



Classical Leakage Resilience from Fault-Tolerant Quantum Computation

Felipe G. Lacerda

Institute for Theoretical Physics, ETH Zurich, Zürich, Switzerland
Departamento de Ciência da Computação, Universidade de Brasília, Brasília, Brazil
fegolac@gmail.com

Joseph M. Renes · Renato Renner

Institute for Theoretical Physics, ETH Zurich, Zürich, Switzerland
renes@phys.ethz.ch
renner@phys.ethz.ch

Communicated by Stefan Wolf.

Received 12 September 2014 / Revised 22 December 2018

Online publication 8 August 2019

Abstract. Physical implementations of cryptographic algorithms leak information, which makes them vulnerable to the so-called side-channel attacks. The problem of secure computation in the presence of leakage is generally known as leakage resilience. In this work, we establish a connection between leakage resilience and fault-tolerant quantum computation. We first prove that for a general leakage model, there exists a corresponding noise model in which fault tolerance implies leakage resilience. Then we show how to use constructions for fault-tolerant quantum computation to implement classical circuits that are secure in specific leakage models.

1. Introduction

Modern theoretical cryptography is primarily concerned with developing schemes that are *provably secure* under reasonable assumptions. While the field has been hugely successful, the threat model considered usually does not allow for the possibility of *side-channel attacks*—attacks on the physical implementation of the cryptographic scheme.

Side-channel attacks have been a worry long before the advent of modern cryptography. As early as 1943, it was discovered that a teletype used for encryption by the American military caused spikes in an oscilloscope that could then be used to recover the plaintext [21]. More recently, side-channel attacks on cryptographic applications widely used in practice have been revealed. One of these is the “Lucky Thirteen” attack on TLS in CBC mode [2], which is based on measuring the time it takes the server to reply to a request over the network. Another recent attack [13] uses acoustic crypt-

analysis to attack the GnuPG implementation of RSA. The authors managed to extract the full RSA key by measuring the noise produced by the computer while it decrypts a set of chosen ciphertexts. Although the relevant software has since then been updated so that these attacks are no longer possible, they highlight the importance of designing implementations with side-channel attacks in mind.

The theoretical approach to side-channel attacks is to design *protocols* that are resilient against them. This is the focus of the area known as *leakage resilience*. In this work, we present a way to perform universal leakage-resilient computation, i.e., we construct a general “leakage-resilient compiler” that takes an arbitrary circuit and produces a new, leakage-resilient version having the same computational functionality.

We take a novel approach to leakage resilience: fault-tolerant quantum computation. The basic idea is that all actions performed in the execution of a classical circuit—as well as leakage attacks on it—can be described using the formalism of quantum mechanics. Then, leakage of the physical state of the computation is equivalent to a so-called phase error in a quantum circuit. Since a fault-tolerant quantum computation must protect against phase errors (as well as more conventional bit-flip errors), it is necessarily leakage-resilient. However, achieving leakage resilience in this way would require a fault-tolerant quantum circuit. Here, we give a method for constructing a leakage-resilient *classical* circuit by modifying an appropriate fault-tolerant *quantum* circuit. The former mimics the latter, inheriting its leakage resilience. Our approach is similar to security proofs of quantum key distribution in which it is shown that the operation of the actual protocol—where the outcomes of measurements on entangled quantum states are processed classically—can be interpreted as mimicking a fully quantum protocol for entanglement distillation, from which the protocol derives its security [25].

Before introducing our general setting for leakage, it is illustrative to present a concrete example of leakage resilience in practice. Smart cards are integrated circuits that have been widely used for authentication and also allow the storing of sensitive data, making them portable carriers of information such as money and medical records. Since smart cards are designed to be portable, they are subject to a variety of physical attacks. Possible attacks include measuring the time and the electrical current used when performing operations (power analysis). Thus, leakage resilience is essential to the design of smart cards.

More specifically, a smart card stores some internal data, provides an interface for external input, performs some computation on the internal data and the external input and sends the output through an output interface. One of the design goals is that if the smart card is given to an adversary, who can send inputs to and read off outputs from it, the adversary should not be able to obtain any information about the internal data in the card, beyond what can be learned from the regular output of the computation. This design goal is what we mean by leakage resilience. However, whereas smart cards are usually designed to be resilient against specific attacks, we are interested in larger classes of attacks.

Our general setting for leakage is as follows: First, an honest party (Alice) inputs a circuit C that computes a function f along with a secret y . Then, the adversary (Eve) is given black-box access to the circuit, so that she can interactively send inputs, denoted x , to the circuit and gets the corresponding outputs ($f(x, y)$). She also receives a description of the circuit. Additionally, with each interaction Eve obtains leaked bits according to the

leakage model. Roughly, the goal is that Eve should not learn any more information about y than what she would have access to $f(\cdot, \cdot)$ alone. This setting will be formalized in Sect. 2 using the abstract cryptography framework [18, 20].

1.1. Previous Work

The question of hiding internal computation from an eavesdropper is related to the problem of *program obfuscation*. Obfuscation can be seen as a “worst-case leakage resilience,” in which the internal state must be protected even if the whole execution leaks to the adversary.¹ However, it is known that obfuscating programs is impossible in general [4]. Thus, if we want any leakage resilience at all, the type of leakage allowed has to be restricted in some way.

This is not a problem in practice because adversaries will typically be restricted in the type and amount of information they can obtain from the system. Here we list a few results in the field, although the list is by no means comprehensive. Ishai et al. [15] considered adversaries that can learn the values of a bounded number of wires in the circuit. Micali and Reyzin [19] introduced the “only computation leaks information” assumption, in which the leakage at each step of the computation only depends on data that were used in the computation. Faust et al. [9] consider a model where the adversary gets a function of all of the circuit’s state (that is, the output and all intermediate computations), the restriction being that the leakage function must be computable in AC^0 , the family of circuits containing only AND, NOT and OR gates (with unbounded fan-in) and having constant depth and polynomial size. They also consider noisy leakage, where the whole state of the circuit leaks, but the adversary only receives a noisy version of the leaked state, each bit being flipped with some probability. A common further assumption is to use a small component of trusted hardware [8, 9, 11].

Despite these advances, relating leakage models to actual leakage seen in practice has so far proved to be challenging. Standaert et al. [27] present a few problems with the “bounded leakage” assumption, in which the leakage at each computation step is assumed to be bounded; see also [26].

Our work is loosely related to the one of Smith et al. [6], where techniques from fault-tolerant quantum computation are used to develop a construction for secure multi-party quantum computation. Given that multi-party computation techniques are commonly used for leakage resilience (in particular, secret sharing [8, 11, 15]), the connection between fault tolerance and leakage resilience is perhaps unsurprising. However, the construction of Smith et al. was inherently quantum, whereas our construction runs on a classical machine.

1.2. Our Contribution

We establish a relation between leakage-resilient (classical) computation and fault-tolerant quantum computation, which are formally defined in Sect. 2. Specifically, we show how methods of the latter can be used to construct leakage-resilient compilers,

¹One explicit connection between obfuscation and a specific leakage model (the “only computation leaks information” model of Micali and Reyzin [19]) has been made in [3]; see also [12].

which transform a given circuit into another (classical) circuit with the same computational functionality as well as resistance to leakage.

The starting point in relating leakage-resilient classical computation and fault-tolerant quantum computation is the observation that any classical logical operation can be regarded as a quantum operation performing the same action in the so-called computational basis. (“Appendix A” provides a brief background on the formalism and tools of quantum information theory necessary here). Then, as we show in Sect. 3, any given leakage model may be interpreted as specific model of *phase noise* afflicting the corresponding quantum circuit.

Phase noise is not the most general type of quantum noise. Nonetheless, as we show in Theorem 1, if fault tolerance is possible for a given noise model—meaning roughly that error correction is performed frequently enough that the encoded information essentially never suffers from errors—the quantum computation is resilient to leakage of the corresponding leakage model. However, we want to make classical circuits leakage-resilient and do not necessarily want to carry out a quantum computation to achieve this goal. Fortunately, the structure of certain quantum error-correcting codes is such that we can mimic the error-correcting steps with classical circuits.

Following this approach, we construct a general leakage-resilient compiler by mimicking the basic fault-tolerant components of the fault-tolerant scheme in [1] using classical components. Section 4 describes how to transform into classical circuits certain simple types of quantum circuits that serve as building blocks for arbitrary circuits. Section 5 then presents a fault-tolerant implementation of the Toffoli gate, a gate that is universal for classical computation [30]. Combining this with the results of Sect. 4 then gives a leakage-resilient compiler.

As in other works, our construction assumes the existence of a small leak-free component: a source of (uniformly) random bits. Using the construction, one can transform an arbitrary classical circuit into a circuit that is resilient to leakage arising in any model for which reliable quantum computation is possible under the corresponding phase noise model. One particular leakage model that translates into a well-studied quantum noise model is that of independent leakage, in which the value of each wire of the circuit leaks with some fixed probability. While potentially too restrictive (in particular, the independence assumption implies that the adversary does not choose the wires that leak), its simplicity allows for an easy interpretation in the quantum scenario as independent phase errors, for which various fault-tolerant constructions are known. This model is used in our construction in Sect. 5. Leakage models that include correlations lead to error models that have yet to be analyzed.

We stress that rather than presenting a specific leakage-resilient scheme, the contribution of this work aims to provide a novel approach to leakage resilience and connect this field of cryptography to the research area of fault-tolerant quantum computation. Our work thus shows how results achieved in one area (e.g., new threshold theorems for quantum fault tolerance) can be translated to the other (e.g., bounds on the performance of leakage-resilient compilers). Our hope is that our result will inspire research developing the relationship between fault tolerance and leakage resilience. In Sect. 6 we discuss possible future directions for this line of work.

2. Definitions

2.1. Abstract Cryptography Framework

In order to define leakage resilience we use the *abstract cryptography* framework [18,20], of which we give a short summary. From an abstract viewpoint, constructing a protocol amounts to assuming that a certain set of *real resources* is available and then using them to build a new resource, termed an *ideal resource*. By way of composition, the ideal resource can then again be used as a real resource in another protocol to build a more complex ideal resource.

For instance, the one-time pad construction assumes that a resource giving out a secret key is available along with an authentic channel. It is then shown that these resources can emulate a secure channel. In this case, the ideal resource is the secure channel, whereas the authentic channel along with a shared secret key is used as a real resource. On the other hand, in a protocol for authentication, the authentic channel takes the role of the ideal resource, and the real resource is a completely insecure channel together with a secret key.

In this framework, resources are a type of *systems*, which are defined as abstract objects which can be composed. Each system has an interface set \mathcal{I} , and interfaces can be connected in order to form new systems. Resources are systems where each interface corresponds to one party that has access to it. As an example of a resource, we can define a private channel between honest parties A and B subject to possible eavesdropping by E as a resource that takes inputs from A and outputs them at B . Since the resource gives E no outputs, the channel is private by definition. Note that this holds whether E acts honestly or dishonestly; in abstract cryptography, the broader goal is to emulate the behavior of ideal resources in all situations, not just when assuming certain parties are honest and others dishonest.

To emulate an ideal resource from given resources, the latter may be composed. Furthermore, each party can act on their interface using a *converter*, which is also modelled as a system. A converter has an “inside” interface, connected to the resource, and an “outside” interface, which is used by the parties. A *protocol* specifies a converter for each party acting honestly and is applied to a real resource. In the security arguments, we will also consider converters applied to the ideal resource by parties acting dishonestly. These are termed *simulators*. We denote composition of systems by juxtaposition, so that, for instance, the resource formed by plugging converter π_A into the resource \mathcal{R} is denoted by $\pi_A \mathcal{R}$.

In order to allow for constructions that do not perfectly match a desired ideal resource but only approximate it, we need a notion of distance d between two resources, which must be a pseudo-metric d on the set of resources. Typically, we consider the distance to be the maximal advantage that a system trying to distinguish between the two resources (the *distinguisher*) can have. Given two resources \mathcal{R} and \mathcal{S} and a distance d , we also write $\mathcal{R} \approx_\varepsilon \mathcal{S}$ to denote $d(\mathcal{R}, \mathcal{S}) \leq \varepsilon$.

In this work, we only ever need to consider two parties, A and E , where A is assumed to be honest. To keep the formal treatment simple, we restrict ourselves to this case. In this scenario, what it means for a protocol to securely emulate an ideal resource \mathcal{S} given a resource \mathcal{R} takes a particularly simple form. It reduces to two different conditions: one

where E is dishonest and another where E is honest—that is, it follows the protocol. We use the following definition of security, which, for our scenario, is sufficient to imply the definition presented in [20].²

Definition 1. Let \mathcal{R} and \mathcal{S} be resources with interface set $\mathcal{I} = \{A, E\}$. We say that a protocol $\pi = \{\pi_A, \pi_E\}$ securely constructs \mathcal{S} from \mathcal{R} within ε if there exists a simulator σ_E such that

$$\pi_A \mathcal{R} \approx_\varepsilon \sigma_E \mathcal{S} \quad \text{and} \quad (1)$$

$$\pi_A \pi_E \mathcal{R} \approx_\varepsilon \mathcal{S}. \quad (2)$$

2.2. Leakage Resilience

Using the abstract cryptography framework, we can now define leakage resilience by specifying the corresponding ideal and real resource. They are two-party resources, with parties that we denote by Alice (A) and Eve (E), where Alice is assumed to be honest.

Ideal resource Our goal is to be able to compute a function that takes an input to be given by the adversary, and an additional input that corresponds to the initial secret. First, we describe intuitively the kind of resource that we want. At the beginning, Alice inputs the circuit to be executed along with the secret input. Then, a description of the circuit is given to Eve, who can execute it as a black box freely.

In light of this informal description, we define the ideal resource \mathcal{S} as follows. Alice initially inputs a secret y and a description of a circuit \mathcal{C} that evaluates a function f . Eve can send inputs x , to which she receives outputs $f(x, y)$. The ideal functionality also outputs \mathcal{C} to Eve. The resource is shown in Fig. 1a.

What this definition implies is that information about y can only leak through $f(x, y)$ and \mathcal{C} . As a concrete example, let \mathcal{C} be a circuit implementing an encryption algorithm f that encrypts inputs x using the secret key y . In establishing security for a cryptographic scheme, one assumes that the secret key is completely hidden from Eve. In the example, this assumption is ensured by the way the ideal resource is defined, since the only way information about y could leak to Eve is through the ciphertext $f(x, y)$.

This example also illustrates that security and leakage resilience are separate goals: We place no restrictions on f , so even if the function reveals some information about y that would be of no consequence for leakage resilience. Instead, leakage resilience only ensures that no additional information about y leaks.

Real resource Informally, the real resource (which we denote by \mathcal{R}) is a “leaky” version of the ideal resource, in which additional information becomes available to Eve. Alice’s interface allows her to input a secret y and a circuit \mathcal{C} that evaluates a function f . Eve’s interface allows her to send inputs x , to which she gets outputs $\mathcal{C}(x, y)$, and additionally send *leakage requests* l , getting leakages l' . As in the ideal resource, Eve also gets \mathcal{C} . The idea is that Eve can use \mathcal{R} as a black box, but also obtain additional information from

²We note, however, that in the general case one needs to use additional constructions, termed *filters*. We refer to [18, 20] for more details.

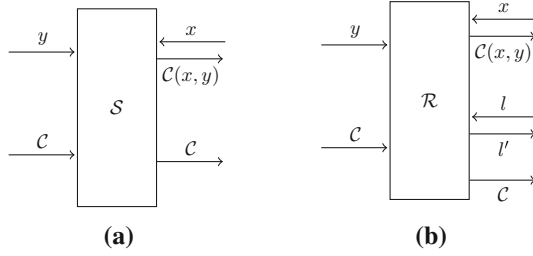


Fig. 1. **a** Ideal resource for leakage resilience and **b** real resource. In both cases, Alice has access to the left interface and Eve has access to the right interface.

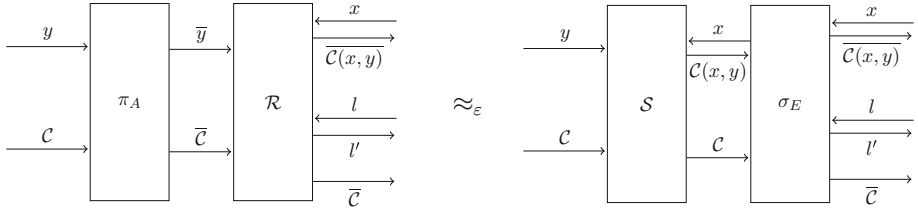


Fig. 2. The first condition of Definition 1 applied to leakage resilience. In order to prove that the converter π_A is part of a leakage-resilient protocol π , one has to show the existence of a simulator σ_E such that the execution of π_A running with the real resource R is indistinguishable from execution of the ideal resource S with σ_E . The second condition is easily satisfied by a converter π_E which inputs a fixed l and ignores the output l' , but forwards the other inputs and outputs unchanged. Here, $\bar{C}(x, y)$ denotes the encoding of the output $C(x, y)$.

the leakage which might reveal something about the secret. This scenario is represented in Fig. 1b.

In order to satisfy the first condition in Definition 1, we need a protocol π to construct the ideal resource S from the real resource R . The relevant setting for Definition 1 is depicted in Fig. 2.

Our description of the real resource captures leakage in a very general form, in the sense that C , l and l' may be arbitrary. However, in order for leakage resilience to be possible, it is necessary to specify how the leakage requests l are chosen. Hence, we construct schemes that are resilient against leakage with some particular structure. In order to formalize this, we will make the scenario presented above more concrete, by restricting the interactions in the following way. Eve interacts with the resource in rounds. In each round, she inputs pairs (x, l) and receives outputs $(f(x, y), l')$.³

The particular strategy used to choose the leakage requests l is referred to as a *leakage model* L . The leakage model is a set of allowed leakage requests. A leakage request is a probability distribution over leakage functions. That is, if the set of possible leakage functions is $\{l_i\}$, Eve sends $l = \{(l_i, p_i)\}$, where l_i is chosen with probability p_i . Each l_i is a function of the set of values assigned to the wires in the circuit, which we denote by $W_C(x, y)$ (when the circuit is given inputs x and y). If the function l_j is chosen, Eve

³We could have the function output additionally be a function of l ; this situation would capture *tampering*, in which Eve can introduce faults to the wires in the circuit [14].

receives $l' = l_j(W_C(x, y))$. Note that the output of l_j may or may not include the index j , depending on the leakage model.

As an example, Eve could choose at every round a bounded number of wires of the circuit to leak. (These are the “probing attacks” considered in [15].) Or one could restrict Eve to leakage functions that are computed with constant-depth circuits [9]. In any case, we define leakage-resilient compilers with respect to a particular leakage model in the following way. In the definition below, \mathcal{S} and \mathcal{R} are, respectively, the ideal resource and the real resource defined above.

Definition 2. The protocol π is an ε -leakage-resilient compiler against leakage model L if it securely constructs \mathcal{S} from \mathcal{R} within ε , where the leakage requests are drawn from L .

Independent leakage In this paper (Sect. 5), we provide as a concrete example an explicit leakage-resilient compiler for a concrete leakage model which we call *independent leakage*. Independent leakage is characterized by having every wire in the circuit potentially leak, each with a fixed probability p .

This model can be formalized in the following way. Let n be the number of wires in the circuit. We label each wire in the circuit with an index i with $1 \leq i \leq n$. Now let w be a binary string of length n and $l_w(x, y)$ be the values of the wires i with $w_i = 1$ when the circuit has (x, y) as input, along with the string w to indicate which wires have leaked. For every round of interaction, the probability p_w that the leakage function L_w was chosen is given by $p_w = \Pr(X = |w|)$, where X follows a binomial distribution with parameters n and p and $|w|$ denotes the Hamming weight of w . The leakage model L is then $L = \{(l_w, p_w) : w \in \{0, 1\}^n\}$. (Note that there is only one possible leakage request.)

Leak-free components Many of the leakage-resilient constructions require a small component of the circuit to be trusted or “leak-free,” meaning its internal wires do not leak to Eve. For our construction, this includes wires coming out of leak-free components. We incorporate this requirement in our definition of leakage resilience by adding the restriction that the leakage function l may not depend on any wires inside the leak-free components, or the wires coming out of them.

The only leak-free component we use in this work is a source of random bits: a gate that takes no input and outputs a uniformly distributed random bit (which is assumed not to leak). This is used in the part of our proof where we transform a quantum circuit into a classical one; we leave open the question of whether this is necessary in general. We note that our requirement is different from than the one used in [9]: We require only uniformly distributed bits, while their transformation requires bits distributed according to an arbitrary (although fixed) distribution. On the other hand, the output wires from their leak-free component are allowed to leak, whereas we require that leak-free components not leak before interacting with other components.

2.3. Fault Tolerance

A noisy quantum circuit \mathcal{E} is an implementation of a unitary U acting on subsystems S (the *data* subsystem) and E (the *environment*). The idea of fault tolerance is to perform

computations on a noisy circuit reliably, by using the noisy circuit to execute encoded operations such that at any point of the computation, the encoded state of the data is close to the state of an ideal circuit that implements U perfectly.

More precisely, let \mathcal{C} be a quantum circuit with inputs and outputs in $\mathcal{H}_2^{\otimes k}$. We assume that the circuit acts on classical inputs; this is done by preparing a known state $|0\rangle^{\otimes k}$ and then encoding a classical input x into it, obtaining the state denoted $|x\rangle$. The possible input states $|x\rangle$ form the computational basis. Then, for any x , the circuit implements the action of a unitary U_x on $|0\rangle^{\otimes k}$.

A fault-tolerant simulation for \mathcal{C} works as follows. Let $\tilde{\mathcal{C}}$ be a noisy quantum circuit acting on $\mathcal{H}_2^{\otimes n}$. The noise incurred on the circuit is specified by a *noise model* \mathcal{N} , the form of which is defined in Sect. 3. The data subsystem is initialized in the state $|0\rangle^{\otimes n}$. The input $x \in \{0, 1\}^k$ is then encoded in a quantum error-correcting code of length n ; we denote the encoded input by $|\bar{x}\rangle$. Execution then proceeds the same way as in the ideal circuit, except that the gates are replaced by encoded gates (i.e., operations on encoded data). We can thus compare the states of \mathcal{C} and $\tilde{\mathcal{C}}$ at an arbitrary step of the computation. Additionally, after each step of the computation, error correction is performed in order to keep the state of the circuit in the encoded space of the code. We say that $\tilde{\mathcal{C}}$ is ε -reliable against noise \mathcal{N} if for every step, the state of \mathcal{C} is equal to the logical value of the state of $\tilde{\mathcal{C}}$ except with probability ε .

We note that the way we use fault tolerance is slightly different from the usual treatment. As defined above (and as it is commonly done), the inputs are encoded as part of the circuit. But our circuits also receive additional encoded inputs. This is done purely for convenience and does not make the definition stronger.

The focus in fault tolerance is in implementing the so-called *gadgets*—components such as logical gates and error correction, that can then be used as building blocks to construct reliable circuits. The goal is usually to implement a set of gates that is universal for quantum computation. However, as we will see in Sect. 5, since we only seek to perform classical computation, we only need a restricted set of gates.

3. Leakage Models and Quantum Noise

In this section, we show that for an arbitrary classical circuit with leakage according to some leakage model L , we can view the circuit as a noisy quantum circuit with a corresponding noise model \mathcal{N} , and that if the quantum circuit is reliable with respect to \mathcal{N} then it is also leakage-resilient against L . To this end, we must first make matters more concrete. A general quantum noise model is an operation \mathcal{N} on quantum states on $\mathcal{H}_2^{\otimes n}$ and takes the form

$$\mathcal{N}(\rho) = \sum_k p_k E_k \rho E_k^\dagger, \quad (3)$$

where E_k are arbitrary operators taking $\mathcal{H}_2^{\otimes n}$ to itself and $p_k \geq 0$ with $\sum_k p_k = 1$. If E_k has the form $E_k = \bigotimes_{i=1}^n Z_i^{a_{i,k}}$ for $a_{i,k} \in \{0, 1\}$ and Z specified by $Z|x\rangle = (-1)^x|x\rangle$, then we call \mathcal{N} a *phase noise* model.

The following lemma relates classical leakage models and quantum noise models.

Lemma 1. *For any circuit \mathcal{C} running with a leakage request $\{l_j, p_j\}_{j=1}^m$ taken from a leakage model L , there exists a noise model \mathcal{N} such that \mathcal{C} , viewed as a quantum circuit, is a quantum circuit subject to noise \mathcal{N} . Furthermore, \mathcal{N} has the form*

$$\mathcal{N}(\rho) = \frac{1}{d} \sum_{j=1}^m p_j \sum_{k=0}^{d-1} F_j^k \rho (F_j^k)^\dagger, \quad (4)$$

where F_j^k is the operator taking $|s\rangle$ to $\omega^{kl_j(s)}|s\rangle$, ω is a primitive d th root of unity, and d is the size of the output of l_j .

Now we can state the formal connection between leakage resilience and reliable quantum computation:

Theorem 1. *For every circuit \mathcal{C} and leakage model L , there exists a noise model \mathcal{N} as specified in Lemma 1 such that for any ε -reliable (against \mathcal{N}) implementation $\tilde{\mathcal{C}}$ of \mathcal{C} with encoding function $y \mapsto |\bar{y}\rangle$, the protocol π that, given (y, \mathcal{C}) as input, outputs $(|\bar{y}\rangle, \tilde{\mathcal{C}})$ is a $2\sqrt{\varepsilon}$ -leakage-resilient compiler against L .*

Proof of Lemma 1. First, consider the case where L is such that at each round, a single leakage function l is chosen. Let S be the subsystem representing the wire assignments $s := W_{\mathcal{C}}(x, y)$ in the circuit and E be Eve's subsystem. In quantum-mechanical terms, the action of the leakage is the transformation

$$|s\rangle^S \otimes |0\rangle^E \rightarrow |s\rangle^S \otimes |l(s)\rangle^E \quad (5)$$

for each s . To determine the action of the leakage on system S itself, consider an arbitrary superposition state $|\psi\rangle^S = \sum_s \sqrt{p_s} |s\rangle^S$ for some probability distribution p_s . After applying the transformation and tracing out subsystem E , $\rho = |\psi\rangle\langle\psi|$ becomes

$$\rho \rightarrow \mathcal{N}(\rho) = \sum_s \sum_{s': l(s')=l(s)} \sqrt{p_s p_{s'}} |s\rangle\langle s'|, \quad (6)$$

as all coherence is lost between parts of the state with different values of l . But this is also the output state if the transformation were instead

$$|s\rangle^S \otimes |0\rangle^E \rightarrow \frac{1}{\sqrt{d}} \sum_k \omega^{kl(s)} |s\rangle^S \otimes |k\rangle^E = \frac{1}{\sqrt{d}} \sum_k F_j^k |s\rangle^S \otimes |k\rangle^E, \quad (7)$$

where d is the size of the output of l , ω a primitive d th root of unity, and F_j^k the operator taking $|s\rangle$ to $\omega^{kl(s)}|s\rangle$. The equivalence of these two maps can be seen by tracing out E .

Thus, the noise model can be expressed as

$$\mathcal{N}(\rho) = \frac{1}{d} \sum_{k=0}^{d-1} F^k \rho (F^k)^\dagger. \quad (8)$$

Now consider the case where the leakage function is chosen probabilistically from a set $\{l_j\}_{j=1}^m$, where l_j is chosen with probability q_j . In this case, a corresponding noise model is a convex combination of noise operators of the form in (8). That is, we can write

$$\mathcal{N}(\rho) = \frac{1}{d} \sum_{j=1}^m q_j \sum_{k=0}^{d-1} F_j^k \rho (F_j^k)^\dagger \quad (9)$$

where F_j^k is the operator taking $|s\rangle$ to $\omega^{kl_j(s)}|s\rangle$. \square

Proof of Theorem 1. Every leakage request from L has a corresponding noise model as given by Lemma 1. The fault-tolerant implementation of the circuit \mathcal{C} includes a compiled circuit $\bar{\mathcal{C}}$ as well as a method to encode inputs. The converter π_A receives the secret y as well as a circuit as input, and outputs the encoded secret $|\bar{y}\rangle$ and $\bar{\mathcal{C}}$. Following the fault-tolerant construction, inputs x received from E are encoded in $|\bar{x}\rangle$, the quantum circuit is used to compute $|\bar{\mathcal{C}}(x, y)\rangle$, and the decoded output is sent back to E . We therefore have a scenario as shown in Fig. 1b.

We now have to show the existence of a simulator σ such that condition (1) (Fig. 2) is satisfied. We will prove that such a simulator exists by showing that for any step of the computation, the leakage received by E is essentially independent of the circuit's current state (that is, the intermediate values encoded by its wires).

Consider an arbitrary timestep in the circuit. Ideally, the evolution of the input systems up to this point is described by a unitary operation \mathcal{U} , the product of all the unitary operators for the individual gates. By Lemma 1, the leakage model results in noise \mathcal{N} on the quantum circuit \mathcal{C} . The actual evolution of gates and noise can be described by a quantum operation \mathcal{T} . Since \mathcal{C} is ε -reliable against \mathcal{N} , there can be no errors on encoded quantum information except with probability ε . (Otherwise, the state of the reliable circuit would be different from the state in the ideal circuit.) Put differently, concatenating the actual evolution with the inverse of the ideal evolution should result in an identity operation on the input. This can be formalized using the completely bounded norm (see [16] for a precise definition) as $\|\mathcal{T}\mathcal{U}^{-1} - \mathbb{1}\|_{\text{cb}} < \varepsilon$.

The actual evolution \mathcal{T} is due to leakage to the eavesdropper, which itself can be described by a quantum operation, call it \mathcal{T}' . Unlike for classical channels, the properties of \mathcal{T} constrain the properties of \mathcal{T}' . In particular, Theorem 3 of [16] shows that if \mathcal{T} is approximately unitary, then \mathcal{T}' must have an output which is approximately independent of the input.⁴ Specifically, there exists an operation \mathcal{T}_S which takes every input to some fixed output such that $\|\mathcal{T}\mathcal{U}^{-1} - \mathbb{1}\|_{\text{cb}} < \varepsilon$ implies $\|\mathcal{T}' - \mathcal{T}_S\|_{\text{cb}} \leq 2\sqrt{\varepsilon}$. Therefore,

⁴It would also be possible to use an uncertainty principle recently derived by one of us for this argument [23].

by the properties of the completely bounded norm, the state at any point in the circuit, which we denote by $|s\rangle$, is independent of $l(W_C(x, y))$ (except with probability $2\sqrt{\varepsilon}$).

We now note that for the case of an adaptive adversary, l' can depend on previous inputs and outputs, as well as previous leakages, since the choice of l can depend on those values. But because those values are also available to the simulator, and because l' is independent of the logical value of s , l' can be generated by the simulator.

Now we describe the converter π_E , used in the case where Eve is honest [condition (2) in Definition (1)]. π_E relays inputs x from Eve to \mathcal{R} and decodes the encoded output $\overline{C(x, y)}$, which is then passed to Eve. It also decodes the encoded circuit \mathcal{C} before passing it to Eve. It is clear that $\pi_A \pi_E \mathcal{R}$ (which can be visualized as plugging π_E to $\pi_A \mathcal{R}$ in Fig. 2) is indistinguishable from \mathcal{S} . \square

4. From Quantum to Classical Circuits

Our goal is to have *classical* leakage resilience, and fault-tolerant techniques yield a circuit that in general does not have a classical translation. In this section, we show how to make a particular set of quantum components classical. Then, in Sect. 5, we use these components to make a fault-tolerant implementation of a quantum gate that is universal for classical computation—namely, the Toffoli gate. Due to the fact that the components have a classical translation, we also have a leakage-resilient classical gate.

In order to show the equivalence between the quantum and classical circuits, we analyze each component in the following scenario: We assume that the quantum component has Z basis inputs (i.e., the inputs are classical), and that after execution, the outputs are measured in the Z basis. We then show that for each component there exists a classical circuit that, when given the same inputs, gives the same outputs as the quantum component after measuring. We then use the components to construct encoded *gadgets* that will be used to implement the Toffoli gate. Since the components have a classical translation, the gadgets and the Toffoli gate have one as well.

In order to be able to combine the classical translation with Theorem 1, we need it to preserve the form of the original quantum circuit, so that the connection between leakage and noise model can be made explicitly. This is achieved by having, for each component, a classical translation that has the same wires as the quantum component.

For the classical scenario, we assume that we can generate random bits in a leak-free manner. We will see that this is needed in order to make state preparation of $|+\rangle$ classical. We prove the following result.

Theorem 2. *Let \mathcal{C} be a quantum circuit accepting classical states as input and containing only X , Z , $CNOT$, CZ and Toffoli gates, state preparation of $|0\rangle$ and $|+\rangle$ and measurement in the X and Z bases. Then each gate G in the circuit can be replaced by a classical circuit G' with the same wires as G . Furthermore, assuming a leak-free source of random bits, there exist procedures for state preparation of $|0\rangle$ and $|+\rangle$ and measurement in the X and Z bases such that the circuit for each procedure can be replaced by a classical circuit with the same wires. If all these replacements are made, then the resulting classical circuit \mathcal{C}' gives the same outputs as applying \mathcal{C} followed by measuring its outputs in the Z basis.*

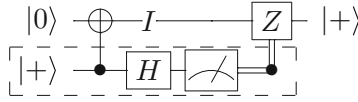


Fig. 3. Preparation of $|+\rangle$.

Proof. Out of the gates we use, X , CNOT and Toffoli are classical, and therefore their classical translations are trivial. The Z gate flips the input's phase; as we are assuming the inputs and outputs to be classical, it corresponds classically to the identity gate. The same holds for CZ .

There are two bare qubits we need to prepare: $|0\rangle$ and $|+\rangle$. The classical translation of preparing a $|0\rangle$ is just preparation of 0. To determine the translation of preparing $|+\rangle$, note that immediate measurement in the Z basis would yield a random bit. We can thus translate state preparation of $|+\rangle$ to the classical scenario by preparing 0 and then adding a random bit r to it, which we assume to be generated in a leak-free manner, so that it is hidden from the adversary.

The output does not correspond univocally to $|+\rangle$; if we had prepared $|-\rangle$, we would get the same output. But we can show that from the point of view of a (classical) adversary that can see all the wires in the circuit, except for the leak-free component, this preparation procedure is equivalent to the one that prepares $|+\rangle$, shown in Fig. 3.

The second wire, shown inside the dashed box, is in the leak-free part of the circuit. The state at the point I is $|+\rangle|+\rangle|-\rangle|-\rangle$; this is the state used in the actual (classical) circuit. The figure shows how we could then correct this state using the leak-free component, so that in the end we get $|+\rangle$. Since the correction operation has no effect on the Z -basis value, it can be omitted. Therefore, if we use the circuit in Fig. 3 for preparation of $|+\rangle$ in the circuit \mathcal{C} , the classical translation amounts to initializing a register to 0 and then adding a (leak-free) random bit to it.

Measurement in the X basis can be done in a similar way. Measuring X projects the state $|\psi\rangle$ onto one of the operator's eigenspaces; the projection operators are given by $P_i = \frac{1}{2}(\mathbb{1} + (-1)^i X)$ for $i \in \{0, 1\}$ and the post-measurement state is given by $\rho = \sum_{i=0}^1 P_i |\psi\rangle\langle\psi| P_i = \frac{1}{2} \sum_{i=0}^1 X^i |\psi\rangle\langle\psi| X^i$. Hence, measuring in the X basis is equivalent to a random bit flip, which we can simulate in the classical circuit by adding a random bit. \square

The significance of Theorem 2 is that we can use the given operations to implement a reliable Toffoli gate for *independent (phase) noise*, in which every wire in the circuit is subject to a phase error, an unwanted Z operator, with a fixed probability. We will see next that this corresponds to the independent leakage model.

5. Leakage-Resilient Gadgets

Our fault-tolerant construction follows [1], which works in the model of independent phase noise where each wire in the circuit is subject to a phase error with probability p . This phase noise model is related to the independent leakage model, where each wire in

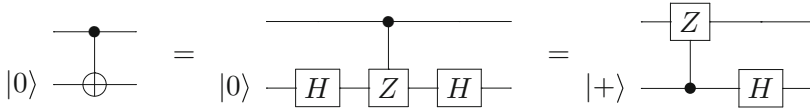


Fig. 4. Leaking one bit (encoded in the top wire) is equivalent to introducing a random phase error.

the circuit leaks with probability $2p$. This follows because leaking the state of one wire is equivalent to introducing a random phase error, as seen in Fig. 4.

This relation also provides a more direct relation between fault tolerance and leakage resilience than the one given in Theorem 1: No bits can leak if all the errors are corrected, and therefore, an ε -reliable circuit is 2ε -leakage-resilient.

In order to construct a leakage-resilient compiler, we just need to implement a set of fault-tolerant gates that is universal for classical computation. We only use the components in the statement of Theorem 2, so that we are able to make the implementation classical. A typical approach is to implement the NAND gate. However, quantum gates are reversible; hence, it is easier to implement a universal set of reversible gates. We have chosen the Toffoli gate, defined by $T(a, b, c) = (a, b, ab \oplus c)$, which is universal [30].

The construction we develop in the remainder of this section only uses the components specified in Theorem 2. By this theorem, for each component there exists a classical translation with the same wires as the original quantum component, and therefore we can apply the relationship between independent phase noise and independent leakage explained above directly (or, in general, we could apply Lemma 1). We have the following result.

Corollary 1. *Let L be the independent leakage model, and assume the probability of leakage p satisfies $p < 10^{-5}$. For any $\varepsilon > 0$ there exists a ε -leakage-resilient compiler π that, given an arbitrary reversible classical circuit C with l locations and depth d as input, outputs a classical circuit \bar{C} with $l' = O(l \text{ polylog}(l/\varepsilon))$ and depth $d' = O(d \text{ polylog}(l/\varepsilon))$.*

This corollary follows from the implementation of the Toffoli gate and the accuracy threshold theorem of [1], along with Theorem 2 and the relationship between independent phase noise and independent leakage.

Given the construction for the Toffoli gate, we have a method of compiling an arbitrary circuit into a private one that works as follows. First, we convert the circuit into one using only Toffoli gates. As we will see, our fault-tolerant implementation of the Toffoli gate involves only X , Z and CNOT gates, state preparation, measurement in the X and Z bases, and error correction. Hence, it has a classical translation due to Theorem 2. Furthermore, by Theorem 1, the resulting circuit is leakage-resilient against independent leakage.

The construction in [1] works by encoding qubits in the Steane [[7, 1, 3]] code [28]. Fault tolerance is achieved by constructing encoded gadgets that are resilient to errors “in first order”—that is, if the physical circuit has probability of failure ε , then the encoded

circuit has probability of failure $O(\varepsilon^2)$. One can then achieve an arbitrary degree of accuracy by code concatenation.

In what follows, we present the fault-tolerant gadgets we use in the construction of the Toffoli gate.

Measurement A Z basis measurement can be done by measuring transversally. For the Steane code, $|\bar{0}\rangle$ is an equal superposition of all the even-weight codewords of the Hamming code, and $|\bar{1}\rangle$ is an equal superposition of the odd-weight codewords. (Section A.3 has a review of the Steane code.) Thus, we can determine which state was prepared by computing the parity of the measurement outcomes. Since phase errors do not affect the outcomes, the procedure is fault-tolerant.

X basis measurement can thus be done by applying $\bar{X} = X_1 X_2 X_3$ controlled by $|+\rangle$ and then measuring in the Z basis. Phase errors only propagate from the data to the ancilla $|+\rangle$ state, which, as argued in the proof of Theorem 2, can be assumed to be prepared perfectly, since classically they correspond to generating a random bit. This ensures that the procedure is fault-tolerant.

Error correction Error correction consists of syndrome extraction, in which the errors are diagnosed, and a recovery step, performed in order to transform the state back into the correct one. First, notice that since we are performing computation on Z basis states and we assume that only phase noise is possible, the recovery step would always consist of applying Z operators and, by the remarks in Theorem 2, can never change the outcome of the computation. Therefore, only the syndrome extraction step is necessary. For that we use a method known as Steane error correction [10, 29], for which syndrome extraction reduces to X basis measurement.

CNOT and X gates For the Steane code, the X gate can be applied transversally, that is, $\bar{X} = \otimes_{i=1}^7 X_i$. The CNOT gate can also be easily seen to be transversal, in the sense that every qubit in the first block only interacts with the corresponding qubit in the second block.

State preparation We first describe the state preparation of $|\bar{0}\rangle$. It is accomplished by preparing 7 $|0\rangle$ states and then performing Steane (phase) error correction. This method consists of taking the state $|0\rangle^{\otimes 7}$ and then performing a CNOT gate with this state as target and an ancilla state $|\bar{0}\rangle$ as control. The ancilla is then measured in the Z basis. This, of course, has an obvious problem—in order to execute the circuit, we need $|\bar{0}\rangle$, which is exactly the state we are trying to prepare. But before addressing this issue, let us see why it works. The outcomes of the measurement of $|\bar{0}\rangle$ after the transversal CNOT with $|0\rangle^{\otimes 7}$ determine the eigenvalues of the X -type stabilizers of the code. The original state $|0\rangle^{\otimes 7}$ is projected onto the subspace associated with the particular measurement outcomes. We can then map the state into the code space by changing the eigenvalues as needed. But in fact recovery is not necessary, since it could only change the phase of the state. As long as we have the information about the eigenvalues, we might as well adopt the resulting state as $|\bar{0}\rangle$.

In order to prepare the ancilla $|\bar{0}\rangle$, we use the circuit shown in [1, Fig. 12]. We note that, since the circuit acts on physical qubits, it is not fault-tolerant. The usual way to prepare an ancilla state fault-tolerantly is to perform a verification step after encoding, where the state is rejected and the procedure is repeated if the verification detects too

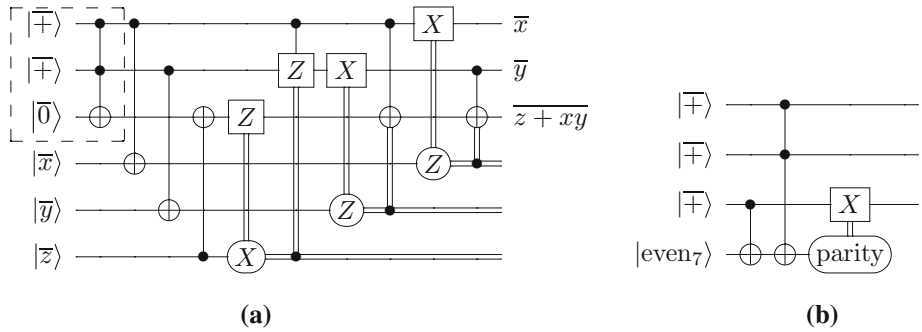


Fig. 5. a Reliable Toffoli gate. **b** Construction of the Toffoli ancilla state. All the gates are executed transversally. After computing the parity, the circuit (Fig. 6b) is executed on the bottom block.

many errors. However, this does not help us because there is no clear classical analogue; the circuit would not “know” when to reject a state, since phase errors do not show up in the classical picture.

Instead, we use the ancilla verification method developed in [7]. In this method, the ancilla state is prepared, interacts with the data transversally and is subsequently decoded. The decoding is done in such a way that errors due to a single fault in the ancilla preparation can be perfectly distinguished and the data block can be corrected. In our case, we can decode by measuring the phase-error syndrome, which for the Steane code reduces to X basis measurement.

Now we can use $|\bar{0}\rangle$ along with $|+\rangle$ to prepare $|\bar{+}\rangle$: We prepare $|\bar{0}\rangle$ and then apply \bar{X} controlled by a random bit. As we have argued for X basis measurement, the circuit is fault-tolerant.

The method used to prepare $|\bar{0}\rangle$ can also be used to prepare the “Shor state” $|\text{even}_7\rangle$, which is a superposition of all the even-weight words in $\{0, 1\}^7$ and is used in the construction of the Toffoli ancilla state (below). Our method to prepare and verify $|\text{even}_7\rangle$ is the same as the one presented in [7] for the cat state, except it is done in the rotated basis ($|\text{even}_7\rangle$ is obtained from the cat state by applying the Hadamard gate transversally). For more details, see “Appendix B.”

Toffoli gate The gadget for the Toffoli gate, proposed by Shor [24] (see also [22]) is shown in Fig. 5a. The correctness of the circuit can be verified by inspection. However, we have not shown how to execute the subcircuit in the dashed rectangle; indeed, this subcircuit uses a Toffoli gate, which is exactly what we are trying to implement. Instead, we use an alternative circuit to prepare the state $|\bar{\Theta}\rangle = |\bar{000}\rangle + |\bar{100}\rangle + |\bar{010}\rangle + |\bar{111}\rangle$, which is the output of the subcircuit in the dashed rectangle. The alternative circuit is shown in Fig. 5b.

All the gates in Fig. 5b are executed transversally, so the circuit is fault-tolerant.⁵ We will show that if transversal CNOT and Toffoli work as the “encoded” circuits in this

⁵Because the Toffoli gate is not a Clifford gate, it does not propagate Pauli errors to Pauli errors; in particular, a phase flip in the third block is propagated to a superposition of phase flips. This would make a classical simulation of the level- k circuit inefficient, because the error patterns to keep track of grow exponentially in k , but again, this is not a problem for our construction because we do not need to perform error recovery.

case, where $|\text{even}_7\rangle$ encodes 0 and the “odd” state $|\text{odd}_7\rangle$ encodes 1, then the circuit prepares $|\bar{\odot}\rangle$. Specifically, we need to check that output in the target of the CNOT is $|\text{even}_7\rangle$ if the control is $|\bar{0}\rangle$ and $|\text{odd}_7\rangle$ if the control is $|\bar{1}\rangle$, and that the Toffoli gate outputs $|\text{odd}_7\rangle$ if and only if both control qubits are $|\bar{1}\rangle$, and $|\text{even}_7\rangle$ otherwise. For the CNOT, $|\bar{0}\rangle$ is an equal superposition of even-weight strings, so XORing them with even-weight strings outputs even-weight strings. $|\bar{1}\rangle$ is an equal superposition of odd-weight strings, and XORing them with even-weight strings produces odd-weight strings.

For the Toffoli gate, it suffices to prove that the transversal product of $|\bar{x}\rangle$ and $|\bar{y}\rangle$ is a superposition of odd-weight strings if and only if $x = y = 1$, and is a superposition of even-weight strings otherwise. The key to show this claim is that the Steane code is based on a classical code \mathcal{C} that contains its own dual \mathcal{C}^\perp . Let H be the parity-check matrix of the code \mathcal{C} . We have $H\mathbf{c} = 0$ for every $\mathbf{c} \in \mathcal{C}$. That is, each codeword is orthogonal to every row H_i of H . Since $\mathcal{C}^\perp \subset \mathcal{C}$, this also applies to the codewords of \mathcal{C}^\perp , and in particular to H_i . That is, each H_i is orthogonal to itself, which can only happen if it has even weight. Hence, all the codewords in \mathcal{C}^\perp have even weight. Furthermore, they must have even overlap (i.e., have 1 in the same position) with every H_i (otherwise they would not be orthogonal), and therefore also among themselves.

Thus, $|\bar{0}\rangle$ is the superposition of all the even-weight codewords of \mathcal{C} , which have even overlap among themselves by the argument above. The state $|\bar{1}\rangle$ is the superposition of the codewords in $\mathcal{C} - \mathcal{C}^\perp$, which all have odd weight. Since they also satisfy the parity check H they have even overlap with the rows of H , and therefore also with the codewords in $|\bar{0}\rangle$. But they have odd overlap among themselves. To prove this, let $x, y \in \mathcal{C} - \mathcal{C}^\perp$. If x and y have even overlap, then $x + y \in \mathcal{C}^\perp$ (since it has even weight) and thus x and $x + y$ would have even overlap. This proves our claim.

Thus, we have a way of preparing the ancilla state $|\bar{\odot}\rangle$ fault-tolerantly. Furthermore, the classical translation is easy: The only new component here is measurement of the parity of the Shor state, which can be done by measuring in the Z basis and then adding up the outcomes. Now every component in Fig. 5a has a classical translation.

6. Discussion

Theorem 1 establishes an explicit relationship between fault-tolerant quantum computation and classical leakage resilience. Although it relates leakage and noise models (via Lemma 1), it is not clear how the properties of the noise model relate to the properties of the leakage model in general. However, in some cases the leakage model resulting from a given noise model has a simple interpretation. In this work, we analyzed the independent leakage model, which corresponds to the independent phase noise model. A possible further direction is to take a leakage model that is used in other leakage resilience proposals and try to understand the corresponding noise model. Conversely, one could take, say, the “local noise” model of [1] and see what the corresponding leakage model looks like.

We developed a concrete implementation of universal leakage-resilient computation based on the fault-tolerant construction of [1]. This construction works in the independent noise model by using the concatenated Steane code. Fault tolerance is achieved assuming

probability of error per wire or gate $p < 10^{-5}$ (the accuracy threshold). Along with our results, this gives us a leakage-resilient construction for universal classical computation assuming independent leakage with probability of leakage per wire $p < 2 \cdot 10^{-5}$.

Taking into account the fact that we only want leakage resilience rather than fault tolerance, this threshold can be improved. An accuracy threshold for concatenated codes depends essentially on the size of the largest gadget in the first level of encoding (a level-1 extended rectangle or 1-exRec in the language of [1]). A lower bound for the threshold is the reciprocal of the number of pairs of locations in the largest 1-exRec. In the implementation of [1], the largest 1-exRec is the CNOT gate, but due to the simplifications in our case, the largest extended rectangles are the gadgets for state preparation: They have 20 locations each. The number of pairs of locations is then $\binom{20}{2} = 190$, which gives a crude leakage threshold estimate of $p \approx 1\%$.

As in [1], we only use the Steane code, but we note that our construction works for any CSS code based on a dual-containing classical code. A promising possibility is to use color codes [5], for which numerical evidence suggests they have good accuracy thresholds [17] but currently lack a rigorous lower bound on the threshold.

We note that while the independent leakage model looks similar to the “noisy leakage” case of [9], they are in fact different models. In the noisy leakage model, all bits leak, but the adversary only receives noisy versions of these bits, each of them being flipped with probability p . Crucially, the adversary does not know which bits have leaked faithfully, whereas in our model every bit that is leaked is sure to arrive correctly at the adversary.

While we do not expect existing results on fault tolerance to give direct constructions for the leakage models commonly studied in the literature (especially since fault tolerance has only been shown to work in very few noise models), we note that quantum fault tolerance is a stronger requirement than classical leakage resilience: As we have seen in this work, translating from the former to the latter allows us to make several simplifications. Additionally, the techniques used in the area of fault tolerance are different from those used for leakage resilience. Hence, we expect our result to shed new light on leakage resilience.

Acknowledgements

The authors would like to thank Christopher Portmann for helpful comments on the abstract cryptography framework. This work was supported by the Swiss National Science Foundation (through the National Centre of Competence in Research “Quantum Science and Technology” and Grant No. 200020-135048), by the European Research Council (Grant No. 258932) and by CNPq/Brazil.

Appendix A: Qubits and Stabilizers

A quantum system can be described by a complex Hilbert space. In this work, we deal only with two-level systems. Let \mathcal{H}_2 be a two-dimensional complex Hilbert space, and let $\{|0\rangle, |1\rangle\}$ denote one orthonormal basis, which we call the standard basis or the

computational basis. An arbitrary element $|\psi\rangle$ of \mathcal{H}_2 can be written as $|\psi\rangle = \alpha|0\rangle + \beta|1\rangle$, where $\alpha, \beta \in \mathbb{C}$. Quantum states are represented by unit vectors, that is, $|\alpha|^2 + |\beta|^2 = 1$.

Composite systems are given by the tensor product, that is, if A, B are two complex Hilbert spaces, then the state of the system composed by A and B is an element of $A \otimes B$.

Systems whose state is not completely known are described properly by *density operators*. Suppose a system is in the state $|\psi_i\rangle$ with probability p_i . We define the density operator ρ for the system by

$$\rho = \sum_i p_i |\psi_i\rangle\langle\psi_i|$$

States that are completely known, that is, states of the form $\rho = |\psi\rangle\langle\psi|$, are called *pure states*, while other states are called *mixed states*.

In the following, as well as in the rest of the text, we only deal with states that are equal superpositions, so we omit the normalizing factor. So, for instance, we write the state $|-\rangle = \frac{1}{\sqrt{2}}(|0\rangle - |1\rangle)$ simply as $|0\rangle - |1\rangle$.

A.1. Pauli Operators

Take the standard basis for \mathcal{H}_2 , and let

$$\mathbb{1} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \quad X = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \quad Y = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix} \quad Z = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}$$

X, Y, Z are called the *Pauli matrices*. They anticommute with each other. Furthermore, an arbitrary matrix A on \mathcal{H}_2 can be written as $A = a_0\mathbb{1} + a_1X + a_2Y + a_3Z$ with $a_0, a_1, a_2, a_3 \in \mathbb{C}$. Similarly, operators on $\mathcal{H}_2^{\otimes n}$ can also be written in terms of Pauli matrices. Let \mathcal{P}_n be the set of operators of the form $i^k \bigotimes_{i=1}^n P_i$, where $k \in \mathbb{Z}$ and $P_i \in \{\mathbb{1}, X, Y, Z\}$. The set \mathcal{P}_n is a non-abelian group; it is called the *Pauli group* on n qubits. Because $Y = iXZ$, the group is generated by X and Z (up to a phase factor); that is, $\mathcal{P}_n = \langle i\mathbb{1}, X_1, \dots, X_n, Z_1, \dots, Z_n \rangle$, where X_i denotes the operator that acts as X on the i th qubit, and similarly for Z_i . An arbitrary matrix on $\mathcal{H}_2^{\otimes n}$ can be written as a linear combination of elements of \mathcal{P}_n .

Consider now a system S subject to noise from the environment E . By the above discussion, the system evolves as

$$|\psi\rangle_S |0\rangle_E \rightarrow \sum_k E_k |\psi\rangle_A |e_k\rangle_E$$

where the states $|e_k\rangle$ are not necessarily orthogonal and $\{E_k\}$ is a set of linearly independent Pauli operators. We call each of them a *Pauli error*. We say a Pauli error E_k has weight t if it acts non-trivially on at most t qubits. Because general errors can always be decomposed into Pauli errors, we only need to design error-correcting codes that can correct Pauli errors.

A.2. Error-Correcting Codes

A quantum error-correcting code \mathcal{C} of length n is a subspace of $\mathcal{H}_2^{\otimes n}$. Let \mathcal{E} be a set of errors. We say that \mathcal{C} corrects \mathcal{E} if there exists a recovery operator R acting on a larger system $\mathcal{C} \otimes A$ such that for every $|\psi\rangle \in \mathcal{C}$ and every $E \in \mathcal{E}$, we have $\text{tr}_A R(E|\psi\rangle_{\mathcal{C}}|a\rangle_A) = |\psi\rangle_{\mathcal{C}}$, where $|a\rangle$ is some ancilla state in A .

A code that corrects all errors of weight t can *detect* all errors of weight $2t$. We define the *distance* of a code as the weight of the error of smallest weight that is not detectable. Thus, a code that can correct t errors has distance $d = 2t + 1$. Analogously to the classical case, we call an error-correcting code of weight n that encodes k qubits and has distance d an $[[n, k, d]]$ code.

Now let \mathcal{S} be a subgroup of the Pauli group \mathcal{P}_n not containing $-\mathbb{1}$. Given such a subgroup, we define a *stabilizer code* \mathcal{C} of length n as

$$\mathcal{C} = \{|\psi\rangle \in \mathcal{H}_2^{\otimes n} \mid s|\psi\rangle = |\psi\rangle, \forall s \in \mathcal{S}\}.$$

The group \mathcal{S} is called the code's stabilizer. If the code encodes k qubits, the stabilizer is generated by $n - k$ elements. Errors that are not detectable are in the centralizer $\mathcal{Z}(\mathcal{S})$, the group of operators that commute with all elements of \mathcal{S} . Since errors act non-trivially on the codewords, they are not in \mathcal{S} ; thus, the distance of the code is given by the weight of the operator in $\mathcal{Z}(\mathcal{S}) - \mathcal{S}$ with smallest weight.

Since \mathcal{P}_n has dimension $2n$, the centralizer has dimension $2n - (n - k) = n + k$. Therefore, $\mathcal{Z}(\mathcal{S}) - \mathcal{S}$ has dimension $2k$. These operators can be regarded as *logical operations* on the codewords. We can always choose them to be the operators $\overline{Z}_1, \dots, \overline{Z}_k, \overline{X}_1, \dots, \overline{X}_k$, satisfying the anticommutation relation $\overline{Z}_i \overline{X}_j = (-1)^{\delta_{ij}} \overline{X}_j \overline{Z}_i$.

Stabilizer generators are the quantum analogue of rows in the parity-check matrix of a classical code. In fact, there is a general class of codes known as *CSS codes* that are constructed from two classical codes C_1, C_2 such that $C_2^\perp \subset C_1$. We will not introduce the general theory of CSS codes here, and instead concern ourselves with a particular CSS code known as the Steane code.

A.3. The Steane Code

The Steane code [28] is a $[[7, 1, 3]]$ code that has stabilizer generators

$$\begin{aligned} g_1 &= Z_1 Z_3 Z_5 Z_7 \\ g_2 &= Z_2 Z_3 Z_6 Z_7 \\ g_3 &= Z_4 Z_5 Z_6 Z_7 \\ g_4 &= X_1 X_3 X_5 X_7 \\ g_5 &= X_2 X_3 X_6 X_7 \\ g_6 &= X_4 X_5 X_6 X_7 \end{aligned}$$

and logical operators

$$\overline{Z} = Z_1 Z_2 Z_3$$

$$\overline{X} = X_1 X_2 X_3$$

The Steane code is a CSS code based on the classical [7, 4, 3] Hamming code, which we denote here by C . The Hamming code contains its own dual, that is, $C^\perp \subset C$. The stabilizer generators g_1, g_2, g_3 and g_4, g_5, g_6 for the Steane code correspond to the rows in the parity-check matrix for C , where a 0 in the parity-check matrix corresponds to the identity and 1 corresponds to Z (for the Z -type stabilizers) or X (for the X -type stabilizers). The codewords are given by

$$\begin{aligned} |\overline{0}\rangle &= \sum_{x \in C^\perp} |x\rangle = |0000000\rangle + |0001111\rangle + |0110011\rangle + |1010101\rangle \\ &\quad + |0111100\rangle + |1011010\rangle + |1100110\rangle \\ |\overline{1}\rangle &= \sum_{x \in C - C^\perp} |x\rangle = |1111111\rangle + |1110000\rangle + |1001100\rangle + |0101010\rangle \\ &\quad + |1000011\rangle + |0100101\rangle + |0011001\rangle. \end{aligned}$$

Notice that $|\overline{0}\rangle$ is an equal superposition of all the even-weight Hamming codewords, and $|\overline{1}\rangle$ is an equal superposition of all the odd-weight ones. This comes from the fact that $C^\perp \subset C$, which implies several properties that make the Steane code useful for fault tolerance. In particular, the logical Hadamard gate can be implemented transversally, that is, $\overline{H} = H^{\otimes 7}$. This fact is also crucial in the construction we use for the Toffoli ancilla state, used to implement the Toffoli gate.

A.4. Quantum Gates and Measurement

X and Z gates correspond to the application of the respective Pauli matrices. CZ and $CNOT$ are the controlled versions of these gates, that is, they are linear operators with

$$\begin{aligned} CNOT(|a\rangle, |b\rangle) &= (|a\rangle, |a + b\rangle) \\ CZ(|a\rangle, \alpha|0\rangle + \beta|1\rangle) &= (|a\rangle, \alpha|0\rangle + \beta(-1)^a|1\rangle) \end{aligned}$$

Measurements are operations that take a quantum state as input and have a classical outcome. A measurement in the Z basis of an arbitrary qubit $\alpha|0\rangle + \beta|1\rangle$ returns 0 with probability $|\alpha|^2$ and 1 with probability $|\beta|^2$. A measurement in the X basis works similarly, by writing the state in the X basis: Measuring $\alpha|+\rangle + \beta|-\rangle$ gives 0 with probability $|\alpha|^2$ and 1 with probability $|\beta|^2$.

Appendix B: Construction of the Shor State

The circuit to construct the Shor state, adapted from the cat state construction of [7], is shown in Fig. 6a. After it interacts with the data, we measure its syndrome using the circuit in Fig. 6b in order to diagnose a possible fault in the preparation.

We need to show that no single error in the encoding and decoding of the Shor state can cause us to incorrectly apply a multi-qubit correction operation to the data. Incorrect single-qubit corrections, in contrast, are not fatal, as they can be corrected in subsequent

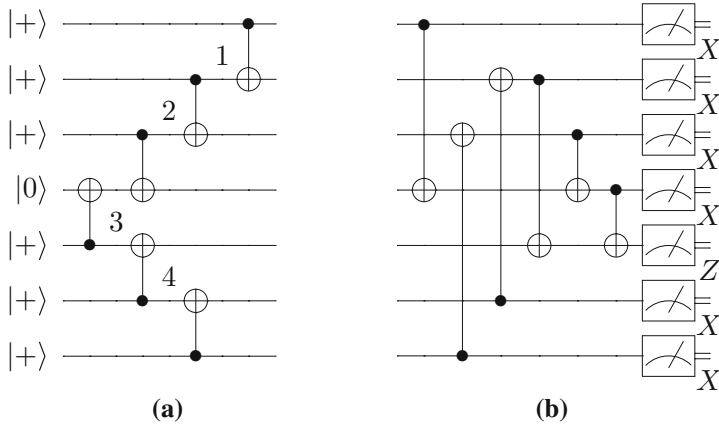


Fig. 6. **a** Preparation of the Shor state $|even_7\rangle$. The numbers indicate positions at which a local error causes errors on multiple output wires. **b** Syndrome measurement for the Shor state .

rounds of error correction. First, let us examine the possible multiple Z error patterns resulting from a single fault that might occur during encoding. Those are shown by the numbers in Fig. 6a. All the other possibilities lead either to a single error in the output, or to the same error pattern as one of those.

The error patterns, in the order given by the figure, are Z_1Z_2 , $Z_1Z_2Z_3$, $Z_5Z_6Z_7$ and Z_6Z_7 . The decoding circuit (Fig. 6b) propagates these errors to $Z_1Z_2Z_6$, $Z_1Z_2Z_3Z_6Z_7$, $Z_2Z_4Z_6Z_7$ and Z_6Z_7 , respectively. Thus, all possible error patterns give different syndromes, making them correctable. Furthermore, it is straightforward to verify that no single fault in the decoding operation leads to any of these patterns. Thus, no single fault causes a multi-qubit correction to be incorrectly applied to the data.

References

- [1] Panos Aliferis, Daniel Gottesman, and John Preskill. Quantum accuracy threshold for concatenated distance-3 codes. *Quantum Information and Computation* **6**(2):97–165, 2006.
- [2] Nadhem J. AlFardan and Kenneth G. Paterson. Lucky 13: Breaking the TLS and DTLS record protocols. In *IEEE Symposium on Security and Privacy*, 2013.
- [3] Nir Bitansky, Ran Canetti, Shafi Goldwasser, Shai Halevi, Yael T. Kalai, and Guy N. Rothblum. Program obfuscation with leaky hardware. In *Advances in Cryptology—ASIACRYPT 2011*, pages 722–739. Springer, Berlin, 2011.
- [4] Boaz Barak, Oded Goldreich, Russell Impagliazzo, Steven Rudich, Amit Sahai, Salil Vadhan, and Ke Yang. On the (im)possibility of obfuscating programs, In *Advances in Cryptology—CRYPTO 2001*, pages 1–18. Springer, Berlin, 2001.
- [5] Hector Bombin and Miguel A. Martin-Delgado. Topological quantum distillation. *Physical Review Letters* **97**(18):180501, 2006.
- [6] Claude Crépeau, Daniel Gottesman, and Adam Smith. Secure multi-party quantum computation. In *Proceedings of the Thirty-Fourth Annual ACM Symposium on Theory of Computing*, ACM, 2002, pages 643–652, 2002.

- [7] David P. DiVincenzo and Panos Aliferis. Effective fault-tolerant quantum computation with slow measurements. *Physical Review Letters* **98**(2):020501, 2007.
- [8] Stefan Dziembowski and Sebastian Faust. Leakage-resilient circuits without computational assumptions. In *Theory of Cryptography*, pages 230–247. Springer, Berlin, 2012.
- [9] Sebastian Faust, Tal Rabin, Leonid Reyzin, Eran Tromer, and Vinod Vaikuntanathan. Protecting circuits from leakage: the computationally-bounded and noisy cases. In *Advances in Cryptology—EUROCRYPT 2010*, pages 135–156. Springer, Berlin, 2010.
- [10] Daniel Gottesman. An introduction to quantum error correction and fault-tolerant quantum computation. In *Quantum Information Science and Its Contributions to Mathematics, Proceedings of Symposia in Applied Mathematics*, volume 68, page 13. 2009.
- [11] Shafi Goldwasser and Guy N. Rothblum. Securing computation against continuous leakage. In *Advances in Cryptology—CRYPTO 2010*, pages 59–79. Springer, Berlin, 2010.
- [12] Shafi Goldwasser and Guy N. Rothblum. How to compute in the presence of leakage. In *2012 IEEE 53rd Annual Symposium on Foundations of Computer Science (FOCS)*, pages 31–40. IEEE, 2012.
- [13] Daniel Genkin, Adi Shamir, and Eran Tromer. RSA key extraction via low-bandwidth acoustic cryptanalysis. In *Advances in Cryptology—CRYPTO 2014, Lecture Notes in Computer Science*, pages 444–461. Springer, Berlin, 2014.
- [14] Yuval Ishai, Manoj Prabhakaran, Amit Sahai, and David Wagner. Private circuits II: Keeping secrets in tamperable circuits. In *Advances in Cryptology-EUROCRYPT 2006*, pages 308–327. Springer, Berlin, 2006.
- [15] Yuval Ishai, Amit Sahai, and David Wagner. Private circuits: Securing hardware against probing attacks. In *Advances in Cryptology-CRYPTO 2003* pages 463–481. Springer, Berlin, 2003.
- [16] D. Kretschmann, D. Schlingemann, and R.F. Werner. The information-disturbance tradeoff and the continuity of Stinespring’s representation. In *IEEE Transactions on Information Theory* **54**(4):1708–1717, 2008.
- [17] Andrew J. Landahl, Jonas T. Anderson, and Patrick R. Rice. Fault-tolerant quantum computing with color codes. arXiv preprint [arXiv:1108.5738](https://arxiv.org/abs/1108.5738) (2011).
- [18] Ueli Maurer. Constructive cryptography—A new paradigm for security definitions and proofs. In *Theory of Security and Applications*, pages 33–56. Springer, Berlin, 2012.
- [19] Silvio Micali and Leonid Reyzin. Physically observable cryptography. In *Theory of Cryptography*, pages 278–296. Springer, Berlin, 2004.
- [20] Ueli Maurer and Renato Renner. Abstract cryptography. In *Innovations in Computer Science*, pages 1–21. 2011.
- [21] National Security Agency, *TEMPEST: A signal problem*, http://www.nsa.gov/public_info/_files/cryptologic_spectrum/tempest.pdf, 2007, Online; Accessed June 5th, 2013.
- [22] John Preskill. Reliable quantum computers. In *Proceedings of the Royal Society A* **454**(1969):385–410, 1998.
- [23] Joseph M. Renes, Volkher B. Scholz, and Stefan Huber. Uncertainty relations: An operational approach to the error-disturbance tradeoff. *Quantum* **1**:20, 2017.
- [24] P.W. Shor. Fault-tolerant quantum computation. In *Proceedings of the 37th Annual Symposium on Foundations of Computer Science*, pages 56–65. 1996.
- [25] Peter W. Shor and John Preskill. Simple proof of security of the BB84 quantum key distribution protocol. *Physical Review Letters* **85**(2):441, 2000.
- [26] Francois-Xavier Standaert, Olivier Pereira, Yu Yu, Jean-Jacques Quisquater, Moti Yung, and Elisabeth Oswald. Leakage resilient cryptography in practice. In *Towards Hardware-Intrinsic Security*, pages 99–134. Springer, Berlin, 2010.
- [27] François-Xavier Standaert, Olivier Pereira, and Yu Yu. Leakage-resilient symmetric cryptography under empirically verifiable assumptions. In *Advances in Cryptology—CRYPTO 2013*, pages 335–352. Springer, Berlin, 2013.
- [28] Andrew Steane. Multiple-particle interference and quantum error correction. In *Proceedings of the Royal Society of London. Series A: Mathematical, Physical and Engineering Sciences* **452**(1954):2551–2577, 1996.
- [29] Andrew M. Steane. Active stabilization, quantum computation, and quantum state synthesis. *Physical Review Letters* **78**(11):2252, 1997.

- [30] Tommaso Toffoli. Reversible computing. In *Seven Colloquium on Automata, Languages and Programming (Noordwijkerhout, the Netherlands), Lecture Notes on Computer Science*, volume 85, pages 632–644. Springer, Berlin, 1980.

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.