Oblivious Transfer in the Bounded Storage Model with Errors

Rafael Dowsley
Institute of Theoretical Informatics
Karlsruhe Institute of Technology
Am Fasanengarten 5, Geb. 50.34
76131 Karlsruhe, Germany
Email: rafael.dowsley@kit.edu

Felipe Lacerda
Department of Computer Science
University of Brasília
Campus Darcy Ribeiro
70910-900, Brasília, Brazil
Email: fegolac@unb.br

Anderson C. A. Nascimento
Department of Electrical Engineering
University of Brasília
Campus Darcy Ribeiro
70910-900, Brasília, Brazil
Email: andclay@ene.unb.br

***Abstract*—In the bounded storage model the memory of the adversarial parties is restricted, instead of their computational power. This different restriction allows the construction of protocols with information-theoretical (instead of only computational) security. We present the first protocol for oblivious transfer in the bounded storage model with errors, i.e., where the public random sources available to the two parties are not exactly the same, but instead are only required to have a small Hamming distance between themselves, and the memory of the (adversarial) receiver is limited. Oblivious transfer protocols were known previously only for the error-free variant of the bounded storage model, which is harder to realize.**

I. Introduction

Oblivious transfer (OT) protocols are fundamental building blocks of secure two- and multi-party computation protocols. OT is a two-party protocol in which Alice inputs two strings s_0 and s_1, and Bob inputs a bit c. Bob's output is the string s_c. The protocol is called secure if Alice never learns the choice bit c and Bob does not learn any information about s_{1-c}. It can be used to realize *any* secure two-party computation [1].

In the setting where the parties only communicate through noiseless channels, unconditionally secure OT is impossible (even if quantum channels are available [2]). However, it is possible to realize it in the context of computational security (in which the adversaries are restricted to be polynomial-time Turing machines), as long as computational hardness assumptions are made. OT can be obtained using generic assumptions such as the existence of dense trapdoor permutations [3] or assuming the hardness of many specific computational problems [4], [5], [6], [7], [8]. One possibility to obtain unconditional secure OT is resorting to physical assumptions such as the existence of noisy channels [9].

In this paper a different setting is considered, the so called *bounded storage model* (BSM) [10], in which the adversary is assumed to have bounded memory. In the BSM, it is assumed that both parties have access to a public random source, and that a dishonest party cannot store the whole source.

A. The Bounded Storage Model

Many cryptographic tasks can be implemented in the BSM. Cachin and Maurer [11] proposed a key agreement protocol in which the parties have a small pre-shared key and also a protocol for key agreement by public discussion (i.e., without a pre-shared key) that requires $\sqrt{n}$ samples from the source and is thus less practical. The first OT protocol was obtained by Cachin et al. [12]. Improvements (in a slightly different model) were presented by Ding [13] and Hong et al. [14]. Ding et al. [15] obtained the first constant-round OT protocol.

Unfortunately the bounded storage model assumes that there exists a random source that can be reliably broadcasted to all parties, without errors in the transmission, and this is hard to realize in practice. Our goal with this work is to study two-party protocols under more realistic assumptions.

B. Our contribution

In this work, a more general variant of the BSM is considered, in which errors can be introduced in the public random source in arbitrary positions. It is only assumed that the fraction of errors, relative to the length of the public string, is not too large. This captures both adversarially introduced errors and natural ones. This model was previously studied by Ding [16], who defined a general paradigm for randomness extraction schemes and showed how to incorporate error correction in key agreement protocols by using fuzzy extractors.

We propose the first protocol for oblivious transfer in this model, thus showing that any multi-party computation protocol can be realized. Our protocol is similar to the one by Ding et al. [15], but uses error correction (i.e., a fuzzy extractor) to ensure correctness.

II. Preliminaries

The probability distribution of a random variable X will be denoted by P_X. The set $\{1, \ldots, n\}$ will be written as $[n]$. If $x = (x_1, \ldots, x_n)$ is a sequence and $S = \{s_1, \ldots, s_t\} \subseteq [n]$, x^S denotes the sequence $(x_{s_1}, \ldots, x_{s_t})$. $u \xleftarrow{\$} U$ denotes that u is drawn from the uniform distribution over the set U and U_r is the uniformly-distributed r-bit random variable. $y \xleftarrow{\$} \mathcal{F}(x)$ denotes the act of running the probabilistic algorithm $\mathcal{F}$ with input x and obtaining the output y. $y \leftarrow \mathcal{F}(x)$ is similarly used for deterministic algorithms.

$\mathrm{HD}(X, Y)$ denotes the hamming distance between X and Y, and $X \oplus Y$ their bitwise exclusive or. $H(X)$ denotes the entropy of X and $I(X;Y)$ the mutual information between X

and Y. The logarithms are in the base 2 and the binary entropy function is denoted by h. $\|P_X - P_Y\|$ denotes the statistical distance between P_X and P_Y.

A. Entropy Measures

The main entropy measure in this work is the *smooth min-entropy*. The min-entropy captures the notion of unpredictability of a random variable (i.e., the extractable private randomness from the variable X, given the correlated random variable Y possessed by an adversary), but has the problem of being sensitive to small changes in the probability distribution. Due to this fact, smooth min-entropy [17] will be used.

Definition 1 (Smooth min-entropy): Let $\varepsilon > 0$ and P_{XY} be a probability distribution. The ε-smooth min-entropy of X given Y is defined by

$$H_\infty^\varepsilon(X|Y) = \max_{X'Y':\|P_{X'Y'} - P_{XY}\| \leq \varepsilon} H_\infty(X'|Y')$$

where $H_\infty(X \mid Y) = \min_{y \in \mathcal{Y}} \min_{x \in \mathcal{X}} (-\log P_{X|Y=y}(x))$. Intuitively, the smooth min-entropy is the maximum min-entropy in the neighborhood of the probability distribution. Analogues of the chain rule for conditional Shannon entropy were establish for smooth min-entropy by Renner and Wolf [17]. X is called a k-source if $H_\infty(X) \geq k$.

Definition 2 (Min-entropy rate): Let X be a random variable with an alphabet $\mathcal{X}$, E a random variable E and $\varepsilon \geq 0$. The min-entropy rate is defined as $R_\infty^\varepsilon(X \mid E) = \frac{H_\infty^\varepsilon(X|E)}{\log|\mathcal{X}|}$.

The following lemma is a restatement of a lemma in [15] and is what makes the bounded assumption useful: it says that a source with high min-entropy rate also has high min-entropy rate when conditioned on a correlated short string.

Lemma 1: Let $X \in \{0,1\}^n$ such that $R_\infty^\varepsilon(X) \geq \rho$ and Y be a random variable over $\{0,1\}^{\phi n}$. Fix $\varepsilon' > 0$. Then

$$R_\infty^{\varepsilon' + \sqrt{8\varepsilon}}(X \mid Y) \geq \rho - \phi - \frac{1 + \log(1/\varepsilon')}{n}.$$

Proof: Let $\psi = \rho - \phi - \frac{1+\log(1/\varepsilon')}{n}$. By lemma 3.16 in [15] we have that if $R_\infty^\varepsilon(X) \geq \rho$ then

$$\Pr_{y \xleftarrow{\$} Y}\left[R_\infty^{\sqrt{2\varepsilon}}(X \mid Y = y) \geq \psi\right] \geq 1 - \varepsilon' - \sqrt{2\varepsilon}.$$

Let $\mathcal{G} = \{y \in \mathcal{Y} \mid R_\infty^{\sqrt{2\varepsilon}}(X \mid Y = y) \geq \psi\}$. Let P'_{XY} be the distribution that is $\sqrt{2\varepsilon}$-close to P_{XY} and is such that $P'(X = x \mid Y = y) \leq 2^{-\psi n}$ for any $x \in \mathcal{X}, y \in \mathcal{G}$. Let P''_{XY} be obtained by letting $P''(X \mid Y = y) = P'(X \mid Y = y)$ for $y \in \mathcal{G}$ and defining $P''(X = x \mid Y = y) = 2^{-n}$ for any $x \in \mathcal{X}$, $y \notin \mathcal{G}$. As $\Pr[\mathcal{G}] \geq 1 - \varepsilon' - \sqrt{2\varepsilon}$, it holds that $\|P''_{XY} - P'_{XY}\| \leq \varepsilon' + \sqrt{2\varepsilon}$ and so $\|P''_{XY} - P_{XY}\| \leq \varepsilon' + 2\sqrt{2\varepsilon}$. Since $P''(X = x \mid Y = y) \leq 2^{-\psi n}$ for every $x \in \mathcal{X}, y \in \mathcal{Y}$, the lemma follows. ∎

B. Averaging Samplers and Randomness Extractors

In the bounded storage model a typical approach for the usage of the source is the sample-then-extact paradigm, in which first some positions of the source are sampled and then an extractor is applied on these positions. Due to the infeasibility of storing the whole source string, any extractor should be locally computable [18]. In this context, averaging samplers are a fundamental tool. They produce samples such that the average value of any function applied to the sampled string is roughly the same as the average when taken over the original string. Important to this work is the fact that averaging samplers roughly preserve the *min-entropy rate*.

Definition 3 (Averaging sampler): A function $\mathrm{Samp}\colon \{0,1\}^r \to [n]^t$ is a (μ, ν, ε)-averaging sampler if for every function $f\colon [n] \to [0,1]$ with average $\frac{\sum_{i=1}^n f(i)}{n} \geq \mu$:

$$\Pr_{\mathcal{S} \xleftarrow{\$} \mathrm{Samp}(U_r)}\left[\frac{1}{t}\sum_{i \in \mathcal{S}} f(i) \leq \mu - \nu\right] \leq \varepsilon.$$

Lemma 2 ([18]): Let $X \in \{0,1\}^n$ be such that $R_\infty(X \mid E) \geq \rho$. Let τ be such that $1 \geq \rho \geq 3\tau > 0$ and $\mathrm{Samp}\colon \{0,1\}^r \to [n]^t$ be an (μ, ν, ε)-averaging sampler with distinct samples for $\mu = (\rho - 2\tau)/\log(1/\tau)$ and $\nu = \tau/\log(1/\tau)$. Then for $\mathcal{S} \xleftarrow{\$} \mathrm{Samp}(U_r)$ and $\varepsilon' = \varepsilon + 2^{-\Omega(\tau n)}$

$$R_\infty^{\varepsilon'}(X^{\mathcal{S}} \mid \mathcal{S}, E) \geq \rho - 3\tau.$$

The (n,t)-*random subset sampler* picks a random subset of $[n]$ of size t. It is an averaging sampler.

Lemma 3: Let $0 < t < n$. For any $\mu, \nu > 0$, the (n,t)-random subset sampler is a $(\mu, \nu, e^{-t\nu^2/2})$-averaging sampler.

Proof: It is just a restatement of Lemma 5.5 in [19]. ∎

A *randomness extractor* is a function that takes a string with high min-entropy as an input and outputs a string that is close (in the statistical distance sense) to a uniformly distributed string. The OT protocol presented in this work will use a variant of a strong extractor, called a *fuzzy extractor* [20]. Intuitively, fuzzy extractors are noise-resilient extractors, that is, extractors such that the extracted string can be reproduced by any party with a string that is close (in the Hamming distance sense) to the original source.

Definition 4 (Fuzzy extractor): A pair of functions $\mathrm{Ext}\colon \{0,1\}^n \times \{0,1\}^r \to \{0,1\}^m \times \{0,1\}^p$, $\mathrm{Rec}\colon \{0,1\}^n \times \{0,1\}^r \times \{0,1\}^p \to \{0,1\}^m$ is a $(k, \varepsilon, \delta, \beta)$-fuzzy extractor if:

- (Security) For every k-source $X \in \{0,1\}^n$, let $R \xleftarrow{\$} U_r$, $(Y, P) \leftarrow \mathrm{Ext}(X, R)$. Then $\|P_{YRP} - P_{U_m} \times P_{RP}\| \leq \varepsilon$.
- (Recovery) For every $X, X' \in \{0,1\}^n$ such that $\mathrm{HD}(X, X') \leq \delta n$, let $R \xleftarrow{\$} U_r$, $(Y, P) \leftarrow \mathrm{Ext}(X, R)$. It should hold that $\Pr[\mathrm{Rec}(X', R, P) = Y] \geq 1 - \beta$.

Since there is a restriction to close strings with respect to the Hamming distance, syndrome-based fuzzy extractors can be used, as summarized in the following lemma from Ding [16].

Lemma 4 ([16]): Let $1 \geq \rho, \psi, \chi > 0$ be arbitrary constants. There is a constant $\delta > 0$ such that for every sufficiently large $n \in \mathbb{N}$, and every $\varepsilon > e^{-n/2^{O(\log^* n)}}$, there is an explicit construction of a $(\rho n, \varepsilon, \delta, 0)$-fuzzy extractor $(\mathrm{Ext}, \mathrm{Rec})$, where $\mathrm{Ext}\colon \{0,1\}^n \times \{0,1\}^r \to \{0,1\}^m \times \{0,1\}^p$ with $m = (1 - \psi)\rho n$, $r = O(\log n + \log(1/\varepsilon))$ and $p \leq \chi m$.

The following lemma shows that random subsets of two sets X and Y have relative Hamming distances that are close to the one between X and Y.

Lemma 5: Let $X, Y \in \{0,1\}^n$, $\mathcal{S}$ be a random subset of $[n]$ of size r and consider any $\nu \in [0,1]$. On one hand, if $\mathrm{HD}(X,Y) \leq \delta n$, then $\mathrm{HD}(X^{\mathcal{S}}, Y^{\mathcal{S}}) < (\delta+\nu)r$ except with probability $e^{-r\nu^2/2}$. On the other hand, if $\mathrm{HD}(X,Y) \geq \delta n$, then $\mathrm{HD}(X^{\mathcal{S}}, Y^{\mathcal{S}}) > (\delta-\nu)r$ except with probability $e^{-r\nu^2/2}$.

Proof: Lets begin with the first part of the Lemma. By Lemma 3, a random subset sampler is an $(\mu, \nu, e^{-r\nu^2/2})$-averaging sampler for any $\mu, \nu > 0$. Let $f(i)$ be defined as 0 if $X_i \neq Y_i$, and 1 otherwise. Fix $\mu = 1-\delta$. Note that $\frac{1}{|\mathcal{S}|}\sum_{i\in\mathcal{S}} f(i) = 1 - \frac{\mathrm{HD}(X^{\mathcal{S}},Y^{\mathcal{S}})}{r}$ and $\frac{1}{n}\sum_{i=1}^{n} f(i) = 1 - \frac{\mathrm{HD}(X,Y)}{n} \geq \mu$. Due to the properties of averaging samplers

$$\begin{aligned} e^{-r\nu^2/2} &\geq \Pr\left[\frac{1}{|\mathcal{S}|}\sum_{i\in\mathcal{S}} f(i) \leq \mu - \nu\right] \\ &= \Pr\left[1 - \frac{\mathrm{HD}(X^{\mathcal{S}},Y^{\mathcal{S}})}{r} \leq 1 - \delta - \nu\right] \\ &= \Pr\left[\mathrm{HD}(X^{\mathcal{S}},Y^{\mathcal{S}}) \geq (\delta+\nu)r\right] \end{aligned}$$

which proves the first part of the Lemma. The second part uses the same idea, but now $f(i) = 0$ iff $X_i = Y_i$. ∎

The following statement of the birthday paradox is standard.

Lemma 6 ([13]): Let $A, B \subset [n]$, chosen independently at random, with $|A| = |B| = 2\sqrt{\ell n}$. Then

$$\Pr[|A \cap B| < \ell] < e^{-\ell/4}.$$

C. Interactive Hashing and Binary Encoding of Subsets

The oblivious transfer protocol proposed in this paper uses *interactive hashing* as a subprotocol. Initially introduced in the context of computationally-secure cryptography [21], interactive hashing was later generalized to the information-theoretic setting, and is particularly useful in the context of oblivious transfer protocols [12], [15], [22], [23]. It is a protocol where Bob inputs a string W and Alice and Bob output two strings W_0, W_1, in such a way that one of the output strings is equal to W, and the other string is completely random from Bob's point of view, even if he is dishonest. A variety of protocols for realizing interactive hashing have been proposed [12], [15], [24]. In this work interactive hashing is used as a black box.

Definition 5 (Interactive hashing): *Interactive hashing* is a protocol between two parties, Alice and Bob, in which Bob inputs $W \in \{0,1\}^m$ and Alice inputs nothing, and both parties output $W_0, W_1 \in \{0,1\}^m$, in lexicographic order, such that $W_d = W$ for some $d \in \{0,1\}$. The protocol is called an η-uniform (t,θ)-secure interactive hashing protocol if: (1) If both parties are honest, then W_{1-d} is η-close to completely random, (2) Alice's view of the protocol is independent of d, and (3) for any $T \subset \{0,1\}^m$ such that $|T| \leq 2^t$, it should hold that $\Pr[W_0, W_1 \in T] \leq \theta$, where the probability is taken over the randomness used by Alice and Bob.

Lemma 7 ([15]): Let t, m be positive integers such that $t \geq \log m + 2$. Then there exists a four-message (2^{-m})-uniform $(t, 2^{-(m-t)+O(\log m)})$-secure interactive hashing protocol.

A secure interactive hashing scheme guarantees that one of the outputs is random; however, in the oblivious transfer protocols, the two binary strings are not used directly, but as encodings of subsets. Thus for the protocol to succeed, both outputs need to be valid encodings of subsets of $\binom{[n]}{\ell}$. We use the 'dense" encoding of subsets technique that ensures that most m-bit strings are valid encodings.

Lemma 8 ([15]): Let $\ell \leq n$, $m \geq \lceil \log \binom{n}{\ell} \rceil$, $t_m = \lfloor 2^m / \binom{n}{\ell} \rfloor$. Then there exists an injective mapping $F_m \colon \binom{[n]}{\ell} \times [t_m] \to [2^m]$ with $|\operatorname{Im}(F_m)| > 2^m - \binom{n}{\ell}$.

III. Oblivious Transfer in the Bounded Storage Model

A. Security Model

a) Transmission Phase.: Prior to the realization of the protocols' main part, a transmission phase is executed. In this phase, the sender (Alice) has access to an αn-source $X \in \{0,1\}^n$ and the receiver (Bob) to $\tilde{X} \in \{0,1\}^n$ such that $\mathrm{HD}(X, \tilde{X}) \leq \delta n$. Note that this captures both the situation where the source is noisy and the situation where Alice controls part of the source. Bob then computes a randomized function $f(\tilde{X})$ with output size smaller or equal to γn for $\gamma < \alpha$, stores its output and discards $\tilde{X}$. [1] This is the so called bounded storage assumption. After that, Alice and Bob output bits A_{TP} and B_{TP}, respectively, with the value 1 if they want to continue the protocol and 0 if they want to abort.

The definition of oblivious transfer used is the one presented in [15]. An oblivious transfer protocol is a protocol between two players, Alice and Bob, in which Alice inputs two strings $s_0, s_1 \in \{0,1\}^m$ and outputs nothing, and Bob inputs $c \in \{0,1\}$ and outputs $s \in \{\perp, s_c\}$. In the following, $\mathrm{view}_{A^*}(s_0, s_1; c)$ denotes the view of Alice using a strategy A^* with honest Bob, and $\mathrm{view}_{B^*}(s_0, s_1; c)$ denotes the view of Bob using a strategy B^* with honest Alice. Bob's strategy has bounded storage, but Alice's strategy can be unbounded.

Intuitively, the protocol is secure for Bob if Alice's view does not depend on the choice bit c, and secure for Alice if Bob cannot obtain any information about s_{1-c}. However this is tricky to formalize since a malicious Bob can choose to play with a different bit depending on the public random source and the messages exchanged before Alice's secret is used.

In order to have a general definition of security, the main part of the oblivious transfer protocols is further divided into two phases: the *setup phase*, consisting of all the communication before Alice uses her secrets, and the *transfer phase*, which goes up until the point where Bob outputs s. By the end of the setup phase, Bob must have chosen a bit i, which may be different from c and can depend on all the messages exchanged thus far. To guarantee Alice's security it is thus required that there is an index i, determined at the setup phase, such that for any two pairs $(s_0, s_1), (s'_0, s'_1)$ with $s_i = s'_i$ the distributions of s_{1-i} and s'_{1-i} are close given Bob's view. Following the terminology of [15], pairs $(s_0, s_1), (s'_0, s'_1)$ satisfying $s_i = s'_i$ will be called i-*consistent*. To account for aborts, it is assumed that at the end of the setup phase, Alice and Bob output bits A_{SP}, B_{SP}, respectively, which are 1 if they want to continue.

[1] In order to achieve security it is not necessary to impose any storage bound on Alice, but in the proposed protocol an honest Alice stores the same amount of information as an honest Bob.

b) Security.: A protocol is called $(\lambda_C, \lambda_B, \lambda_A)$-secure if it satisfies the following properties:

1) λ_C-correct: if Alice and Bob are honest, then
$$\Pr[A_{TP} = B_{TP} = A_{SP} = B_{SP} = 1 \wedge s = s_c] \geq 1 - \lambda_C$$
2) λ_B-secure for Bob: for any strategy A^* used by Alice,
$$\|\{\text{view}_{A^*}(s_0, s_1; 0)\} - \{\text{view}_{A^*}(s_0, s_1; 1)\}\| \leq \lambda_B$$
3) λ_A-secure for Alice: for any strategy B^* used by Bob with input c, there exists a random variable i, defined at the end of the setup stage, such that for every two i-consistent pairs $(s_0, s_1), (s'_0, s'_1)$, we have
$$\|\{\text{view}_{B^*}(s_0, s_1; c)\} - \{\text{view}_{B^*}(s'_0, s'_1; c)\}\| \leq \lambda_A$$

B. An Oblivious Transfer Protocol

The idea of the protocol is that initially both parties samples some positions from the public random source. Then an interactive hashing protocol (with an associated dense encoding) is used to select two subsets of the positions sampled by Alice. The input of Bob to the interactive hashing is one subset for which he has also sampled the public random source in that positions. The other subset is out of Bob's control due to the properties of the interactive hashing protocol. Finally the two subsets are used as input to a fuzzy extractor in order to obtain one-time pads. Bob sends one bit indicating which input string should be xored with which one-time pad. The security for Alice is guarantee by the fact that one of the subsets is out of Bob's control and will have high min-entropy given his view, thus resulting in a good one-time pad. The security for Bob follows from the security of the interactive hashing. The correctness follows from the correctness of the fuzzy extractor.

The protocol is defined below. We assume that in the transmission phase, an αn-source $X \in \{0,1\}^n$ is available to Alice, and $\tilde{X} \in \{0,1\}^n$ with $\text{HD}(X, \tilde{X}) \leq \delta n$ is available to Bob. The security parameter is ℓ and k is set as $k = 2\sqrt{\ell n}$. Fix $\varepsilon', \hat{\varepsilon}, \xi > 0$ and let $\rho = \alpha - \gamma - \frac{1+\log(1/\varepsilon')}{n}$. Fix $0 < \zeta < 1$ and τ such that $\frac{\rho}{3} \geq \tau > 0$. Let $\varepsilon'' = e^{-\ell\nu^2/2} - 2^{-\Omega(\tau n)}$, where the last term comes from Lemma 2, and let $\tilde{\varepsilon} = (\varepsilon' + \varepsilon'')^{1-\zeta}$. It is assumed that the following functionalities, which are possible due to the lemmas in Section II, are available to the parties:

- A pair of functions $\text{Ext}\colon \{0,1\}^\ell \times \{0,1\}^r \to \{0,1\}^{m_F \ell} \times \{0,1\}^p$ and $\text{Rec}\colon \{0,1\}^\ell \times \{0,1\}^r \times \{0,1\}^p \to \{0,1\}^{m_F \ell}$ that constitutes an $(k_F \ell, \varepsilon_F, \delta + \xi, 0)$-fuzzy extractor with $k_F = \rho - 3\tau - 2m_F - \frac{1+\log(1/\hat{\varepsilon})}{\ell}$ and ε_F negligible in ℓ. Notice that we should have $\delta + \xi$ in the range allowed by Lemma 4.
- An 2^{-m}-uniform (t, θ)-secure interactive hashing protocol where θ is negligible in ℓ, $t \geq m - \zeta \log(1/(\varepsilon' + \varepsilon''))$ and $m \geq 2\ell \log k$. Let F_m be a dense encoding of the subsets of size ℓ of a set of size k.

Transmission phase:

- Alice chooses uniformly k positions from X. Similarly, Bob samples k positions from $\tilde{X}$. We call their sets of positions A and B, respectively.

Setup phase:

- Alice sends A to Bob. Bob computes $D = A \cap B$. If $|D| < \ell$, Bob aborts. Otherwise, Bob picks a random subset C of D of size ℓ.
- Bob computes the encoding W of C (as a subset of A). Alice and Bob interactively hash W, producing two strings W_0, W_1. They compute the subsets $C_0, C_1 \subset A$ that are respectively encoded in W_0, W_1. If either encoding is invalid, they abort.

Transfer phase:

- Bob sends $e = c \oplus d$, where $W_d = W$.
- For $i \in \{0,1\}$, Alice picks $R_i \xleftarrow{\$} \{0,1\}^r$, computes $(Y_i, P_i) \leftarrow \text{Ext}(X^{C_i}, R_i)$ and $Z_i \leftarrow s_{i \oplus e} \oplus Y_i$, and sends (Z_i, R_i, P_i) to Bob.
- Bob outputs $s = \text{Rec}(\tilde{X}^C, R_d, P_d) \oplus Z_d$.

C. Proof of security for the oblivious transfer protocol

In this section it is proved that the protocol presented above is $(\lambda_C, 0, \lambda_A)$–secure for λ_C and λ_A negligible in ℓ.

Lemma 9: The protocol is λ_C-correct for λ_C negligible in ℓ.

Proof: The probability of an abort is analyzed first. The protocol will abort if either $|D| < \ell$, or if one string obtained in the interactive hashing protocol is an invalid encoding of subsets of A. By Lemma 6, $\Pr[|D| < \ell] < e^{-\ell/4}$. Out of the two outputs of the interactive hashing protocol, one of them is always a valid encoding (since $W_d = W$, which is the encoding of C). The other output W_{1-d} is 2^{-m}-close to distributed uniformly over the $2^{-m} - 1$ strings different from W_d. Since it is a dense encoding, Lemma 8 implies that the probability that it is not a valid encoding is thus less than or equal to $2^{-m} + \frac{\binom{k}{\ell}}{2^{-m}-1} \leq 2^{-m} + 2^{\ell \log k - m + 1} \leq 4k^{-\ell}$ for $m \geq 2\ell \log k$.

If both parties are honest and there is no abort, then $s = s_c$ if and only if $\text{Rec}(\tilde{X}^C, R_d, P_d) = Y_d$. By the properties of the employed fuzzy extractor, this last event happens if $\text{HD}(X^C, \tilde{X}^C) \leq (\delta + \xi)\ell$. By Lemma 5, $\text{HD}(X^C, \tilde{X}^C) > (\delta + \xi)\ell$ with probability at most $e^{-\xi^2 \ell/2}$. Putting everything together this concludes the proof. ∎

Lemma 10: The protocol is 0-secure for Bob.

Proof: There are two possibilities: either the protocol aborts or not. If the protocol aborts in the setup phase, Bob still has not sent $e = c \oplus d$, so Alice's view is independent from c. On the other hand, if the protocol does not abort, then W_{1-d} is a valid encoding of some set C'. Due to the properties of the interactive hashing protocol, Alice's view is then consistent with both Bob choosing c and C, and Bob choosing $1 - c$ and C'. Hence Alice's view is independent of c. Thus the protocol is 0-secure for Bob. ∎

Lemma 11: The protocol is λ_A-secure for Alice for λ_A negligible in ℓ.

Proof: There should be an index i (determined at the setup stage) such that for any $(s_0, s_1), (s'_0, s'_1)$ with $s_i = s'_i$, Bob's view when the protocol is executed with (s_0, s_1) is close to his view when executed with (s'_0, s'_1). The view of Bob is

given by the function computed from the public random source $f(\tilde{X})$ along with all the messages exchanged and his local randomness. The proof's strategy is to show that for i, $X^{C_{1-i}}$ has high enough min-entropy, given Bob's view, in such a way that Y_{1-i} is indistinguishable from an uniform distribution. Indistinguishability of Bob's views will then follow.

By the bounded storage assumption, $|f(\tilde{X})| \leq \gamma n$ with $\gamma < \alpha$. Then, by Lemma 1,

$$R_{\infty}^{\varepsilon'}(X \mid f(\tilde{X})) \geq \alpha - \gamma - \frac{1+\log(1/\varepsilon')}{n} = \rho.$$

Since Alice is honest, A is randomly chosen. Lets consider a random subset $\tilde{C}$ of A such that $|\tilde{C}| = \ell$. This is an $(\mu, \nu, e^{-\ell\nu^2/2})$-averaging sampler for any $\mu, \nu > 0$ according to Lemma 3. By setting $\mu = \frac{\rho-2\tau}{\log(1/\tau)}$, $\nu = \frac{\tau}{\log(1/\tau)}$, we have by Lemma 2 that

$$R_{\infty}^{\varepsilon'+\varepsilon''}(X^{\tilde{C}} \mid A, \tilde{C}, f(\tilde{X})) \geq \rho - 3\tau$$

for $\varepsilon'' = e^{-\ell\nu^2/2} - 2^{-\Omega(\tau n)}$. For $\tilde{\varepsilon} = (\varepsilon' + \varepsilon'')^{1-\zeta}$, let Bad be the set of $\tilde{C}$'s such that $R_{\infty}(X^{\tilde{C}} \mid A, \tilde{C}, f(\tilde{X}))$ is not $\tilde{\varepsilon}$-close to $(\rho - 3\tau)$-min entropy rate. Due to the above equation the density of Bad is at most $(\varepsilon' + \varepsilon'')^{\zeta}$. Then the size of the set $T \subset \{0,1\}^m$ of strings that maps to subsets in Bad is at most $(\varepsilon' + \varepsilon'')^{\zeta} 2^m \leq 2^t$. Hence the properties of the interactive hashing protocol guarantee that with probability greater than or equal to $1 - \theta$ there will be an i such that

$$R_{\infty}^{\tilde{\varepsilon}}(X^{C_{1-i}} \mid A, C_{1-i}, f(\tilde{X}), M_{IH}) \geq \rho - 3\tau$$

where M_{IH} are the messages exchanged during the interactive hashing protocol. We now show that $X^{C_{1-i}}$ has high min-entropy even when given Z_i, Y_i. We can see (Z_i, Y_i) as a random variable over $\{0,1\}^{2m_F\ell}$. Then, by Lemma 1,

$$R_{\infty}^{\tilde{\varepsilon}+\sqrt{8\tilde{\varepsilon}}}(X^{C_{1-i}} \mid A, C_{1-i}, f(\tilde{X}), M_{IH}, Z_i, Y_i) \geq \rho - 3\tau - 2m_F - \frac{1+\log(1/\hat{\varepsilon})}{\ell} = k_F.$$

Thus setting ε' and $\hat{\varepsilon}$ to be negligible in ℓ, the use of the $(k_F\ell, \varepsilon_F, \delta+\xi, 0)$-fuzzy extractor to obtain Y_i (that is used as an one-time pad) guarantees that only negligible information about $s_{i\oplus e}$ can be leaked and that the protocol is λ_A-secure for Alice for negligible λ_A. ∎

IV. Conclusion

This work presented the first protocol for oblivious transfer in the bounded storage model with errors. As expected, the protocol works for a limited range of the noise parameter δ.

References

[1] J. Kilian, "Founding crytpography on oblivious transfer," ser. STOC '88. New York, NY, USA: ACM, 1988, pp. 20–31.

[2] D. Mayers, "Unconditionally secure quantum bit commitment is impossible," *Physical review letters*, vol. 78, no. 17, pp. 3414–3417, 1997.

[3] I. Haitner, "Implementing oblivious transfer using collection of dense trapdoor permutations," in *TCC 2004*, ser. LNCS, M. Naor, Ed., vol. 2951. Cambridge, MA, USA: Springer, Berlin, Germany, Feb. 19–21, 2004, pp. 394–409.

[4] M. O. Rabin, "How to exchange secrets by oblivious transfer," Technical Report TR-81, Harvard Aiken Computation Laboratory, Tech. Rep., 1981.

[5] M. Bellare and S. Micali, "Non-interactive oblivious transfer and spplications," in *CRYPTO'89*, ser. LNCS, G. Brassard, Ed., vol. 435. Santa Barbara, CA, USA: Springer, Berlin, Germany, Aug. 20–24, 1990, pp. 547–557.

[6] Y. T. Kalai, "Smooth projective hashing and two-message oblivious transfer," in *EUROCRYPT 2005*, ser. LNCS, R. Cramer, Ed., vol. 3494. Aarhus, Denmark: Springer, Berlin, Germany, May 22–26, 2005, pp. 78–95.

[7] C. Peikert, V. Vaikuntanathan, and B. Waters, "A framework for efficient and composable oblivious transfer," in *CRYPTO 2008*, ser. LNCS, D. Wagner, Ed., vol. 5157. Santa Barbara, CA, USA: Springer, Berlin, Germany, Aug. 17–21, 2008, pp. 554–571.

[8] R. Dowsley, J. van de Graaf, J. Müller-Quade, and A. C. A. Nascimento, "Oblivious transfer based on the mceliece assumptions," in *ICITS 2008*, ser. LNCS, R. Safavi-Naini, Ed., vol. 5155. Calgary, Canada: Springer, Berlin, Germany, Aug. 10–13, 2008, pp. 107–117.

[9] C. Crépeau and J. Kilian, "Achieving oblivious transfer using weakened security assumptions," in *FOCS 1988*, 1988, pp. 42–52.

[10] U. M. Maurer, "Conditionally-perfect secrecy and a provably-secure randomized cipher," *Journal of Cryptology*, vol. 5, no. 1, pp. 53–66, 1992.

[11] C. Cachin and U. M. Maurer, "Unconditional security against memory-bounded adversaries," in *CRYPTO'97*, ser. LNCS, B. S. Kaliski Jr., Ed., vol. 1294. Santa Barbara, CA, USA: Springer, Berlin, Germany, Aug. 17–21, 1997, pp. 292–306.

[12] C. Cachin, C. Crépeau, and J. Marcil, "Oblivious transfer with a memory-bounded receiver," in *39th FOCS*. Palo Alto, California, USA: IEEE Computer Society Press, Nov. 8–11, 1998, pp. 493–502.

[13] Y. Z. Ding, "Oblivious transfer in the bounded storage model," in *CRYPTO 2001*, ser. LNCS, J. Kilian, Ed., vol. 2139. Santa Barbara, CA, USA: Springer, Berlin, Germany, Aug. 19–23, 2001, pp. 155–170.

[14] D. Hong, K.-Y. Chang, and H. Ryu, "Efficient oblivious transfer in the bounded-storage model," in *ASIACRYPT 2002*, ser. LNCS, Y. Zheng, Ed., vol. 2501. Queenstown, New Zealand: Springer, Berlin, Germany, Dec. 1–5, 2002, pp. 143–159.

[15] Y. Z. Ding, D. Harnik, A. Rosen, and R. Shaltiel, "Constant-round oblivious transfer in the bounded storage model," in *TCC 2004*, ser. LNCS, M. Naor, Ed., vol. 2951. Cambridge, MA, USA: Springer, Berlin, Germany, Feb. 19–21, 2004, pp. 446–472.

[16] Y. Z. Ding, "Error correction in the bounded storage model," in *TCC 2005*, ser. LNCS, J. Kilian, Ed., vol. 3378. Cambridge, MA, USA: Springer, Berlin, Germany, Feb. 10–12, 2005, pp. 578–599.

[17] R. Renner and S. Wolf, "Simple and tight bounds for information reconciliation and privacy amplification," in *ASIACRYPT 2005*, ser. LNCS, B. K. Roy, Ed., vol. 3788. Chennai, India: Springer, Berlin, Germany, Dec. 4–8, 2005, pp. 199–216.

[18] S. P. Vadhan, "Constructing locally computable extractors and cryptosystems in the bounded-storage model," *Journal of Cryptology*, vol. 17, no. 1, pp. 43–77, Jan. 2004.

[19] L. Babai and T. P. Hayes, "Near-independence of permutations and an almost sure polynomial bound on the diameter of the symmetric group," ser. SODA '05. Philadelphia, PA, USA: Society for Industrial and Applied Mathematics, 2005, pp. 1057–1066.

[20] Y. Dodis, L. Reyzin, and A. Smith, "Fuzzy extractors: How to generate strong keys from biometrics and other noisy data," in *EUROCRYPT 2004*, ser. LNCS, C. Cachin and J. Camenisch, Eds., vol. 3027. Interlaken, Switzerland: Springer, Berlin, Germany, May 2–6, 2004, pp. 523–540.

[21] R. Ostrovsky, R. Venkatesan, and M. Yung, "Fair games against an all-powerful adversary," in *Sequences II*, R. Capocelli, A. Santis, and U. Vaccaro, Eds. Springer New York, 1993, pp. 418–429.

[22] C. Crépeau and G. Savvides, "Optimal reductions between oblivious transfers using interactive hashing," in *EUROCRYPT 2006*, ser. LNCS, S. Vaudenay, Ed., vol. 4004. St. Petersburg, Russia: Springer, Berlin, Germany, May 28 – Jun. 1, 2006, pp. 201–221.

[23] A. C. B. Pinto, R. Dowsley, K. Morozov, and A. C. A. Nascimento, "Achieving oblivious transfer capacity of generalized erasure channels in the malicious model," *Information Theory, IEEE Transactions on*, vol. 57, no. 8, pp. 5566–5571, 2011.

[24] M. Naor, R. Ostrovsky, R. Venkatesan, and M. Yung, "Perfect zero-knowledge arguments for NP using any one-way permutation," *Journal of Cryptology*, vol. 11, no. 2, pp. 87–108, 1998.