates.

Communication channel. A means for transmitting information from one place to another. It can be associated with a physical system in which the information to be transmitted is stored by the sender. The system is subsequently conveyed to the receiver, who can then make use of the information.

Correctable error set. For a given code, a set of errors such that there is an implementable procedure R that, after any one of the errors E acts on a state x in the code, returns the system to the state $x = REx$. What procedures are implementable depends on the type of information represented by the system and, if it is a physical system, its physics.

Decoding. The process of transferring information from an encoded form to its "natural" form. In the context of error correction, decoding is often thought of as consisting of two steps: one which removes the errors' effects (sometimes called the recovery procedure) and one that extracts the information (often also called decoding in a narrower sense).

Depolarizing errors. An error model for qubits in which random Pauli operators are applied independently to each qubit.

Detectable error. For a given code, an error that has no effect if the state is observed to have remained in the code. If the state is no longer in the code, the error is said to have been detected, and the state no longer represents valid information.

Deterministic information. The type of information based on bits and bit strings. This is the same as classical information but explicitly excludes probabilistic information.

Encoding. The process of transferring information from its natural form to an encoded form. It requires an identification of the valid states associated with the information and the states of a code. The process acts on an information unit and replaces it with the system whose state space contains the code.

Environment. In the context of information encoded in a physical system, it refers to other physical systems that may interact with the information-carrying system.

Environmental noise. Noise due to unwanted interactions with the environment.

Error. Any unintended effect on the state of a system, particularly in storing or otherwise processing information.

Error basis. A set of state transformations that can be used to represent any error. For quantum systems, errors can be represented as operators acting on the system's state space, and an error basis is a maximal, linearly independent set of such operators.

Error control. The term for general procedures that limit the effects of errors on information represented in noisy, physical systems.

Error correction. The process of removing the effects of errors on encoded information.

Error-correcting code. A code with additional properties that enable a decoding procedure to remove the effects of the dominant sources of errors on encoded information. Any code is error correcting for some error model in this sense. To call a code error correcting emphasizes the fact that it was designed for this purpose.

Error model. An explicit description of how and when errors happen in a given system. Typically, a model is specified as a probability distribution over error operators. More general models may need to be considered, particularly in the context of fault-tolerant computation, for which correlations in time are important.

Fault tolerance. A property of encoded information that is being processed with gates. It means that errors occurring during processing, including control errors and environmental noise, do not seriously affect the information of interest.

Gate. An operation applied to information for the purpose of information processing.

Hamming distance. The Hamming distance between two binary words (sequences of 0 and 1) is the number of positions in which the two words disagree.

Hilbert space. An n-dimensional Hilbert space consists of all complex n-dimensional vectors. A defining operation in a Hilbert space is the inner product. If the vectors are thought of as column vectors, then the inner product $\langle x, y\rangle$ of x and y is obtained by forming the conjugate transpose $x^\dagger$ of x and calculating $\langle x, y\rangle = x^\dagger y$. The inner product induces the usual norm $|x|^2 = \langle x, x\rangle$.

Information. Something that can be recorded, communicated, and computed with. Information is fungible, which implies that its meaning can be identified regardless of the particulars of the physical realization. Thus, information in one realization (such as ink on a sheet of paper) can be easily transferred to another (for example, spoken words). Types of information include deterministic, probabilistic, and quantum information. Each type is characterized by information units, which are abstract systems whose states represent the simplest information of this type. These define the natural representation of the information. For deterministic information, the unit is the bit, whose states are symbolized by 0 and 1. Information units can be put together to form larger systems and can be processed with basic operations acting on a small number of units at a time.

Length. For codes on n basic information units, the length of the code is n.

Minimum distance. The smallest number of errors that is not detectable by a code. In this context, the error model consists of a set of error operators without specified probabilities. Typically, the concept is used for codes on n information units, and the error model consists of operators acting on any one of the units. For a classical binary code, the minimum distance is the smallest Hamming distance between two code words.

Noise. Any unintended effect on the state of a system, particularly an effect with a stochastic component due to incomplete isolation of the system from its environment.

Operator. A function transforming the states of a system. Operators may be restricted, depending on the system's properties. For example, operators acting on quantum systems are always assumed to be linear.

Pauli operators. The Hermitian matrices σ_x, σ_y, and σ_z—refer to Equation (7)—acting on qubits. It is often convenient to consider the identity operator to be included in the set of Pauli operators.

Physical system. A system explicitly associated with a physical device or particle. The term is used to distinguish between abstract systems used to define a type of information and specific realizations, which are subject to environmental noise and errors due to other imperfections.

Probabilistic bit. The basic unit of probabilistic information. It is a system whose state space consists of all probability distributions over the two states of a bit. The states can be thought of as describing the outcome of a biased coin flip before the coin is flipped.

Probabilistic information. The type of information obtained when the state spaces of deterministic information are extended with arbitrary probability distributions over the deterministic states. This is the main type of classical information with which quantum information is compared.

Quantum information. The type of information obtained when the state space of deterministic information is extended with arbitrary superpositions of deterministic states. Formally, each deterministic state is identified with one of an orthonormal basis vector in a Hilbert space, and superpositions are unit-length vectors that are expressible as complex linear sums of the chosen basis vectors. Ultimately, it is convenient to extend this state space again by permitting probability distributions over the quantum states. This is still called quantum information.

Qubit. The basic unit of quantum information. It is the quantum extension of the deterministic bit; that is, its state space consists of the unit-length vectors in a two-dimensional Hilbert space.

Repetition code. The classical, binary repetition code of length n consists of the two words 00 ... 0 and 11 ... 1. For quantum variants of this code, one applies the superposition principle to obtain the states consisting of all unit-length complex linear combinations of the two classical code words.

Scalability. A property of physical implementations of information processing that implies that there are no bounds on accurate information processing. That is, arbitrarily many information units can be realized, and they can be manipulated for an arbitrarily long amount of time without loss of accuracy. Furthermore, the realization is polynomially efficient in terms of the number of information units and gates used.

States. The set of states for a system characterizes the system's behavior and possible configurations.

Subspace. For a Hilbert space, a subspace is a linearly closed subset of the vector space. The term can be used more generally for a system Q of any information type: A subspace of Q or, more specifically, of the state space of Q is a subset of the state space that preserves the properties of the information type represented by Q.

Subsystem. A typical example of a subsystem is the first (qu)bit in a system consisting of two (qu)bits. In general, to obtain a subsystem of system Q, one first selects a subset C of Q's state space and then identifies C as the state space of a pair of systems. Each member of the pair is then a subsystem of Q. Restrictions apply, depending on the types of information carried by the system and subsystems. For example, if Q is quantum and so are the subsystems, then C has to be a linear subspace and the identification of the subsystems' state space with C has to be unitary.

Subsystem identification. The mapping or transformation that identifies the state space of two systems with a subset C of states of a system Q. In saying that L is a subsystem of Q, we also introduce a second subsystem and identify the state space of the combined system with the subset of states C.

Syndrome. One of the states of a syndrome subsystem. It is often used more narrowly for one of a distinguished set of basis states of a syndrome subsystem.

Syndrome subsystem. In identifying an information-carrying subsystem in the context of error correction, the other member of the pair of subsystems required for the subsystem identification is called the syndrome subsystem. The terminology comes from classical error correction, in which the syndrome is used to determine the most likely error that has occurred.

System. An entity that can be in any of a specified number of states. An example is a desktop computer whose states are determined by the contents of its various memories and disks. Another example is a qubit, which can be thought of as a particle whose state space is identified with complex, two-dimensional length-one vectors. Here, a system is always associated with a type of information, which in turn determines the properties of the state space**.** For example, for quantum information, the state space is a Hilbert space. For deterministic information, it is a finite set called an alphabet.

Twirling. A randomization method for ensuring that errors act like a depolarizing error model. For one qubit, it involves applying a random Pauli operator before the errors occur and then undoing the operator by applying its inverse.

Unitary operator. A linear operator U on a Hilbert space that preserves the inner product. That is, for all x and y, $\langle Ux, Uy\rangle = \langle x, y\rangle$. If U is given in matrix form, then this condition is equivalent to $U^{\dagger}U = \mathbb{1}$.

Weight. For a binary word, the weight is the number of 1s in the word. For an error operator acting on n systems by applying an operator to each one of them, the weight is the number of nonidentity operators applied.