Achieving the physical limits of the bounded-storage model

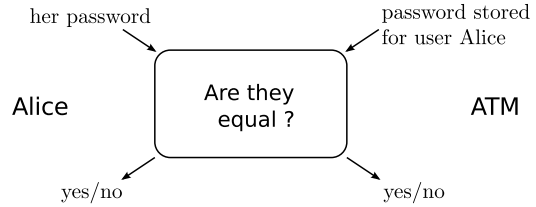
Prabha Mandayam
Institute for Quantum Information, California Institute of Technology, Pasadena CA 91125, USA

Stephanie Wehner
Centre for Quantum Technologies, National University of Singapore, 2 Science Drive 3, 117543 Singapore

(Dated: September 9, 2010)

Secure two-party cryptography is possible if the adversary's quantum storage device suffers imperfections. For example, security can be achieved if the adversary can store strictly less then half of the qubits transmitted during the protocol. This special case is known as the bounded-storage model, and it has long been an open question whether security can still be achieved if the adversary's storage were any larger. Here, we answer this question positively and demonstrate a two-party protocol which is secure as long as the adversary cannot store even a small fraction of the transmitted pulses. We also show that security can be extended to a larger class of noisy quantum memories.

Two-party cryptography enables Alice and Bob to solve problems in cooperation even if they do not trust each other. Important examples of such tasks include auctions and secure identification. In the latter, Alice wants to identify herself to Bob (possibly a fraudulent ATM machine) without revealing her password.



More generally, Alice and Bob wish to solve problems where Alice holds an input x (e.g. her password) and Bob holds an input y (e.g. the password an honest Alice should possess), and they want to obtain the value of some function $f(x, y)$ (e.g. 'yes' if $x = y$, and 'no' otherwise). Security means that Alice should not learn anything about y and Bob should not learn anything about x, apart from what can be inferred from $f(x, y)$ [1].

Contrary to quantum key distribution where honest Alice and Bob can work together to detect the presence of an outside eavesdropper [2, 3], two-party cryptography is made difficult by the fact that Alice and Bob do not trust each other and have to fend for themselves. Indeed, two-party cryptography is impossible without making assumptions about the adversary, even when we allow quantum communication [4]. The security of most cryptographic systems in use today is based on the premise that certain computational problems are hard to solve for the adversary. Concretely, the security relies on the assumption that the adversary's computational resources are limited, and the underlying problem is hard in some precise complexity-theoretic sense. While the former assumption may be justified in practice, the latter statement is usually an unproven mathematical conjecture.

It is thus a natural question whether other, more physical assumptions regarding the two parties' resources allow us to obtain security without relying on any additional hardness results. This is indeed known to be possible if we assume that the adversary's classical [5–7] or quantum storage is limited [8–10] or more generally if his memory is simply imperfect [11–13].

Concretely, the assumption of the noisy-storage model entails that during waiting times Δt in a protocol, the adversary has to measure/discard all his quantum information except what he can encode (arbitrarily) into his quantum memory. Any quantum storage can be modeled as a completely positive trace preserving map $\mathcal{F} : \mathcal{B}(\mathcal{H}_{\text{in}}) \to \mathcal{B}(\mathcal{H}_{\text{out}})$, that maps states $\rho \in \mathcal{H}_{\text{in}}$ to some noisy states $\mathcal{F}(\rho) \in \mathcal{H}_{\text{out}}$. In this paper, we focus on the case where the input space is an n-fold tensor product $\mathcal{H}_{\text{in}} \cong (\mathbb{C}^d)^{\otimes \nu \cdot n}$, the protocols involve n-qu*d*its of communication and the noise is of the form $\mathcal{F} \equiv \mathcal{N}^{\otimes \nu \cdot n}$ with $\mathcal{N} : \mathcal{B}(\mathbb{C}^d) \to \mathcal{B}(\mathbb{C}^d)$. The constant $\nu > 0$ is referred to as the *storage rate* as it captures the fraction of the transmitted qu*d*its that could could potentially be stored by the adversary.

Clearly, the storage rate ν plays a crucial role in deciding whether security can be obtained from a particular storage device. For example, in the case of bounded storage where we have no noise ($\mathcal{N} = \mathbb{I}$), we can never hope to obtain security if the adversary can store all quantum information made available to him during the protocol, that is, if $\nu = 1$ and the input space is $\mathcal{H}_{\text{in}} = (\mathbb{C}^d)^{\otimes n}$. Apart from this trivial condition, however, no bounds were known that restrict our ability to obtain security. In [9] it was shown that security can be achieved in a protocol based on qubits ($d = 2$) as long as $\nu < 1/4$. This was improved to $\nu < 1/2$ in [13]. More generally, it was shown that security in the noisy-storage model can be obtained [13] if

$$C_{\mathcal{N}} \cdot \nu < \frac{1}{2} \,, \tag{1}$$

where $C_{\mathcal{N}}$ is the classical capacity of the quantum channel $\mathcal{N}$.

Results Here, we show that for the case of bounded storage, security can indeed be obtained if the cheating party can store all but a constant fraction of the transmitted pulses. That is, the trivial condition $\nu < 1$ stated above is in fact optimal. The honest players thereby need no quantum storage at all in order to execute the protocol. This not only settles the question, but also highlights the sharp contrast to the case of classical bounded storage, where it was shown that security can only be obtained if the adversary's classical storage is at most quadratic in the storage required by the honest players [14]. Unlike the protocols in [8, 9, 11, 13] which use BB84 encoded qubits, we make use of states encoded in higher-dimensional mutually unbiased bases [34]. Of course, we also scale the storage size accordingly to $\mathcal{H}_{\text{in}} = (\mathbb{C}^d)^{\otimes \nu \cdot n}$ when sending d dimensional states. More specifically, we show that security in the setting of bounded storage is possible as long as

$$\nu < \frac{\log(d+1) - 1}{\log d} \to 1 \ , \tag{2}$$

where the r.h.s. approaches 1 for large d. We stress that for large values of d, the resulting protocols will be much harder to implement experimentally, and even though the errors decrease exponentially with n they converge very slowly for large d. Note, however, that here we are merely interested in exploring the fundamental physical limitations of this model.

For the general setting of noisy quantum storage we further show that security is possible for devices $\mathcal{F} = \mathcal{N}^{\otimes \nu \cdot n}$, where the channel $\mathcal{N} : \mathcal{B}(\mathbb{C}^d) \to \mathcal{B}(\mathbb{C}^d)$ satisfies the strong converse property [16], whenever

$$C_{\mathcal{N}} \cdot \nu < \log(d+1) - 1 \ , \tag{3}$$

thus extending the range of storage devices for which we can prove security [13].

Techniques We first give an overview of the steps involved in obtaining our result. The constant 1/2 in the bound (1) is a result of using BB84-states [2] in the protocol, and stems from an uncertainty relation for measurements in these two bases [17]. It is thus natural to consider a protocol that uses more than two mutually unbiased bases (MUBs) for which uncertainty relations are known to exist [18]. Our first step is to obtained a modified protocol for the simple two-party primitive weak string erasure which was originally introduced in [13], using the full set of $d+1$ MUBs that are known to exist in prime power dimensions [19, 20]. Next, we show that there is still a secure protocol for the cryptographic primitive of oblivious transfer using this variant of weak string erasure. Since it is known that any two-party cryptographic problem can be solved using oblivious transfer [21], this concludes our claim.

WEAK STRING ERASURE

We first show how to obtain a variant of the very simple cryptographic primitive weak string erasure (WSE) introduced in [13], which we will call *non-uniform* weak string erasure; a formal definition can be found in the appendix. Intuitively, this primitive provides Alice with a string $X^n = (X_1, ..., X_n) \in \{0, 1, \ldots, d-1\}^n$, where each entry X_i takes on one of d possible values. Bob obtains a set of index locations $\mathcal{I} = \{i_1, ..., i_{|\mathcal{I}|} \mid i_j \in [n]\}$, where any index $i \in \{1, \ldots, n\} =: [n]$ is chosen to be in $\mathcal{I}$ with some probability p. In addition, Bob receives the entries of the string X^n corresponding to the indices $\mathcal{I}$, which we denote by the substring $X_{\mathcal{I}} = (X_{i_1}, X_{i_2}, \ldots, X_{i_{|\mathcal{I}|}})$. Security here means that even if Alice is dishonest, she cannot learn which entries are known to Bob, i.e., she cannot learn anything about the index set $\mathcal{I}$. Conversely, if Bob is dishonest, then his information about the entire string X^n should still be limited in the sense that the probability that he can guess all of X^n given his information B' is small. That is,

$$P_{\text{guess}}(X|B') \leq 2^{-\lambda n} \ . \tag{4}$$

for some $\lambda > 0$. This is equivalent [22] to demanding that his min-entropy [23] obeys

$$\mathrm{H}_\infty(X^n|B')_\rho = -\log P_{guess}(X^n|B') \geq \lambda n \ . \tag{5}$$

In practice, we allow this condition to fail with error parameter ε, which corresponds to demanding that the *smooth* min-entropy defined as

$$\mathrm{H}^\varepsilon_\infty(X|B')_\rho := \sup_{\substack{\bar{\rho}_{XB'} \geq 0 : \frac{1}{2}\|\bar{\rho}_{XB'} - \rho_{XB'}\|_1 \leq \mathrm{tr}(\rho_{XB'})\cdot\varepsilon \\ \mathrm{tr}(\bar{\rho}_{XB'}) \leq \mathrm{tr}(\rho_{XB'})}} \mathrm{H}_\infty(X|B')_{\bar{\rho}} \ ,$$

satisfies $\mathrm{H}^\varepsilon_\infty(X|B') \geq \lambda n$.

Next we outline a simple protocol that achieves the functionality described above. It is a straightforward generalization of the original protocol in [13] to multiple encodings, the main difference being that the indices in $\mathcal{I} \subseteq [n]$ are no longer chosen uniformly at random. Instead, the probability p that honest Bob learns the value of X_i for $i \in [n]$ is equal to the probability that he chooses the same basis as Alice, that is, $p = 1/(d+1)$.Let $2^{[n]}$ denote the set of all subsets of $[n]$.

Protocol 1: Non-uniform WSE
Outputs: $x^n \in \{0, 1, \ldots, d-1\}^n$ to Alice, $(\mathcal{I}, z^{|\mathcal{I}|}) \in 2^{[n]} \times \{0, 1, \ldots, d-1\}^{|\mathcal{I}|}$ to Bob.

1: Alice: Picks an n-$d$it string uniformly at random, $x^n \in \{0, 1, ..., d-1\}^n$. She encodes each $d$it into one of the $d+1$ MUBs, that is, she chooses a basis string $\theta^n \in \{0, ..., d\}^n$ uniformly at random, so that the $d$it x_j is encoded in basis $\mathcal{B}_{\theta_j}$, and sends it to Bob.

2: Bob: Chooses a basis string $\tilde{\theta}^n \in \{0,1,...,d\}^n$ uniformly at random. When receiving the i-th qu$d$it, he measures it in the basis $\mathcal{B}_{\tilde{\theta}_i}$, to obtain outcome $\tilde{x}_i$.

Both parties wait time Δt.

3: Alice: Sends the basis information θ^n to Bob, and outputs x^n.

4: Bob: Computes $\mathcal{I} := \{i \in [n] | \theta_i = \tilde{\theta}_i\}$, and outputs $(\mathcal{I}, \tilde{x}_{\mathcal{I}})$.

We now formally state our claim that this protocol does indeed implement non-uniform WSE, with the parameters ε, λ and d.

Theorem 1. *Let Bob's storage be given by $\mathcal{F} = \mathcal{N}^{\otimes \nu n}$ with a storage rate $\nu > 0$, where $\mathcal{N}$ satisfies the strong converse property [16] and the capacity $C_{\mathcal{N}}$ of the channel $\mathcal{N}$ bounded by*

$$C_{\mathcal{N}} \cdot \nu < \log(d+1) - 1 \ . \tag{6}$$

Let $\delta \in]0, \frac{1}{2} - C_{\mathcal{N}} \cdot \nu[$. Then, Protocol 1 securely implements weak string erasure for sufficiently large n with

$$\lambda(\delta, d) = \nu \cdot \gamma^{\mathcal{N}} \left(\frac{\log(d+1) - 1 - \delta}{\nu} \right) \tag{7}$$

and error $\varepsilon(\delta, d) = 2\exp(-f(\delta, d)n)$ with

$$f(\delta, d) \propto -n\delta^2 / \left(\log((d+1)\cdot d) + \log 4/\delta\right)^2 > 0, \tag{8}$$

where $\gamma^{\mathcal{N}}(\cdot)$ is the strong converse parameter of $\mathcal{N}$ [16].

Note that for bounded storage, where $\mathcal{N}$ is simply the identity channel over Bob's d-dimensional input Hilbert space, $C_{\mathcal{N}} = \log d$ in (6), which is our central result (2).

It is easy to see that the protocol is correct if both parties are honest: if Alice is honest, her string $X^n = x^n$ is chosen uniformly at random from $\{0, 1, \ldots, d-1\}^n$ as desired, and if Bob is honest, he clearly obtains $\tilde{x}_i = x_i$ whenever $i \in \mathcal{I}$ for a random subset $\mathcal{I} \subseteq [n]$. In the remainder of this section, we prove security if one of the parties is dishonest.

Dishonest Bob The security proof proceeds in three steps: First, we consider Bob's information about the string X^n given only his classical information K, and the basis information Θ^n he receives. This can be quantified using entropic uncertainty relations in terms of the Shannon entropy for $d+1$ MUBs in $\mathbb{C}_d$ [18]. Using [24, Theorem 4.22] these uncertainty relations imply a bound on Bob's information in terms of the smooth min-entropy

$$H_\infty^{\varepsilon/2}(X^n | K\Theta^n)_\rho \geq \left(\log(d+1) - 1 - \frac{\delta}{2} \right) n \ , \tag{9}$$

for any $0 < \delta < \frac{1}{2}$ with $\varepsilon = 2\exp\left(-f(\delta, d)n\right)$. That is, the error decreases exponentially with n, as desired.

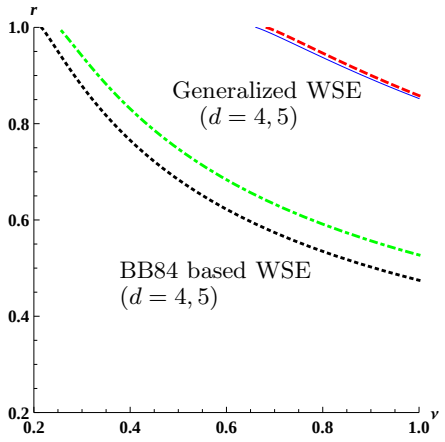


FIG. 1: Security regions (r, ν) for weak string erasure (WSE) with depolarizing noise $\mathcal{N}(\rho) = r\rho + (1-r)\mathbb{I}/d$, in dimensions $d = 4, 5$. Previously [13], security was shown in the regions below the dotted black curve for $d = 4$ and the dot-dashed green curve for $d = 5$. Our analysis extends the security region to the solid blue curve ($d = 4$) and the dashed red curve ($d = 5$) respectively.

Next we consider Bob's information, when in addition he is given the output of his storage device $\mathcal{F}(Q)$. We know from [13] that the uncertainty relation (9) determines the rate at which Bob needs to send information through his storage device. Using [13, Lemma 2.2] together with (9) we obtain

$$H_\infty^\varepsilon(X^n | \Theta^n K \mathcal{F}(Q))_\rho \geq -\log P_{succ}^{\mathcal{F}} \left[n \left(\log(d+1) - 1 - \frac{\delta}{2} \right) \right] \ , \tag{10}$$

where, $P_{\text{succ}}^{\mathcal{F}}(nR)$ is the average probability of sending a randomly chosen string $x \in \{0,1\}^{nR}$ through the storage $\mathcal{F}$ [35]. For noise of the form $\mathcal{F} = \mathcal{N}^{\otimes \nu n}$, the r.h.s. of (10) is the success probability of sending νn bits at a rate $R = (\log(d+1) - 1 - \delta/2)/\nu$. The final step is to note that for channels satisfying the strong converse property [16], this success probability drops off exponentially with n according to the parameter $\gamma^{\mathcal{N}}(\cdot)$, whenever $R > C_{\mathcal{N}}$. This gives the bound

$$C_{\mathcal{N}} \cdot \nu < \log(d+1) - 1 - \frac{\delta}{2} \ . \tag{11}$$

Theorem 1 then follows by noting that exponential security (in n) is possible for any constant $\delta > 0$.

Dishonest Alice When Alice is dishonest, it is intuitively obvious that she is unable to gain any information about the index set $\mathcal{I}$, since she never receives any information from Bob during our protocol. However, a more careful security analysis is required if we want to use weak string erasure to build more complicated primitives like oblivious transfer. This argument is essentially analogous to [13], as outlined in the appendix.

OBLIVIOUS TRANSFER

Ultimately, we would like to use WSE to solve arbitrary two-party cryptographic problems. To this end, it suffices to implement the primitive oblivious transfer [36], which can solve any two-party problem [21]. Informally, this primitive outputs two strings $S_0^\ell, S_1^\ell \in \{0,1\}^\ell$ to Alice, and a choice bit $C \in \{0,1\}$ and S_C^ℓ to Bob. Security means that if Alice is dishonest, she should not learn anything about C. If Bob is dishonest, we demand that there exists some random variable C such that Bob is entirely ignorant about S_{1-C}^ℓ. That is, he may learn at most one of the two strings which are generated.

Here, we state a simplified version of the actual protocol which executes fully randomized oblivious transfer from WSE. It contains all the essential ingredients to understand the main steps of our security proof. A formal definition, as well as the full protocol can be found in the appendix. The only difference from the protocol presented in [13] is the fact that $\mathcal{I}$ is no longer uniform, and honest Bob only learns about pn entries x_j, whereas in the case of uniform WSE [13] he could learn roughly $n/2$. We hence introduce a new parameter $\eta = 2(d+1)$ in the protocol, such that with high probability Bob learns at least n/η of the indices.

Protocol 2: Oblivious Transfer
Outputs: $(s_0^\ell, s_1^\ell) \in \{0,1\}^\ell \times \{0,1\}^\ell$ to Alice, and $(c, y^\ell) \in \{0,1\} \times \{0,1\}^\ell$ to Bob

1: Alice and Bob: Execute WSE. Alice gets a string $x^n \in \{0,1,\ldots,d-1\}^n$, Bob a set $\mathcal{I} \subset [n]$ and a string $s = x_{\mathcal{I}}$. If $|\mathcal{I}| < n/\eta$, Bob chooses uniformly at random a set $\mathcal{I}_{\text{tr}}$ of size $|\mathcal{I}_{\text{tr}}| = n/\eta$. Otherwise, he randomly truncates $\mathcal{I}$ to $|\mathcal{I}_{\text{tr}}|$ of size n/η, and deletes the corresponding values in s.

2: Alice and Bob: Execute interactive hashing with Bob's input w equal to a description of $\mathcal{I}_{\text{tr}} = \mathsf{Enc}(w)$. Interpret the outputs w_0 and w_1 as descriptions of subsets $\mathcal{I}_0$ and $\mathcal{I}_1$ of $[n]$.

3: Alice: Chooses $r_0, r_1 \in_R \mathcal{R}$ and sends them to Bob.

4: Alice: Outputs $(s_0^\ell, s_1^\ell) := (\mathsf{Ext}(x_{\mathcal{I}_0}, r_0), \mathsf{Ext}(x_{\mathcal{I}_1}, r_1))$ using the 2-universal hash function known from quantum key distribution [23] $\mathsf{Ext} : \{0,\ldots,d-1\}^{n/\eta} \times \mathcal{R} \to \{0,1\}^\ell$.

5: Bob: Computes $c \in \{0,1\}$ with $\mathcal{I} = \mathcal{I}_c$, and $x_{\mathcal{I}}$ from s. He outputs $(c, y^\ell) := (c, \mathsf{Ext}(s, r_c))$.

We provide only an overview of our proof since it closely follows the steps in [13]; details can be found in the appendix. To show that the protocol is correct we first use Hoeffding's inequality [25] to show that except with exponentially small probability $\exp(-2n/\eta)$, Bob learns a sufficient number of indices to retrieve the desired string S_C.

Dishonest Alice The properties of weak string erasure ensure that a dishonest Alice does not know which *d*its $x_{\mathcal{I}}$ of x^n are known to Bob, that is, she is ignorant about the index set $\mathcal{I}$. Finally, the properties of interactive hashing [26] ensure that she cannot gain any information which of the two subsets $\mathcal{I}_0$ and $\mathcal{I}_1$ of bits are known to Bob. Hence, she cannot learn C as desired.

A significant change from the original security proof in [13] is in showing that the truncated set $\mathcal{I}_{\text{tr}}$ is actually uniform over subsets of size n/η, since $\mathcal{I}$ is no longer distributed uniformly over $2^{[n]}$. Formally, the probability of a given truncated set $\mathcal{I}_{\text{tr}}$ can be written in terms of the probability $p(\bar{A})$ that $|\mathcal{I}| \geq n/\eta$, as follows:

$$p(\mathcal{I}_{\text{tr}}|\bar{A}) = \sum_{\substack{\mathcal{I} \\ |\mathcal{I}| \geq n/\eta}} \frac{p(\mathcal{I}|\bar{A})}{\binom{|\mathcal{I}|}{n/\eta}} = \frac{1}{p(\bar{A})} \sum_{\mathcal{I}} \frac{p(\mathcal{I})}{\binom{|\mathcal{I}|}{n/\eta}} , \qquad (12)$$

independent of the choice of truncation as desired. Here, $1/\binom{|\mathcal{I}|}{n/\eta}$ is the probability of choosing the particular subset $\mathcal{I}_{\text{tr}}$ from $\mathcal{I}$ and $p(\mathcal{I}|\bar{A})$ is the conditional probability of a set $\mathcal{I}$, given that $|\mathcal{I}| \geq n/\eta$. The final equality is simply an application of Bayes' rule, $p(\bar{A})p(\mathcal{I}|\bar{A}) = p(\bar{A}|\mathcal{I})p(\mathcal{I})$.

Dishonest Bob Again, it follows from weak string erasure that a dishonest Bob gains only a limited amount of information about the string X^n. The properties of interactive hashing ensure that Bob has very little control over the subset $\mathcal{I}_{1-c}$ chosen by the interactive hashing. Therefore, by the results on min-entropy sampling [27], Bob has only limited information about the *d*its in this subset. Privacy amplification [23, 28] can then be used to turn this into almost complete ignorance.

CONCLUSION

We have shown that any two-party cryptographic primitive can be implemented securely in the setting of bounded quantum storage, even if the adversary can store all but a fraction of the transmitted pulses. This is optimal, in the sense that we could never hope to achieve security if the cheating party could store *all* quantum communication made available to him. This demonstrates that there is no physical principle that prevents us from achieving security even with a very high storage rate $\nu < 1$. We have also shown in the noisy-storage setting that security is possible for a much larger range of noisy quantum memories.

To achieve our result we use higher dimensional states which are very difficult to create in practice. It is therefore an interesting open question, whether the same result could be obtained using merely BB84 encoded

qubits. Note, however, that our approach merely relies on the existence of entropic uncertainty relations for multiple encodings, and our protocols and proofs are completely analogous if we were to use any other encodings for which strong uncertainty relations are known to exist. For example, it is conceivable that uncertainty relations for multiple encodings can be based on top of BB84 encoded qubits [29], which would immediately lead to a protocol that is easy to implement experimentally.

We thank Robert König and Jürg Wullschleger for many interesting and useful discussions, as well as comments on an earlier draft. We also thank Christian Schaffner for comments. PM and SW were supported by NSF grants PHY-04056720 and PHY-0803371. SW was supported by the National Research Foundation and the Ministry of Education, Singapore. Part of this work was done while SW was at the Institute for Quantum Information, Caltech.

[1] A. C. Yao, in *Proceedings of the 23rd Annual IEEE FOCS* (1982), pp. 160–164.
[2] C. H. Bennett and G. Brassard, in *Proceedings of the IEEE International Conference on Computers, Systems and Signal Processing* (1984), pp. 175–179.
[3] A. Ekert, Physical Review Letters **67**, 661 (1991).
[4] H.-K. Lo, Physical Review A **56**, 1154 (1997).
[5] U. Maurer, Journal of Cryptology **5**, 53 (1992).
[6] U. Maurer, in *Advances in Cryptology - EUROCRYPT* (1990), vol. 473 of *Lecture Notes in Computer Science*, pp. 361–373.
[7] C. Cachin and U. M. Maurer, in *Proceedings of CRYPTO 1997* (1997), pp. 292–306.
[8] I. B. Damgård, S. Fehr, L. Salvail, and C. Schaffner, in *Proceedings of 46th IEEE FOCS* (2005), pp. 449–458.
[9] I. B. Damgård, S. Fehr, R. Renner, L. Salvail, and C. Schaffner, in *Advances in Cryptology—CRYPTO '07* (Springer-Verlag, 2007), vol. 4622 of *Lecture Notes in Computer Science*, pp. 360–378, quant-ph/0612014.
[10] I. B. Damgård, S. Fehr, L. Salvail, and C. Schaffner, in *Advances in Cryptology—CRYPTO '07* (Springer-Verlag, 2007), vol. 4622 of *Lecture Notes in Computer Science*, pp. 342–359.
[11] S. Wehner, C. Schaffner, and B. M. Terhal, Physical Review Letters **100**, 220502 (2008).
[12] C. Schaffner, B. Terhal, and S. Wehner (2008), arXiv:0807.1333.
[13] R. König, S. Wehner, and J. Wullschleger (2009), arXiv:0906.1030.
[14] S. Dziembowski. U. Maurer, in *Proceedings of EUROCRYPT*(2004), vol. 3027 of *Lecture Notes in Computer Science*, pp. 126-137.
[15] Y. Z. Ding, D. Harnik, A. Rosen, and R. Shaltiel, in *Proceedings of TCC* (2004), vol. 2951 of *Lecture Notes in Computer Science*, pp. 446–472.
[16] R. König and S. Wehner (2009), arXiv:0903.2838.
[17] H. Maassen and J. Uffink, Physical Review Letters **60** (1988).
[18] J. Sanchez, Physics Letters A **173**, 233 (1993).
[19] S. Bandyopadhyay, P. Boykin, V. Roychowdhury, and F. Vatan, Algorithmica **34**, 512 (2002).
[20] W. Wootters and B. Fields, Ann. Phys. **191** (1989).
[21] J. Kilian, in *Proceedings of 20th ACM STOC* (1988), pp. 20–31.
[22] R. König, R. Renner, and C. Schaffner (2008), arXiv:0807.1338.
[23] R. Renner, Ph.D. thesis, ETH Zurich (2005), quant-ph/0512258.
[24] C. Schaffner, Ph.D. thesis, University of Aarhus (2007), http://arxiv.org/abs/0709.0289.
[25] W. Hoeffding, Journal of the American Statistical Association **58**, 13 (1963).
[26] G. Savvides, Ph.D. thesis, McGill University, Montreal (2007).
[27] R. König and R. Renner (2007), arXiv:0712.4291.
[28] R. Renner and R. König, in *Proceedings of TCC 2005* (Springer, 2005), vol. 3378 of *Lecture Notes in Computer Science*, pp. 407–425.
[29] P. Hayden (2010), personal communication.
[30] S. Fehr and C. Schaffner (2008), arXiv:0804.1059.
[31] C. H. Bennett, G. Brassard, C. Crépeau, and H. Skubiszewska, in *Advances in Cryptology — CRYPTO '91* (Springer, 1992), vol. 576 of *Lecture Notes in Computer Science*, pp. 351–366.
[32] D. Beaver, in *Advances in Cryptology — EUROCRYPT '95* (Springer-Verlag, 1995), vol. 963 of *Lecture Notes in Computer Science*, pp. 97–109.
[33] S. Wehner and A. Winter, New Journal of Physics p. 025009 (2010), special Issue on Quantum Information and Many-Body Systems.
[34] Two orthonormal bases $\mathcal{B}_1 = \{|x^{(1)}\rangle \mid x \in \{0,\dots,d-1\}\}$ and $\mathcal{B}_2 = \{|x^{(2)}\rangle \mid x \in \{0,\dots,d-1\}\}$ are called mutually unbiased if for all $x, y \in \{0,\dots,d-1\}$ we have $|\langle x^{(1)}|y^{(2)}\rangle| = 1/\sqrt{d}$.
[35] $P^{\mathcal{F}}_{\text{succ}}(nR)$ is given by $P^{\mathcal{F}}_{\text{succ}}(nR) = \max_{\{M_x\}_x,\{\rho_x\}_x} \frac{1}{2^{nR}} \sum_{x\in\{0,1\}^n} \text{tr}\,(M_x \mathcal{F}(\rho_x))$, where the maximization is taken over encodings $\{\rho_x\}_x$ and decoding POVMs $\{M_x\}_x$.
[36] Here we execute fully randomized oblivious transfer [9, 30] which can easily be converted into standard 1-2 oblivious transfer [31, 32].

For completeness, we provide the technical details of our protocols and security proofs in the appendix.

Weak String Erasure

To formally state our result, let us now first define non-uniform weak string erasure. This definition closely follows the one of [13], except that the string X^n is now chosen from a larger alphabet and the indices in $\mathcal{I} \subseteq [n]$ are not chosen uniformly at random. Instead, the probability p that honest Bob learns the value of X_i for $i \in [n]$ is equal to the probability that he chooses the same basis as Alice, i.e., $p = 1/(d+1)$. In the definition below, we will need to talk about distributions over subsets $\mathcal{I} \subseteq [n]$. Clearly, the probability that Bob learns a particular subset $\mathcal{I}$ satisfies

$$\Pr(\mathcal{I}) = p^{|\mathcal{I}|}(1-p)^{n-|\mathcal{I}|} \tag{13}$$

Note that we can write the subset $\mathcal{I}$ as a string $(y_1, \ldots, y_n) \in \{0,1\}^n$ where $y_i = 1$ if and only if $i \in \mathcal{I}$, allowing us to identify $|\mathcal{I}\rangle := |y_1\rangle \otimes \ldots \otimes |y_n\rangle$. The probability distribution over subsets $\mathcal{I} \subseteq [n]$ can then be expressed as (see also [13])

$$\Psi(p) := \sum_{\mathcal{I} \subseteq 2^{[n]}} p^{|\mathcal{I}|}(1-p)^{n-|\mathcal{I}|} |\mathcal{I}\rangle\langle\mathcal{I}| \ . \tag{14}$$

Furthermore, we will follow the notation of [13] and use

$$\tau_{\mathcal{S}} := \frac{1}{|\mathcal{S}|} \sum_{s \in \mathcal{S}} |s\rangle\langle s| \ , \tag{15}$$

to denote the uniform distribution over a set $\mathcal{S}$.

Definition 2 (Non-uniform WSE). *An $(n, \lambda, \varepsilon, p, d)$-weak string erasure scheme is a protocol between A and B satisfying the following properties:*

Correctness: *If both parties are honest, then there exists an ideal state $\sigma_{X^n \mathcal{I} X_{\mathcal{I}}}$ is defined such that*

1. *The joint distribution of the n-dit string X^n and subset $\mathcal{I}$ is given by*

$$\sigma_{X^n \mathcal{I}} = \tau_{\{0,1,\ldots,d-1\}^n} \otimes \Psi(p) \ , \tag{16}$$

2. *The joint state ρ_{AB} created by the real protocol is equal to the ideal state: $\rho_{AB} = \sigma_{X^n \mathcal{I} X_{\mathcal{I}}}$ where we identify (A, B) with $(X^n, \mathcal{I} X_{\mathcal{I}})$.*

Security for Alice: *If A is honest, then there exists an ideal state $\sigma_{X^n B'}$ such that*

1. *The amount of information B' gives Bob about X^n is limited:*

$$\frac{1}{n} H_\infty(X^n|B')_\sigma \geq \lambda \tag{17}$$

2. *The joint state $\rho_{AB'}$ created by the real protocol is ε-close to the ideal state, i.e. $\sigma_{X^n B'} \approx_\varepsilon \rho_{AB'}$ where we identify (X^n, B') with (A, B').*

Security for Bob: *If B is honest, then there exists ideal state $\sigma_{A' \hat{X}^n \mathcal{I}}$ where $\hat{X}^n \in \{0, 1, ..., d-1\}^n$ and $\mathcal{I} \subseteq [n]$ such that*

1. *The random variable $\mathcal{I}$ is independent of $A' \hat{X}^n$ and distributed over $2^{[n]}$ according to the probability distribution given by* (13)*:*

$$\sigma_{A' \hat{X}^n \mathcal{I}} = \sigma_{A' \hat{X}^n} \otimes \Psi(p) \ . \tag{18}$$

2. *The joint state $\rho_{A'B}$ created by the real protocol is equal to the ideal state: $\rho_{A'B} = \sigma_{A'(\mathcal{I} \hat{X}_{\mathcal{I}})}$, where we identify (A', B) with $(A', \mathcal{I} \hat{X}_{\mathcal{I}})$.*

We are now ready to state our result for non-uniform weak string erasure more formally. We first state the general result for quantum memories, and then focus on the tensor-product channels of the type $\mathcal{F} = \mathcal{N}^{\otimes \nu \cdot n}$.

Theorem 3. *(i) Let $\delta \in]0, \frac{1}{2}[$ and let Bob's storage be given by $\mathcal{F} : \mathcal{B}(\mathcal{H}_{\text{in}}) \to \mathcal{B}(\mathcal{H}_{\text{out}})$. Then Protocol 1 is an $(n, \lambda(\delta, d), \varepsilon(\delta, d), 1/(d+1), d)$-weak string erasure protocol with min-entropy rate*

$$\lambda(\delta, d) = -\lim_{n \to \infty} \frac{1}{n} P^{\mathcal{F}}_{succ}\left((\log(d+1) - 1 - \delta) \cdot n\right) \ ,$$

and error $\varepsilon(\delta, d) = 2\exp(-f(\delta, d)n)$ with

$$f(\delta, d) := \frac{(\delta/4)^2}{32\left(\log((d+1) \cdot d) + \log \frac{4}{\delta}\right)^2} > 0. \tag{19}$$

(ii) Suppose $\mathcal{F} = \mathcal{N}^{\otimes \nu n}$ for a storage rate $\nu > 0$, $\mathcal{N}$ satisfying the strong-converse property and having capacity $C_{\mathcal{N}}$ bounded by

$$C_{\mathcal{N}} \cdot \nu < \log(d+1) - 1 \ .$$

Let $\delta \in]0, \frac{1}{2} - C_{\mathcal{N}} \cdot \nu[$. Then Protocol 1 is an $(n, \tilde{\lambda}(\delta, d), \varepsilon(\delta, d), 1/(d+1), d)$-weak string erasure protocol for sufficiently large n, where

$$\tilde{\lambda}(\delta, d) = \nu \cdot \gamma^{\mathcal{N}}\left(\frac{\log(d+1) - 1 - \delta}{\nu}\right) \ .$$

Note that when $\mathcal{N} = \mathbb{I}_d$ then $C_{\mathcal{N}} = \log d$, so that the bound in (26) holds for a storage rate of

$$\nu < \frac{\log(d+1) - 1}{\log d} \approx 1, \text{ for large } d.$$

Thus for the case of bounded storage, security can in principle be obtained for any storage rate $\nu < 1$, provided we choose a large enough system size d.

Security for honest Alice

Let us now first consider the case of dishonest Bob. The main difference from [13] in proving security lies in the use of the uncertainty relation for the full set of $d+1$ mutually unbiased bases in prime power dimensions [18]. To see where we will make use of this relation, note that analogous to [13] we can model Bob's attack as a CPTP map $\mathcal{E} : \mathcal{B}((\mathbb{C}^d)^{\otimes n}) \to \mathcal{B}(\mathcal{H}_{\text{in}} \otimes \mathcal{H}_K)$ so that for any input state $\rho \in (\mathbb{C}^d)^{\otimes n}$ provided by Alice before the waiting time, he obtains an output state $\zeta_{Q_{\text{in}} K} = \mathcal{E}(\rho)$, where Q_{in} is the quantum information he puts into his quantum storage and K is any additional classical information he retains. Hence, the joint state of Alice and Bob before his storage noise is applied is of the form

$$\rho_{X^n \Theta^n K Q_{in}} = \frac{1}{d^n (d+1)^n} \sum_{x^n, \theta^n, k} P_{K|X^n = x^n, \Theta^n = \theta^n}(k) \underbrace{|x^n\rangle\langle x^n| \otimes |\theta^n\rangle\langle\theta^n|}_{\text{Alice}} \otimes \underbrace{|k\rangle\langle k| \otimes \zeta_{x^n \theta^n k}}_{\text{Bob}} \tag{20}$$

where $\zeta_{x^n\theta^n k}$ is the state on $\mathcal{H}_{\text{in}}$ depending on Alice's choice of string x^n, bases θ^n and Bob's classical information k. Bob's storage then undergoes noise described by $\mathcal{F} : \mathcal{B}(\mathcal{H}_{\text{in}}) \to \mathcal{B}(\mathcal{H}_{\text{out}})$, and the state evolves to $\rho_{X^n\Theta^n K\mathcal{F}(Q_{\text{in}})}$. After time Δt Bob also receives the basis info $\Theta^n = \theta^n$. Then their joint state is

$$\rho_{X^n\Theta^n K\mathcal{F}(Q_{\text{in}})} = \frac{1}{d^n(d+1)^n} \sum_{x^n,\theta^n,k} P_{K|X^n=x^n,\Theta^n=\theta^n}(k) \underbrace{|x^n\rangle\langle x^n|}_{\text{Alice}} \otimes \underbrace{|\theta^n\rangle\langle\theta^n| \otimes \mathcal{F}(\zeta_{x^n\theta^n k})}_{\text{Bob B'}} \tag{21}$$

where Bob now holds $B' = \Theta^n K\mathcal{F}(Q_{\text{in}})$.

Our goal is to show that for any cheating strategy of dishonest Bob, his min-entropy about the string $X^n = (X_1, ..., X_n)$ is large, using an entropic uncertainty relation. Recall that the set of $(d+1)$ MUBs in $\mathbb{C}_d$ satisfies [18] (see [33] for a simple proof)

$$\frac{1}{d+1}\sum_{i=1}^{d+1} H(\mathcal{B}_i|\rho) \geq \log(d+1) - 1, \forall\, \rho \in \mathbb{C}_d, \tag{22}$$

where

$$H(\mathcal{B}_i|\rho) = -\sum_x \text{Tr}(|b_i^x\rangle\langle b_i^x|\rho) \log \text{Tr}(|b_i^x\rangle\langle b_i^x|\rho) \tag{23}$$

is the Shannon entropy of the probability distribution induced by measuring the state ρ in the basis $\mathcal{B}_i$. This lower bound, along with the uncertainty relation for the smooth min-entropy from [24, Theorem 4.22] gives

$$H_\infty^{\varepsilon/2}(X^n|\Theta^n)_\rho \geq \left(\log(d+1) - 1 - \frac{\delta}{2}\right) n \tag{24}$$

for any $0 < \delta < \frac{1}{2}$ with

$$\varepsilon = 2\exp\left(-\frac{(\delta/4)^2 n}{32\left(\log((d+1)\cdot d) + \log\frac{4}{\delta}\right)^2}\right). \tag{25}$$

Finally, we make use of Lemma 2.2 in [13] that relates the smooth min-entropy to the maximal decoding probability, $P_{succ}^{\mathcal{F}}$, to get,

$$\begin{aligned} &H_\infty^\varepsilon(X^n|\Theta^n K\mathcal{F}(Q_{in}))_\rho \\ &\geq -\log P_{succ}^{\mathcal{F}}\left(n\left(\log(d+1) - 1 - \tfrac{\delta}{2}\right) - \log\tfrac{2}{\varepsilon}\right) \\ &\geq -\log P_{succ}^{\mathcal{F}}\left(n(\log(d+1) - 1) - n\tfrac{\delta}{2}\right) \end{aligned}$$

where the second inequality follows from the monotonicity of $P_{\text{succ}}^{\mathcal{F}}$ and the fact that $\log\frac{2}{\varepsilon} < \frac{\delta}{2}n$ for $0 < \delta < \frac{1}{2}$. By definition of the smooth min-entropy, this implies that there exists an ideal state $\sigma_{X^nB'}$ such that

1. $\sigma_{X^nB'} \approx_\varepsilon \rho_{X^nB'}$,
2. $\frac{1}{n}H_\infty(X^n|B')_\sigma \geq -\frac{1}{n}\log P_{succ}^{\mathcal{F}}\left(n\log(d+1) - n - \frac{\delta}{2}n\right)$,

which proves part (i) of Theorem 3.

In the special case that $\mathcal{F}$ is the tensor product channel $\mathcal{F} = \mathcal{N}^{\otimes\nu n}$, where $\mathcal{N}$ satisfies the strong converse property and $C_{\mathcal{N}} \cdot \nu < \log(d+1) - 1$, following the same steps as in [13] we obtain that there exists an ideal state $\sigma_{X^nB'}$ that is ε-close to $\rho_{X^nB'}$ and has a min-entropy

$$\frac{1}{n}H_\infty(X^n|B')_\sigma \geq \nu.\gamma^{\mathcal{N}}\left(\frac{\log(d+1) - 1 - \frac{\delta}{2}}{\nu}\right) > 0\ , \tag{26}$$

where $\gamma^{\mathcal{N}}(\cdot)$ is the strong converse parameter of the channel $\mathcal{N}$ [16]. This proves part (ii) of Theorem 3.

Security for honest Bob

The proof of security when Alice is dishonest is essentially analogous to [13] (see Section 3.4 and Figures 7 and 8), where we introduce an imaginary "simulator" with perfect quantum memory to define the desired ideal state. We hence merely state how to adapt the proof of [13]: here we naturally obtain $\Psi(p)$ in place of the uniform distribution $\tau_{2^{[n]}}$ in our simulation. Similarly, the subset $\mathcal{I}$ is not chosen uniformly at random, but with probability

$$\Pr(\mathcal{I}) := \left(\frac{1}{d+1}\right)^{|\mathcal{I}|}\left(\frac{d}{d+1}\right)^{n-|\mathcal{I}|}\ . \tag{27}$$

1-2 Oblivious Transfer from Weak String Erasure

We are now ready to show how oblivious transfer can be obtained even from the non-uniform variant of weak string erasure. To formally state our result we thereby include the definition of oblivious transfer from [13].

Definition 4. *An (ℓ, ε)-fully randomized oblivious transfer (FROT) scheme is a protocol between Alice and Bob satisfying the following:*

Correctness: *If both parties are honest, then the ideal state $\sigma_{S_0^\ell S_1^\ell C S_C^\ell}$ is defined such that*

1. *The distribution over S_0^ℓ, S_1^ℓ and C is uniform:*

   $$\sigma_{S_0^\ell S_1^\ell C} = \tau_{\{0,1\}^\ell} \otimes \tau_{\{0,1\}^\ell} \otimes \tau_{\{0,1\}}\ .$$

2. *The real state $\rho_{S_0^\ell S_1^\ell C Y^\ell}$ created during the protocol is ε-close to the ideal state:*

   $$\rho_{S_0^\ell S_1^\ell C Y^\ell} \approx_\varepsilon \sigma_{S_0^\ell S_1^\ell C S_C^\ell}\ , \tag{28}$$

   where we identify $A = (S_0^\ell, S_1^\ell)$ and $B = (C, Y^\ell)$.

Security for Alice: *If Alice is honest, then there exists an ideal state $\sigma_{S_0^\ell S_1^\ell B' C}$, where C is a random variable on $\{0,1\}$, such that*

1. *Bob is ignorant about* S^ℓ_{1-C}*:*

$$\sigma_{S^\ell_{1-C}S^\ell_C B'C} \approx_\varepsilon \tau_{\{0,1\}^\ell} \otimes \sigma_{S^\ell_C B'C} \ .$$

2. *The real state* $\rho_{S^\ell_0 S^\ell_1 B'}$ *created during the protocol is* ε*-close to the ideal state:*

$$\rho_{S^\ell_0 S^\ell_1 B'} \approx_\varepsilon \sigma_{S^\ell_0 S^\ell_1 B'} \ .$$

Security for Bob: *If Bob is honest, then there exists an ideal state* $\sigma_{A'S^\ell_0 S^\ell_1 C}$ *such that*

1. *Alice is ignorant about* C*:*

$$\sigma_{A'S^\ell_0 S^\ell_1 C} = \sigma_{A'S^\ell_0 S^\ell_1} \otimes \tau_{\{0,1\}} \ .$$

2. *The real state* $\rho_{A'CY^\ell}$ *created during the protocol is* ε*-close to the ideal state:*

$$\rho_{A'CY^\ell} \approx_\varepsilon \sigma_{A'CS^\ell_C} \ ,$$

where we identify $B = (C, Y^\ell)$*.*

In the main part of this text, we had restricted ourselves to considering a simplified protocol containing all the essential ideas of the protocol below. The actual protocol is very similar, but for technical reasons we will work with m blocks of β elements each, instead of sampling individual elements X_j. Fortunately, the protocol we will use for the case of non-uniform weak string erasure remains the same as in the case of weak string erasure with a small modification. Since $p \neq 1/2$, the expected number of $d$its X_j that Bob will learn is of course no longer roughly $n/2$ as in the original setting [13]. This requires the introduction of a new parameter η such that with high probability Bob will learn n/η of the string's entries. We again require an encoding of subsets as strings. Since our subsets will now be smaller, we choose t such that $2^t \leq \binom{m}{m/\eta} \leq 2 \cdot 2^t$, and an injective encoding $\mathsf{Enc} : \{0,1\}^t \to \mathcal{T}$, where $\mathcal{T}$ is the set of possible subsets of size m/η. Note that this again means that not all subsets can be encoded but at least half of them will.

Protocol 2: WSE-to-FROT

Parameters: Set $\eta := 2(d+1)$. Integers n, β such that $m := n/\beta$ is a multiple of η. Outputs: $(s^\ell_0, s^\ell_1) \in \{0,1\}^\ell \times \{0,1\}^\ell$ to Alice, and $(c, y^\ell) \in \{0,1\} \times \{0,1\}^\ell$ to Bob.

1: Alice and Bob: Execute $(n, \lambda, \varepsilon, 1/(d+1), d)$-WSE. Alice gets a string $x^n \in \{0, 1, \ldots, d-1\}^n$, Bob a set $\mathcal{I} \subset [n]$ and a string $s = x_{\mathcal{I}}$. If $|\mathcal{I}| < n/\eta$, then Bob simply chooses $\mathcal{I}_{\text{tr}}$ from all subsets of size $|\mathcal{I}| = n/\eta$ uniformly at random. Otherwise, he randomly truncates $\mathcal{I}$ to $\mathcal{I}_{\text{tr}}$ of size n/η, and deletes the corresponding values in s.

We arrange x^n into a matrix $\mathbf{z} \in \mathbb{M}_{m\times\beta}(\{0, 1, \ldots, d-1\})$, by $\mathbf{z}_{j,\alpha} := x_{(j-1)\cdot\beta+\alpha}$ for $(j, \alpha) \in [m] \times [\beta]$.

2: Bob:

1. Randomly chooses a string $w^t \in_R \{0,1\}^t$ corresponding to an encoding of a subset $\mathsf{Enc}(w^t)$ of $[m]$ with m/η elements.
2. Randomly partitions the n $d$its of x^n into m blocks of β $d$its each: He randomly chooses a permutation $\pi : [m] \times [\beta] \to [m] \times [\beta]$ of the entries of $\mathbf{z}$ such that he knows $\pi(\mathbf{z})_{\mathsf{Enc}(w^t)}$ (that is, these $d$its are permutation of the $d$its of s). Formally, π is uniform over permutations satisfying the following condition: for all $(j, \alpha) \in [m] \times [\beta]$ and $(j', \alpha') := \pi(j, \alpha)$, we have $(j-1)\cdot\beta + \alpha \in \mathcal{I} \Leftrightarrow j' \in \mathsf{Enc}(w^t)$.
3. Bob sends π to Alice.

3: Alice and Bob: Execute interactive hashing with Bob's input equal to w^t. They obtain $w^t_0, w^t_1 \in \{0,1\}^t$ with $w^t \in \{w^t_0, w^t_1\}$.

4: Alice: Chooses $r_0, r_1 \in_R \mathcal{R}$ and sends them to Bob.

5: Alice: Outputs $(s^\ell_0, s^\ell_1) := (\mathsf{Ext}(\pi(\mathbf{z})_{\mathsf{Enc}(w^t_0)}, r_0), \mathsf{Ext}(\pi(\mathbf{z})_{\mathsf{Enc}(w^t_1)}, r_1))$.

6: Bob: Computes c, where $w^t = w^t_c$, and $\pi(\mathbf{z})_{\mathsf{Enc}(w^t)}$ from s. He outputs $(c, y^\ell) := (c, \mathsf{Ext}(\pi(\mathbf{z})_{\mathsf{Enc}(w^t)}, r_c))$.

Theorem 5 (Oblivious Transfer)**.** *For any* $\omega \geq (d+1)$ *and* $\beta \geq \max\{67, 256\omega^2/\lambda^2\}$*, the protocol WSE-to-FROT implements an* $(\ell, 43 \cdot 2^{-\frac{\lambda^2}{512\omega^2\beta}n} + 2\varepsilon)$*-FROT from one instance of of* $(n, \lambda, \varepsilon, p, d)$*-non-uniform WSE, where* $\ell := \left\lfloor \left(\left(\frac{\omega-1}{\omega}\right) \frac{\lambda}{4(d+1)} - \frac{\lambda^2}{512\omega^2\beta} \right) n - \frac{1}{2} \right\rfloor$.

Security for Bob

We first show that the protocol is secure against a cheating Alice. This can again be done following the steps of [13] taking the non-uniformity into account. Formally, let $\tilde{\rho}_{A''CY^\ell}$ denote the joint state at the end of the protocol, where A'' is the quantum output of a malicious Alice and (C, Y^ℓ) is the classical output of an honest Bob. Following the same steps as in [13] we can construct an ideal state $\tilde{\sigma}_{A''W^\ell_0 W^\ell_1 C} = \tilde{\sigma}_{A''W^\ell_0 W^\ell_1} \otimes \tau_{\{0,1\}}$ that satisfies $\tilde{\rho}_{A''CY^\ell} = \tilde{\sigma}_{A''CW^\ell_C} = \tilde{\sigma}_{A''CS^\ell_C}$.

It now again remains to be shown that Alice does not learn anything about C, that is, $\tilde{\sigma}_{A''S^\ell_0 S^\ell_1 C} = \tilde{\sigma}_{A''S^\ell_0 S^\ell_1} \otimes \tau_{\{0,1\}}$. From the properties of non-uniform WSE it follows that $\sigma_{A'\hat{X}^n\mathcal{I}} = \sigma_{A'\hat{X}^n} \otimes \Psi(1/(d+1))$. Since Bob randomly truncates $\mathcal{I}$ to $\mathcal{I}_{\text{tr}}$ such that $|\mathcal{I}_{\text{tr}}| = n/\eta$, the truncated set is independent of A'. Furthermore, although $\mathcal{I}$ is not distributed uniformly over $2^{[n]}$, we can show that

the truncated set $\mathcal{I}_{\rm tr}$ is indeed distributed uniformly over all subsets of size n/η. Intuitively this follows from the fact that the distribution of the set $\mathcal{I}$ depends only on $|\mathcal{I}|$, the number of elements in $\mathcal{I}$. Formally, the probability of a given truncated set $\mathcal{I}_{\rm tr}$ can be written in terms of the probability $p(\bar{A})$ that $|\mathcal{I}| \geq n/\eta$ as follows

$$\begin{aligned} p(\mathcal{I}_{\rm tr}|\bar{A}) &= \sum_{\substack{\mathcal{I}\subseteq[n] \\ |\mathcal{I}|\geq n/\eta}} \frac{p(\mathcal{I}|\bar{A})}{\binom{|\mathcal{I}|}{n/\eta}} p(|\mathcal{I}| \geq n/\eta) \qquad (29) \\ &= \frac{1}{p(\bar{A})} \sum_{\mathcal{I}} \frac{p(\mathcal{I})}{\binom{|\mathcal{I}|}{n/\eta}}, \end{aligned}$$

independent of the choice of truncation. Here $1/\binom{|\mathcal{I}|}{n/\eta}$ is the probability that we pick a particular $\mathcal{I}_{\rm tr}$ from the original $\mathcal{I}$ and $p(\mathcal{I}|\bar{A})$ is the conditional probability of a set $\mathcal{I}$, given that Bob obtains a sufficient number of indices. The last step is simply an application of Bayes' rule, $p(\bar{A})p(\mathcal{I}|\bar{A}) = p(\bar{A}|\mathcal{I})p(\mathcal{I})$ where $p(\bar{A}|\mathcal{I}) = 1$ for the subsets $\mathcal{I}$ in the sum. Note that if $|\mathcal{I}| < n/\eta$ then Bob chooses a subset of the desired size uniformly at random from all subsets of size $|\mathcal{I}| = n/\eta$ and hence $\Pr(\mathcal{I}_{\rm tr})$ is always uniform. Hence, conditioned on any fixed $W^t = w^t$, the permutation Π is uniform and independent of A'. It follows that the string W^t is also uniform and independent of A' and Π. From the properties of interactive hashing we are guaranteed that C is uniform and independent of Alice's view afterwards, and hence,

$$\tilde{\sigma}_{A''S_0^\ell S_1^\ell C} = \tilde{\sigma}_{A''S_0^\ell S_1^\ell} \otimes \tau_{\{0,1\}} \ .$$

Security for Alice

The security proof for the case that Bob is dishonest is analogous to [13], this time employing [13, Lemma 2.5] with a subset size of $|\mathcal{S}| = m/\eta$.

Correctness

It remains to prove that if both parties are honest, then honest Bob can indeed learn the desired S_C. This requires us to show that for our choice of η, Bob can learn sufficiently many indices $i \in [n]$.

Lemma 6 (Correctness)**.** *Protocol WSE-to-FROT satisfies correctness with an error of*

$$43 \cdot 2^{-\frac{\lambda^2}{512\omega^2\beta}n} \ .$$

First we show using the Hoeffding bound [25], that the probability that a subset of $[n]$ where each entry is chosen with probability $p = 1/(d+1)$ has less than n/η elements is at most $\exp(-2n/\eta^2)$. Consider a sequence of independent random variables $\{X_1, \ldots, X_n\}$, which are bounded as follows: $\Pr(X_i - \mathbf{E}(X_i) \in [a_i, b_i]) = 1$, $\forall 1 < i < n$. Then, Hoeffding's inequality states that the sum $S = X_1 + \ldots + X_n$ satisfies,

$$\Pr(\mathbf{E}(S) - S \geq t) \leq \exp\left(-\frac{2t^2}{\sum_{i=1}^n (b_i - a_i)^2}\right) \qquad (30)$$

In our context, X_i is the binary variable which takes on the value 1 if the index $i \in \mathcal{I}$, and 0 otherwise. The sum S is thus simply equal to $|\mathcal{I}|$, the number of elements in the index set $\mathcal{I}$, which is a random subset of $[n]$. For the case of $d+1$ encodings, $\Pr(X_i = 1) = 1/(d+1)$ and $\Pr(X_i = 0) = d/(d+1)$, so that the expectation value satisfies

$$\mathbf{E}(S) = \mathbf{E}(|\mathcal{I}|) = \frac{n}{d+1} \ . \qquad (31)$$

Furthermore, we can take $a_i = 0$ and $b_i = 1$ for all i. Applying Hoeffding's inequality to the sum $S = |\mathcal{I}|$ gives

$$\begin{aligned} \Pr(\frac{n}{d+1} - |\mathcal{I}| \geq \frac{n}{d+1} - \frac{n}{\eta}) \leq \\ \exp\left(-2n\left[\frac{1}{d+1} - \frac{1}{\eta}\right]^2\right). \qquad (32) \end{aligned}$$

Rearranging terms, we obtain the probability that a random set $\mathcal{I}$ has less than n/η elements:

$$\Pr(|\mathcal{I}| \leq \frac{n}{\eta}) \leq \exp\left(-2n\left[\frac{1}{d+1} - \frac{1}{\eta}\right]^2\right). \qquad (33)$$

Since our work is mainly a proof of principle, we do not yet care about optimality or efficiency. We simply pick a choice of η that will satisfy this condition, and set $\eta = 2(d+1)$. Thus, the probability that a random subset of $[n]$ has less than n/η elements is at most $\exp(-2n/\eta^2)$.

Let $\xi := 2^{-n/\eta^2}$. We have to show that the state $\tilde{\rho}_{S_0^\ell S_1^\ell C Y^\ell}$ at the end of the protocol is close to the given ideal state $\tilde{\sigma}_{S_0^\ell S_1^\ell C S_C^\ell}$. As shown above, the probability that a subset of $[n]$ has less than n/η elements is at most

$$\exp(-2n/\eta^2) \leq \xi \ . \qquad (34)$$

Hence, the probability that Bob does not learn sufficiently many indices when both parties are honest is at most ξ. Let $\mathcal{A}$ be the event that $|\mathcal{I}| \geq n/\eta$. It remains to show that the state $\tilde{\rho}_{S_0^\ell S_1^\ell C Y^\ell|\mathcal{A}}$ is close to the given ideal state $\sigma_{S_0^\ell S_1^\ell C S_C^\ell}$.

Note that the correctness condition of WSE ensures that the state created by WSE is equal to $\rho_{X^n \mathcal{I} X_{\mathcal{I}}} = \sigma_{X^n \mathcal{I} X_{\mathcal{I}}}$, where $\sigma_{X^n \mathcal{I}} = \tau_{\{0,1,\ldots,d-1\}^n} \otimes \Psi(1/(d+1))$. Since $\mathcal{I}_0$ and $\mathcal{I}_1$ are chosen independently of X^n, $X_{\mathcal{I}_0}$ and $X_{\mathcal{I}_1}$ have a min-entropy of n/η each. Since $\ell \leq n/2\eta \leq n/\eta - 2\log 1/\xi$, it follows from privacy amplification that S_C^ℓ is independent and ξ-close to uniform. Since dishonest Bob is only more powerful than honest Bob, we furthermore have from the proof against dishonest Bob

that S^ℓ_{1-C} is independent and uniform except with an error of at most $\hat{\varepsilon} = 41 \cdot 2^{-\frac{\lambda^2}{512\omega^2\beta}n}$, where we used the fact that Bob is also honest during weak string erasure ($\varepsilon = 0$). Finally, by the same arguments showing security for Bob we have that C is uniform and independent of S^ℓ_0 and S^ℓ_1. Hence,

$$\rho_{S^\ell_0 S^\ell_1 C|\mathcal{A}} \approx_{\xi+\hat{\varepsilon}} \sigma_{S^\ell_0 S^\ell_1 C} \cdot$$

Since the extra condition on the permutation Π implies that Bob can indeed calculate $\Pi(\mathbf{Z})_{\mathsf{Enc}(W)}$ from $X_{\mathcal{I}}$, we have that $Y^\ell = S^\ell_C$. Using $\Pr[\mathcal{A}] \geq 1 - \xi$, we get

$$\rho_{S^\ell_0 S^\ell_1 C Y^\ell} \approx_{2\xi+\hat{\varepsilon}} \sigma_{S^\ell_0 S^\ell_1 C S^\ell_C} \cdot$$

Finally, $\lambda \leq 1$, $\beta > 1$ and $\omega \geq (d+1)$ give us $1/\eta^2 = 1/(4(d+1)^2) > \lambda^2/(512\omega^2\beta)$. Adding up all errors and noting that

$$2 \cdot 2^{-\frac{1}{\eta^2}n} \leq 2 \cdot 2^{-\frac{\lambda^2}{512\omega^2\beta}n} ,$$

gives our claim.