

# Commitment Capacity of Discrete Memoryless Channels

Andreas Winter\*    Anderson C. A. Nascimento†    Hideki Imai†

31st March 2003

## Abstract

In extension of the bit commitment task and following work initiated by Crépeau and Kilian, we introduce and solve the problem of characterising the optimal rate at which a discrete memoryless channel can be used for bit commitment. It turns out that the answer is very intuitive: it is the maximum equivocation of the channel (after removing trivial redundancy), even when unlimited noiseless bidirectional side communication is allowed. By a well-known reduction, this result provides a lower bound on the channel's capacity for implementing coin tossing, which we conjecture to be an equality.

The method of proving this relates the problem to Wyner's wire-tap channel in an amusing way. We also discuss extensions to quantum channels.

## 1 Introduction

Chess masters Alice and Bob are playing for the world chess championship and, after playing for several hours, realize that they will have to stop the game and resume it on the next morning. However, a problem arises: if Alice plays her turn before stopping the game, Bob will have the entire night to think of his next move, giving him an unfair advantage. If Alice does not play, she will have the entire night to think of her move. How can they get out of this problem?

If there is a trusted referee, Alice can write down her move and put it into an envelope and give it to the referee, who will announce it to Bob in the next morning. As the referee is trusted, Alice will be unable to change her move after writing it down, also Bob will be unable to learn Alice's move before the next morning. Can Alice and Bob solve this problem without the help of a trusted referee?

---

\*Department of Computer Science, University of Bristol, Merchant Venturers Building, Woodland Road, Bristol BS8 1UB, United Kingdom. Email: [winter@cs.bris.ac.uk](mailto:winter@cs.bris.ac.uk)

†Imai Laboratory, Information and Systems, Institute of Industrial Science, University of Tokyo, 4-6-1 Komaba, Meguro-ku, Tokyo 153-8505, Japan. Email: [anderson@imailab.iis.u-tokyo.ac.jp](mailto:anderson@imailab.iis.u-tokyo.ac.jp), [imai@iis.u-tokyo.ac.jp](mailto:imai@iis.u-tokyo.ac.jp)

To solve these kind of problems without the use of an active trusted party, Blum introduced commitment schemes in [7]. In a commitment scheme, Alice commits to an information by sending some piece of information to Bob during a commit phase. Later on, she can unveil the information she committed to by sending some opening information to Bob during an unveiling (also called reveal) phase. The protocol is said to be concealing if the information sent by Alice during the commit phase does not help Bob to learn a non-negligible amount of information on the value Alice is committing to. It is said to be binding if Alice is unable to commit to a certain information (which is usually a string of bits) and later on unveil a different one.

Without any kind of computational assumptions and assuming noiseless communications, commitment schemes are impossible (see e.g. [15]; the generalisation to quantum protocols is due to Mayers [20]). Therefore, research has mostly focused on schemes where the receiver is computationally bounded (computationally concealing schemes) or schemes where the sender is computationally bounded (computationally binding schemes). Examples of computationally binding but unconditionally concealing schemes are [7], [8], [16] and [17]. Examples of computationally concealing but unconditionally binding schemes are [21] and [23].

It is now known that noise is a powerful resource for the implementation of cryptographic primitives: it allows for the construction of *information theoretically secure* cryptographic protocols — a task typically impossible without the noise, and in practice done by relaxing to *computational security*, assuming conjectures from complexity theory.

In his famous paper [26], Wyner was the first to exploit noise in order to establish a secure channel in the presence of an eavesdropper. These results were extended in studies of secret key distillation by Maurer [19], Ahlswede and Csiszár [1] and followers. The noise in these studies is assumed to affect the eavesdropper: thus, to work in practice, it has to be guaranteed or certified somehow. This might be due to some — trusted — third party who controls the channel (and thus prevents the cryptographic parties from cheating), or due to physical limitations, as in quantum key distribution [4, 5]. Recently, Crépeau and Kilian [13] showed how information theoretically secure bit commitment can be implemented using a binary symmetric channel, their results being improved in [12] and [15].

The object of the present study is to optimise the use of the noisy channel, much as in Shannon’s theory of channel capacities: while the previous studies have concentrated on the *possibility* of bit commitment using noisy channels, here we look at committing to one out of a larger message set, e.g. a bit string. We are able, for a general discrete memoryless channel, to characterise the commitment capacity by a simple (single-letter) formula (theorem 2), stated in section 2, and proved in two parts in sections 3 and 4. A few specific examples are discussed in section 5 to illustrate the main result. In section 6 results on an extension to quantum channels are related, and we close with a discussion (section 7). An appendix collects some facts about typical sequences used in the main proof.

## 2 Definitions and main result

In the commitment of a message there are two parties, called *Alice* and *Bob*, the first one given the message  $a$  from a certain set  $\mathcal{A}$ . The whole procedure consists of two stages: first the *commit phase*, in which Alice (based on  $a$ ) and Bob exchange messages, according to a protocol. This will leave Bob with a record (usually called *view*), to be used in the second stage, the *reveal phase*. This consists of Alice disclosing  $a$  and other relevant information to Bob. Bob performs a test on all his recorded data which accepts if Alice followed the rules and disclosed the correct information in the second stage, and rejects if a violation of the rules is discovered.

To be useful, such a scheme has to fulfill two requirements: it must be “concealing” as well as “sound” and “binding”: the first property means that after the commit phase Bob has no or almost no information about  $a$  (i.e., even though Alice has “committed” herself to something by the communications to Bob, this commitment remains secret), and this has to hold even if Bob does not follow the protocol, while Alice does. Soundness means that if both parties behave according to the protocol, Bob’s test will accept (with high probability) after the reveal phase. The protocol to be binding means that Bob’s test is such that whatever Alice did in the commit phase (with Bob following the rules) there is only at most one  $a$  she can “reveal” which passes Bob’s test.

In our present consideration there is an unlimited bidirectional noiseless channel available between Alice and Bob, and in addition a discrete memoryless noisy channel  $W : \mathcal{X} \rightarrow \mathcal{Z}$  from Alice to Bob, which may be used  $n$  times: on input  $x^n = x_1 \dots x_n$ , the output distribution on  $\mathcal{Z}^n$  is  $W_{x^n} = W_{x_1} \otimes \dots \otimes W_{x_n}$ .

**Definition 1** *The channel  $W$  is called non-redundant, if none of its output distributions is a convex combination of its other output distributions:*

$$\forall y \forall P \text{ s.t. } P(y) = 0 \quad W_y \neq \sum_x P(x)W_x.$$

*In geometric terms this means that all distributions  $W_x$  are distinct extremal points of the polytope  $\mathcal{W} = \text{conv} \{W_x : x \in \mathcal{X}\}$ , the convex hull of the output distributions within the probability simplex over  $\mathcal{Z}$ . Clearly, we can make  $W$  into a non-redundant channel  $\tilde{W}$  by removing all input symbols  $x$  whose output distribution  $W_x$  is not extremal. The old channel can be simulated by the new one, because by feeding it distributions over input symbols one can generate the output distributions of the removed symbols.*

*The channel  $W$  is called trivial, if after making it non-redundant its output distributions have mutually disjoint support. This means that from the output one can infer the input with certainty.*

With this we can pass to a formal definition of a protocol: this, consisting of the named two stages, involves creation on Alice’s side of either messages intended for the noiseless channel, or inputs to the noisy channel, based on previous messages received from Bob via the noiseless channel, which themselves

are based on data received before, etc. Both agents may employ probabilistic choices, which we model by Alice and Bob each using a random variable,  $M$  and  $N$ , respectively. This allows them to use *deterministic* functions in the protocol. Note that this makes all messages sent and received into well-defined random variables, *dependent on  $a$* .

*Commit Phase:* The protocol goes for  $r$  rounds of Alice-to-Bob and Bob-to-Alice noiseless communications  $U_j$  and  $V_j$ . After round  $r_i$  ( $r_1 \leq \dots \leq r_n \leq r$ ) Alice will also send a symbol  $X_i$  down the noisy channel  $W$ , which Bob receives as  $Z_i$ . Setting  $r_0 = 0$  and  $r_{n+1} = r$ :

Round  $r_i + k$  ( $1 \leq k \leq r_{i+1} - r_i$ ): Alice sends  $U^{r_i+k} = f_{r_i+k}(a, M, V^{r_i+k-1})$  noiselessly. Bob answers  $V_{r_i+k} = g_{r_i+k}(Z^i, N, U^{r_i+k})$ , also noiselessly. After round  $r_i$  and before round  $r_i + 1$  ( $1 \leq i \leq n$ ), Alice sends  $X_i = F_i(a, M, V^{r_i})$ , which Bob receives as  $Z_i = W(X_i)$ .

*Reveal Phase:* A similar procedure as the Commit Phase, but without the noisy channel uses, including Alice's sending  $a$  to Bob. At the end of the exchange Bob performs a test as to whether to accept Alice's behaviour or not. It is easily seen that this procedure can be simulated by Alice simply telling Bob  $a$  and  $M$ , after which Bob performs his test  $\beta(Z^n, N, U^r; a, M) \in \{\text{ACC}, \text{REJ}\}$ . I.e., requiring Alice to reveal  $M$  and  $a$  makes cheating for her only more difficult.

We shall, for technical reasons, impose the condition that the range of the variable  $U^r$  is bounded:

$$|U^r| \leq \exp(Bn), \quad (1)$$

with a constant  $B$ . Note that  $\exp$  and  $\log$  in this paper are always to basis 2, unless otherwise stated.

Now, the mathematical form of the conditions for concealing as well as for soundness and binding is this: we call the above protocol  $\epsilon$ -concealing if for any two messages  $a, a' \in \mathcal{A}$  and any behaviour of Bob during the commit phase,

$$\frac{1}{2} \left\| \text{Distr}_a(Z^n N U^r) - \text{Distr}_{a'}(Z^n N U^r) \right\|_1 \leq \epsilon, \quad (\text{A})$$

where  $\text{Distr}_a(Z^n N U^r)$  is the distribution of the random variables  $Z^n N U^r$  after completion of the commit phase which Alice entered with the message  $a$  and the randomness  $M$ , and with the  $\ell_1$ -norm  $\|\cdot\|_1$ ; the above expression is identical to the total variational distance of the distributions. This is certainly the strongest requirement one could wish for: it says that no statistical test of Bob immediately after the commit phase can distinguish between  $a$  and  $a'$  with probability larger than  $\epsilon$ . Note that  $V^r$  is a function of  $Z^n N U^r$ , and hence could be left out in eq. (A). Assuming any probability distribution on the messages,  $a$  is the value of a random variable  $A$ , and it is jointly distributed with all other variables of the protocol. Then, whatever Bob's strategy,

$$I(A \wedge Z^n N U^r) \leq \epsilon' = H(2\epsilon, 1 - 2\epsilon) + 2n\epsilon(\log B + \log |\mathcal{Z}|), \quad (\text{A}')$$

where

$$I(X \wedge Y) = H(X) + H(Y) - H(XY)$$

is the (Shannon) mutual information between  $X$  and  $Y$ , and

$$H(X) = - \sum_x \Pr\{X = x\} \log \Pr\{X = x\}$$

is the (Shannon) entropy of  $X$  [25].

We call the protocol  $\delta$ -*sound and -binding* ( $\delta$ -*binding* for short), if for Alice and Bob following the protocol, for all  $a \in \mathcal{A}$ ,

$$\Pr\{\beta(Z^n NU^r; aM) = \text{ACC}\} \geq 1 - \delta, \quad (\text{B1})$$

and, whatever Alice does during the commit phase, governed by a random variable  $S$  with values  $\sigma$  (which determines the distribution of  $Z^n NU^r$ ), for all  $A = a(S, V^r)$ ,  $A' = a'(S, V^r)$ ,  $\widetilde{M} = \mu(S, V^r)$  and  $\widetilde{M}' = \mu'(S, V^r)$  such that  $A \neq A'$  with probability 1,

$$\Pr\left\{\beta(Z^n NU^r; \widetilde{AM}) = \text{ACC} \ \& \ \beta(Z^n NU^r; A'\widetilde{M}') = \text{ACC}\right\} \leq \delta. \quad (\text{B2})$$

Note that by convexity the cheating attempt of Alice is w.l.o.g. *deterministic*, which is to say that  $S$  takes on only one value  $\sigma$  with non-zero probability, hence  $\Pr\{S = \sigma\} = 1$ .

We call  $\frac{1}{n} \log |\mathcal{A}|$  the (*commitment*) *rate* of the protocol. A rate  $R$  is said to be *achievable* if there exist commitment protocols for every  $n$  with rates converging to  $R$ , which are  $\epsilon$ -concealing and  $\delta$ -binding with  $\epsilon, \delta \rightarrow 0$  as  $n \rightarrow \infty$ . The *commitment capacity*  $C_{\text{com}}(W)$  of  $W$  is the supremum of all achievable rates.

The main result of this paper is the following theorem:

**Theorem 2** *The commitment capacity of the discrete channel  $W$  (assumed to be non-redundant) is*

$$C_{\text{com}}(W) = \max\{H(X|Z) : X, Z \text{ RVs, Distr}(Z|X) = W\},$$

*i.e., the maximal equivocation of the channel over all possible input distributions.*

**Corollary 3** *Every non-trivial discrete memoryless channel can be used to perform bit commitment.*  $\square$

By invoking the well-known reduction of coin tossing to bit commitment [7] we obtain:

**Corollary 4** *The channel  $W$  can be used for secure two-party coin tossing at rate at least  $C_{\text{com}}(W)$ . I.e., for the naturally defined coin tossing capacity  $C_{\text{c.t.}}(W)$ , one has  $C_{\text{c.t.}}(W) \geq C_{\text{com}}(W)$ .*  $\square$

This theorem will be proved in the following two sections (propositions 8 and 9): first we construct a protocol achieving the equivocation bound, showing

that exponential decrease of  $\epsilon$  and  $\delta$  is possible, and without using the noiseless side channels at all during the commit phase. Then we show the optimality of the bound.

To justify our allowing small errors both in the concealing and the binding property of a protocol, we close this section by showing that demanding too much trivialises the problem:

**Theorem 5** *There is no bit-commitment via  $W$  which is  $\epsilon$ -concealing and 0-binding with  $\epsilon < 1$ . I.e., not even two distinct messages can be committed:  $|\mathcal{A}| = 1$ .*

*Proof.* If the protocol is 0-sound, this means that for every value  $\mu$  attained by  $M$  with positive probability, Bob will accept the reveal phase if Alice behaved according to the protocol. On the other hand, that the protocol is 0-binding means that for  $a \neq a'$  and arbitrary  $\mu'$ , Bob will never accept if Alice behaves according to the protocol in the commit phase, with values  $a\mu$  but tries to “reveal”  $a'\mu'$ . This opens the possibility of a decoding method for  $a$  based on  $Z^n NU^r$ : Bob simply tries out all possible  $a\mu$  with his test  $\beta$  — the single  $a$  which is accepted must be the one used by Alice. Hence the scheme cannot be  $\epsilon$ -concealing with  $\epsilon < 1$ .  $\square$

**Remark 6** *By contrast, it is easy to construct schemes both 0-concealing and  $\delta$ -binding with  $\delta < 1$ , for an appropriately defined channel:*

*Consider the channel  $F$  with input and output alphabets  $\mathcal{X} = \mathcal{Z} = \{0, 1, 2, 3\}$  and signals defined by*

$$F_x(z) = \begin{cases} \frac{1}{2} & \text{if } z - x \equiv 0 \text{ or } 1 \pmod{4}, \\ 0 & \text{otherwise.} \end{cases}$$

*Alice commits to the bit  $b$  by picking  $c \in \{0, 1\}$  at random and sending  $x = b + 2c$  (for which Bob receives a random  $z$  such that  $z - x \equiv 0 \text{ or } 1 \pmod{4}$ ). To reveal, she tells him  $x$  (which decodes to a unique  $b$ ) and he accepts iff  $z - x \equiv 0 \text{ or } 1 \pmod{4}$ .*

*Clearly, this scheme is 0-concealing because for both  $b = 0$  and  $b = 1$  Bob sees the uniform distribution on  $\mathcal{Z}$ . Equally obviously, it is 0-sound, but it is also  $\frac{1}{2}$ -binding: for if Alice wants to “reveal”  $b' \neq b$  (with corresponding  $x' \neq x$ ), she has only a probability of  $\frac{1}{2}$  to pick  $x'$  with  $z - x' \equiv 0 \text{ or } 1 \pmod{4}$ .*

This simple scheme is in fact the template for the coding scheme of proposition 8: put differently, on sufficiently large scale (block length) every channel looks like  $F$ .

### 3 A scheme meeting the equivocation bound

Here we describe and prove security bounds of a scheme which is very simple compared to the generality we allowed in section 4: in the commit phase it

consists only of a single block use of the noisy channel  $W^n$ , with no public discussion at all, where the input  $X^n$  is a random one-to-many function of  $a$  (in particular,  $a$  is a function of the  $x^n$  chosen). In the reveal phase Alice simply announces  $X^n$  to Bob.

**Proposition 7** *Given  $\sigma, \tau > 0$ , and a distribution  $P$  of  $X \in \mathcal{X}$ , with the output  $Z = W(X)$ ,  $Q = \text{Distr}(Z)$ . Then there exists a collection of codewords*

$$(\xi_{a\mu} \in \mathcal{X}^n : a = 1, \dots, K, \mu = 1, \dots, L)$$

with the following properties:

1. For all  $(a, \mu) \neq (a', \mu')$ ,  $d_{\text{H}}(\xi_{a\mu}, \xi_{a'\mu'}) \geq 2\sigma n$ .

2. For every  $a$ :

$$\frac{1}{2} \left\| \frac{1}{L} \sum_{\mu=1}^L W_{\xi_{a\mu}}^n - Q^{\otimes n} \right\|_1 \leq 25|\mathcal{X}||\mathcal{Z}| \exp(-n\tau).$$

3. There are constants  $G, G'$  and a continuous function  $G''$  vanishing at 0 such that

$$K \geq \frac{1}{2n} (3 + \log |\mathcal{X}| + \log |\mathcal{Z}|)^{-1} \exp(nH(X|Z) - n\sqrt{2\tau}G' - nG''(\sigma)),$$

$$L \leq n(3 + \log |\mathcal{X}| + \log |\mathcal{Z}|) \exp(nI(X \wedge Z) + n\sqrt{2\tau}G).$$

*Proof.* To get the idea, imagine a wiretap channel [26] with  $W$  as the stochastic matrix of the eavesdropper and a symmetric channel  $S_\sigma : \mathcal{X} \rightarrow \mathcal{Y} = \mathcal{X}$  for the legal user:

$$S_\sigma(y|x) = \begin{cases} 1 - \sigma & \text{if } x = y, \\ \frac{1}{|\mathcal{X}|} \sigma & \text{if } x \neq y. \end{cases}$$

The random coding strategy for such a channel, according to Wyner's solution [26] (but see also [14] and [10]) will produce a code with the properties 2 and 3. Because the code for the legal user must fight the noise of the symmetric channel  $S_\sigma$ , we can expect its codewords to be of large mutual Hamming distance, i.e., we should get property 1.

In detail: pick the  $\xi_{a\mu}$  i.i.d. according to the distribution  $\tilde{P}^n$ , which is 0 outside  $\mathcal{T}_{P, \sqrt{2\tau}}^n$  (the typical sequences, see appendix A) and  $P^{\otimes n}$  within, suitably normalised. Also introduce the subnormalised measures  $\widehat{W}_{x^n}^n$ : this is identical to  $W_{x^n}^n$  within  $\mathcal{T}_{W, \sqrt{2\tau}}^n(x^n)$  and 0 outside. We will show that with high probability we can select codewords with properties 2 and 3, and only a small proportion of which violate property 1; then by an expurgation argument will we obtain the desired code.

By eqs. (9) and (14) in the appendix we have

$$\frac{1}{2} \left\| E \widehat{W}_{\xi_{a\mu}}^n - Q^{\otimes n} \right\| \leq 3|\mathcal{X}||\mathcal{Z}| \exp(-n\tau), \quad (2)$$

with the expectation referring to the distribution  $\tilde{P}^n$  of the  $\xi_{a\mu}$ . Observe that the support of all  $\widehat{W}_{\xi_{a\mu}}^n$  is contained in  $\mathcal{T}_{Q,2|\mathcal{X}|\sqrt{\tau}}^n$ , using eq. (18) of the appendix. Now,  $\mathcal{S}$  is defined as the set of those  $z^n$  for which

$$E\widehat{W}_{\xi_{a\mu}}^n(z^n) \geq T := \exp(-n\tau) \exp(-nH(Q) - n\sqrt{2\tau}|\mathcal{X}|F),$$

with  $F = \sum_{z:Q(z)\neq 0} -\log Q(z)$ , and define  $\widetilde{W}_{x^n}^n(z^n) = \widehat{W}_{x^n}^n(z^n)$  if  $z^n \in \mathcal{S}$  and 0 otherwise. With the cardinality estimate eq. (??) of the appendix and eq. (2) we obtain

$$\frac{1}{2} \left\| E\widetilde{W}_{\xi_{a\mu}}^n - Q^{\otimes n} \right\| \leq 4|\mathcal{X}||\mathcal{Z}| \exp(-n\tau). \quad (3)$$

The Chernoff bound allows us now to efficiently sample the expectation  $\tilde{Q}^n := E\widetilde{W}_{\xi_{a\mu}}^n$ : observe that all the values of  $\widetilde{W}_{\xi_{a\mu}}^n$  are upper bounded by

$$t := \exp(-nH(W|P) + n\sqrt{2\tau}|\mathcal{X}| \log |\mathcal{Z}| + n\sqrt{2\tau}E),$$

using eq. (15). Thus, rescaling the variables, by lemma 13 and with the union bound, we get

$$\Pr \left\{ \forall a \forall z^n \in \mathcal{S} \frac{1}{L'} \sum_{\mu=1}^{L'} \widetilde{W}_{\xi_{a\mu}}^n(z^n) \in [(1 \pm \exp(-n\tau))\tilde{Q}^n(z^n)] \right\} \leq 2K|\mathcal{S}| \exp\left(-L \frac{T \exp(-2n\tau)}{2 \ln 2}\right), \quad (4)$$

which is smaller than 1/2 if

$$L' > 2 + n(\log |\mathcal{X}| + \log |\mathcal{Z}|) \exp(nI(X \wedge Z) + n\sqrt{2\tau}G),$$

with  $G = 3 + |\mathcal{X}|F + |\mathcal{X}| \log |\mathcal{Z}| + E$ . Note that in this case, there exist values for the  $\xi_{a\mu}$  such that the averages  $\frac{1}{L'} \sum_{\mu} W_{\xi_{a\mu}}^n$  are close to  $Q^{\otimes n}$ .

Now we have to enforce property 1: in a random batch of  $\xi_{a\mu}$  we call  $a\mu$  *bad* if  $\xi_{a\mu}$  has Hamming distance less than  $2\sigma n$  from another  $\xi_{a'\mu'}$ . The probability that  $a\mu$  is bad is easily bounded:

$$\begin{aligned} \Pr\{a\mu \text{ bad}\} &\leq 2|\mathcal{X}| \exp(-n\tau) + P^{\otimes n} \left( \bigcup_{a'\mu' \neq a\mu} B_{2n\sigma}(\xi_{a'\mu'}) \right) \\ &\leq 2|\mathcal{X}| \exp(-n\tau) + \max \left\{ P^{\otimes n}(\mathcal{A}) : |\mathcal{A}| \leq K'L' \binom{n}{2n\sigma} |\mathcal{X}|^{2n\sigma} \right\} \\ &\leq 5|\mathcal{X}| \exp(-n\tau), \end{aligned}$$

by eq. (13) in the appendix, because we choose

$$\begin{aligned} K' &\leq \frac{1}{n} (3 + \log |\mathcal{X}| + \log |\mathcal{Z}|)^{-1} \\ &\quad \exp(nH(X|Z) - n\sqrt{2\tau}G - 2n\sqrt{2\tau}D - nH(2\sigma, 1 - 2\sigma) - 2n\sigma \log |\mathcal{X}|), \end{aligned}$$

hence

$$K'L' \binom{n}{2n\sigma} |\mathcal{X}|^{2n\sigma} < \exp(nH(P) - 2n\sqrt{2\tau}D).$$

Thus, with probability at least  $1/2$ , only a fraction of  $10|\mathcal{X}|\exp(-n\tau)$  of the  $a\mu$  are bad. Putting this together with eq. (4), we obtain a selection of  $\xi_{a\mu}$  such that

$$\forall a \quad \frac{1}{2} \left\| \frac{1}{L'} \sum_{\mu} W_{\xi_{a\mu}}^n - Q^{\otimes n} \right\|_1 \leq 5|\mathcal{X}||\mathcal{Z}|\exp(-n\tau) \quad (5)$$

and only a fraction of  $10|\mathcal{X}|\exp(-n\tau)$  of the  $a\mu$  are bad.

This means that for at least half of the  $a$ , w.l.o.g.  $a = 1, \dots, K = K'/2$ , only a fraction  $20|\mathcal{X}|\exp(-n\tau)$  of the  $\mu$  form bad pairs  $a\mu$ , w.l.o.g. for  $\mu = L+1, \dots, L'$ , with  $L = (1 - 20|\mathcal{X}|\exp(-n\tau))L'$ . Throwing out the remaining  $a$  and the bad  $\mu$ , we are left with a code as desired.  $\square$

Observe that a receiver of  $Z^n$  can efficiently check claims about the input  $\xi_{a\mu}$  because of property 1, that distinct codewords have “large” Hamming distance. The non-redundancy of  $W$  shuns one-sided errors in this checking, as we shall see. The test  $\beta$  is straightforward: it accepts iff  $Z^N \in \mathcal{T}_{W, \sqrt{2\tau}}^n(\xi_{a\mu})$ , the set of *conditional typical sequences*, see appendix A. This ensures soundness; for the bindingness we refer to the following proof.

We are now in a position to describe a protocol, having chosen codewords according to proposition 7:

*Commit phase:* To commit to a message  $a$ , Alice picks  $\mu \in \{1, \dots, L\}$  uniformly at random and sends  $\xi_{a\mu}$  through the channel. Bob obtains a channel output  $z^n$ .

*Reveal phase:* Alice announces  $a$  and  $\mu$ . Bob performs the test  $\beta$ : he accepts if  $z^n \in \mathcal{B}_{a\mu} := \mathcal{T}_{W, \sqrt{2\tau}}^n(\xi_{a\mu})$  and rejects otherwise.

**Proposition 8** *Assume that for all  $x \in \mathcal{X}$  and distributions  $P$  with  $P(x) = 0$ ,*

$$\left\| W_x - \sum_y P(y)W_y \right\|_1 \geq \eta.$$

*Let  $\tau = \frac{\sigma^4 \eta^2}{8|\mathcal{X}|^4 |\mathcal{Z}|^2}$ : then the above protocol implements an  $\epsilon$ -concealing and  $\delta$ -binding commitment with rate*

$$\frac{1}{n} \log K \geq H(X|Z) - \sqrt{2\tau}G' - H(2\sigma, 1 - 2\sigma) - 2\sigma \log |\mathcal{X}| - \frac{\log n}{n} - O\left(\frac{1}{n}\right)$$

*and exponentially bounded security parameters:*

$$\begin{aligned} \epsilon &= 50|\mathcal{X}||\mathcal{Z}|\exp(-n\tau), \\ \delta &= 2|\mathcal{X}||\mathcal{Z}|\exp(-2n\tau^2). \end{aligned}$$

*Proof.* That the protocol is  $\epsilon$ -concealing is obvious from property 2 of the code in proposition 7: Bob's distribution of  $Z^n$  is always  $\epsilon/2$ -close to  $Q^{\otimes n}$ , whatever  $a$  is.

To show  $\delta$ -bindingness observe first that if Alice is honest, sending  $\xi_{a\mu}$  in the commit phase and later revealing  $a\mu$ , the test  $\beta$  will accept with high probability:

$$\begin{aligned} \Pr\{Z^n \in \mathcal{B}_{a\mu}\} &= W_{\xi_{a\mu}}^n(\mathcal{T}_{W, \sqrt{2\tau}}^n(\xi_{a\mu})) \\ &\geq 1 - 2|\mathcal{X}||\mathcal{Z}|\exp(-n\tau) \geq 1 - \delta, \end{aligned}$$

by eq. (14) in the appendix.

On the other hand, if Alice cheats, we may — in accordance with our definition — assume her using a deterministic strategy: i.e., she “commits” sending some  $x^n$  and later attempts to “reveal” either  $a\mu$  or  $a'\mu'$ , with  $a \neq a'$ . Because of property 1 of the code in proposition 7, at least one of the codewords  $\xi_{a\mu}$ ,  $\xi_{a'\mu'}$  is at Hamming distance at least  $\sigma n$  from  $x^n$ : w.l.o.g., the former of the two. But then the test  $\beta$  accepts “revelation” of  $a\mu$  with small probability:

$$\Pr\{Z^n \in \mathcal{B}_{a\mu}\} = W_{x^n}^n(\mathcal{T}_{W, \sqrt{2\tau}}^n(\xi_{a\mu})) \leq 2\exp(-2n\tau^2) \leq \delta,$$

by lemma 14 in the appendix.  $\square$

## 4 Upper bounding the achievable rate

We assume that  $W$  is non-redundant. We shall prove the following assertion, assuming a uniformly distributed variable  $A \in \mathcal{A}$  of messages.

**Proposition 9** *Consider an  $\epsilon$ -concealing and  $\delta$ -binding commitment protocol with  $n$  uses of  $W$ . Then*

$$\begin{aligned} \log |\mathcal{A}| \leq n \max\{H(X|Z) : \text{Distr}(Z|X) = W\} \\ + n(\epsilon(\log B + \log |\mathcal{Z}|) + 5\sqrt[3]{\delta} \log |\mathcal{X}|) + 2. \end{aligned} \quad (6)$$

The key, as it turns out, of its proof, is the insight that in the above protocol, should it be concealing and binding,  $x^n$  together with Bob's view of the commit phase (essentially) determine  $a$ . In the more general formulation we permitted in section 2, we prove :

$$H(A|Z^n NU^r; X^n) \leq \delta' = H\left(5\sqrt[3]{\delta}, 1 - 5\sqrt[3]{\delta}\right) + 5\sqrt[3]{\delta} \log |\mathcal{A}|. \quad (B')$$

Intuitively, this means that with the items Alice entered into the commit phase of the protocol and those which are accessible to Bob, not too many values of  $A$  should be consistent — otherwise Alice had a way to cheat.

*Proof of eq. (B').* For each  $a\mu$  the commit protocol (both players being honest) creates a distribution  $\Delta_{a\mu}$  over conversations  $(x^n v^r; z^n u^r)$ . We leave out Bob's random variable  $N$  here, noting that he can create its correct conditional distribution from  $z^n u^r; v^r$ , which is his view of the conversation. The only other

place where he needs it is to perform the test  $\beta$ . We shall in the following assume that it includes this creation of  $N$ , which makes  $\beta$  into a probabilistic test, depending on  $(a\mu v^r; z^n u^r)$ .

The pair  $a\mu$  has a probability  $\alpha_{a\mu}$  that its conversation with subsequent revelation of  $a\mu$  is accepted. By soundness, we have

$$\sum_{\mu} \Pr\{M = \mu\} \alpha_{a\mu} \geq 1 - \delta,$$

for every  $a$ . Hence, by Markov inequality, there exists (for every  $a$ ) a set of  $\mu$  of total probability  $\geq 1 - \sqrt[3]{\delta^2}$  for which  $\alpha_{a\mu} \geq 1 - \sqrt[3]{\delta}$ . We call such  $\mu$  *good for a*.

From this we get a set  $\mathcal{C}_{a\mu}$  of “partial” conversations  $(x^n v^r; u^r)$ , with probability  $\Delta_{a\mu}(\mathcal{C}_{a\mu}) \geq 1 - \sqrt[3]{\delta}$ , which are accepted with probability at least  $1 - \sqrt[3]{\delta}$ . (In the test also  $Z^n$  enters, which is distributed according to  $W_{x^n}^n$ .)

Let us now define the set

$$\mathcal{C}_a := \bigcup_{\mu \text{ good for } a} \mathcal{C}_{a\mu},$$

which is a set of “partial conversations” which are accepted with probability at least  $1 - \sqrt[3]{\delta}$  and

$$\Delta_a(\mathcal{C}_a) \geq 1 - 2\sqrt[3]{\delta},$$

with the distribution

$$\Delta_a := \sum_{\mu} \Pr\{M = \mu\} \Delta_{a\mu}$$

over “partial conversations”: it is the distribution created by the commit phase give the message  $a$ .

We claim that

$$\Delta_a \left( X^n V^r; U^r \in \bigcup_{a' \neq a} \mathcal{C}_{a'} \right) \leq 3\sqrt[3]{\delta}. \quad (7)$$

Indeed, if this were not the case, Alice had the following cheating strategy: in the commit phase she follows the protocol for input message  $a$ . In the reveal phase she looks at the “partial conversation”  $x^n v^r; u^r$  and tries to “reveal” some  $a'\mu'$  for which the partial conversation is in  $\mathcal{C}_{a'\mu'}$  (if these do not exist,  $a'\mu'$  is arbitrary). This defines random variables  $A'$  and  $\widetilde{M}'$  for which it is easily checked that

$$\Pr\{\beta(Z^n N U^r; aM) = \text{ACC} \ \& \ \beta(Z^n N U^r; A' \widetilde{M}') = \text{ACC}\} > \delta,$$

contradicting the  $\delta$ -bindingness condition.

Using eq. (7) we can build a decoder for  $A$  from  $X^n V^r; U^r$ : choose  $\widehat{A} = a$  such that  $X^n V^r; U^r \in \mathcal{C}_a$  — if there exists none or more than one, let  $\widehat{A}$  be arbitrary. Clearly,

$$\Pr\{A \neq \widehat{A}\} \leq 5\sqrt[3]{\delta},$$

and invoking Fano's inequality we are done.  $\square$

Armed with this, we can now proceed to the *Proof of proposition 9*. We can successively estimate,

$$\begin{aligned}
H(X^n|Z^n) &\geq H(X^n|Z^nNU^r) \\
&= H(AX^n|Z^nNU^r) - H(A|Z^nNU^r; X^n) \\
&\geq H(A|Z^nNU^r) - H(A|Z^nNU^r; X^n) \\
&\geq H(A|Z^nNU^r) - \delta' \\
&= H(A) - I(A \wedge Z^nNU^r) - \delta' \\
&\geq H(A) - \epsilon' - \delta',
\end{aligned}$$

using eq. (B') in the fourth, eq. (A') in the sixth line. On the other hand, subadditivity and the conditioning inequality imply

$$H(X^n|Z^n) \leq \sum_{k=1}^n H(X_k|Z_k),$$

yielding the claim, because  $H(A) = \log |\mathcal{A}|$ .

The application to the proof of the converse of theorem 2 is by observing that  $\epsilon', \delta' = o(n)$ .  $\square$

Note that for the proof of the proposition we considered only a very weak attempt of Alice to cheat: she behaves according to the protocol during the commit phase, and only at the reveal stage she tries to be inconsistent. Similarly, our concealingness condition considered only passive attempts to cheat by Bob, i.e., he follows exactly the protocol, and tries to extract information about  $A$  only by looking at his view of the exchange.

Thus, even in the model of *passive cheating*, which is less restrictive than our definition in section 2, we obtain the upper bound of proposition 9

## 5 Examples

In this section we discuss some particular channels, which we present as stochastic matrices with the rows containing the output distributions.

**1. Binary symmetric channel  $B_p$ :** Let  $0 \leq p \leq 1$ . Define

$$B_p := \begin{array}{c|cc} \cdot & 0 & 1 \\ \hline 0 & 1-p & p \\ \hline 1 & p & 1-p \end{array}$$

The transmission capacity of this channel is easily computed from Shannon's formula [25]:  $C(B_p) = 1 - H(p, 1-p)$ , which is non-zero iff  $p \neq 1/2$ . The optimal input distribution is the uniform distribution  $(1/2, 1/2)$  on  $\{0, 1\}$ . Note that this channel is trivial if  $p \in \{0, 1/2, 1\}$ , hence  $C_{\text{com}}(B_p) = 0$  for these values of  $p$ . We may thus, w.l.o.g., assume that  $0 < p < 1/2$ , for which  $B_p$  is

non-redundant. It is not hard to compute the optimal input distribution as the uniform distribution, establishing  $C_{\text{com}}(B_p) = H(p, 1 - p)$ .

The result is in accordance with our intuition: the noisier the channel is, the worse it is for transmission, but the better for commitment.

**2. A trivial channel:** Consider the channel

$$T := \begin{array}{c|cc} \cdot & 0 & 1 \\ \hline a & 1/2 & 1/2 \\ \hline b & 1 & 0 \\ \hline c & 0 & 1 \end{array}$$

Clearly,  $T$  is trivial, hence  $C_{\text{com}}(T) = 0$ . Still it is an interesting example in the light of our proof of proposition 8: for assume a wiretap channel for which  $T$  is the stochastic matrix of the eavesdropper, while the legal user obtains a noiseless copy of the input. Then clearly the wiretap capacity of this system is 1, with optimal input distribution  $(1/2, 1/4, 1/4)$ .

**3. Transmission and commitment need not be opposites:** We show here an example of a channel where the optimising input distributions for transmission and for commitment are very different:

$$V := \begin{array}{c|cc} \cdot & 0 & 1 \\ \hline 0 & 1/2 & 1/2 \\ \hline 1 & 1 & 0 \end{array}$$

It can be easily checked that the maximum of the mutual information, i.e. the transmission capacity, is attained for the input distribution

$$P(0) = \frac{2}{5} = 0.4, \quad P(1) = \frac{3}{5} = 0.6,$$

from which we obtain  $C(V) \approx 0.3219$ . On the other hand, the equivocation is maximised for the input distribution

$$P'(0) = 1 - \sqrt{\frac{1}{5}} \approx 0.5528, \quad P'(1) = \sqrt{\frac{1}{5}} \approx 0.4472,$$

from which we get that  $C_{\text{com}}(V) \approx 0.6942$ . The maximising distributions are so different that the sum  $C(V) + C_{\text{com}}(V) > 1$ , i.e. it exceeds the maximum input and output entropies of the channel.

## 6 Quantum channels

The construction of section 3 can be carried over to a class of quantum channels, namely so-called *cq-channels* (classical-quantum channels):

$$W : \mathcal{X} \longrightarrow \mathcal{S}(\mathcal{H}),$$

a map from an input alphabet (here assumed to be finite) into the set of states on a Hilbert space  $\mathcal{H}$ , also assumed to be finite-dimensional in the present discussion. (For an overview of quantum information theory see [6] and the

textbook [22].) *Non-redundancy* means the same here, only that the convex structure is now the convex compact set of states, instead of the probability simplex. We assume this property of  $W$  silently in the following.

**Theorem 10** *For a distribution  $P$  on the input alphabet, one can achieve the commitment rate*

$$H(P) - \chi(\{P(x); W_x\}), \quad (8)$$

*with the Holevo mutual information [18]*

$$\chi(\{P(x); W_x\}) = S\left(\sum_x P(x)W_x\right) - \sum_x P(x)S(W_x),$$

where  $S(\rho) = -\text{Tr}\rho \log \rho$  is the von Neumann entropy of a state.

*The maximum of the expression (8) is optimal in the case of no noiseless side communication during the commit phase.*

*Proof (Sketch).* For the achievability one proves a coding result similar to proposition 7, with the  $\|\cdot\|_1$ -norm denoting trace norm. The most crucial point is property 2: our proof used two things: restricting the distributions  $W_{x^n}$  to typical sequences — this can be done also for states by constructing typical subspaces — and Chernoff bound to obtain a “small sample”. We use an analogue of this for operators from [2], stated below as lemma 11. This technique is actually used in the work of Cai and Yeung [10] to construct codes for the quantum wiretap channel

For the optimality, it is not hard to prove the quantum analogues of eqs. (A’) and (B’), and then the upper bound follows exactly as in our proof of proposition 9.  $\square$

**Lemma 11 (Ahlsvede, Winter [2])** *Let  $X_1, \dots, X_L$  be i.i.d. random variables taking values in the operators  $\mathcal{B}(\mathcal{H})$  on the  $D$ -dimensional Hilbert space  $\mathcal{H}$ ,  $0 \leq X_\ell \leq \mathbb{1}$ , with  $A = EX_\ell \geq \alpha \mathbb{1}$ , and let  $\eta > 0$ . Then*

$$\Pr \left\{ \frac{1}{L} \sum_{\ell=1}^L X_\ell \notin [(1-\eta)A; (1+\eta)A] \right\} \leq 2D \exp\left(-L \frac{\alpha \eta^2}{2 \ln 2}\right),$$

where  $[A; B] = \{X : A \leq X \leq B\}$  is an interval in the operator order.  $\square$

**Example 12** Assume any set of distinct pure qubit states  $W_x = |\psi_x\rangle\langle\psi_x|$ . Then, with  $\rho = \sum_x P(x)W_x$ , the rate

$$\max_P \{H(P) - S(\rho)\}$$

is achievable. Because of  $S(\rho) \leq 1$  this is positive if the input alphabet has at least three symbols; in the case of two input symbols it is positive iff the two states are non-orthogonal.

This is no contradiction to Mayer’s no-go theorem for quantum bit commitment even though the channel might appear to be noiseless: it is, however, not a noiseless qubit channel, because the states are restricted to a set of pure states. Modelled as a completely positive map, it would be a measurement–prepare channel of the form

$$W : \mathcal{B}(C\mathcal{X}) \longrightarrow \mathcal{B}(\mathcal{H})$$

$$\sigma \longmapsto \sum_x \langle x|\sigma|x\rangle W_x.$$

I.e., its use involves a “guaranteed (von Neumann) measurement” on all messages which come from Alice.

Regarding theorem 10, we conjecture the achievable rate stated there to remain optimal even if unlimited noiseless quantum communication is allowed. There is however the much more interesting question of more general quantum channels, for example a depolarising qubit channel, the quantum analogue of a binary symmetric channel: does it allow bit commitment, and if so, at which rate?

This generalisation may be significant because first of all, information theoretically secure bit commitment is not possible with noiseless quantum communication [20]. Here we have shown that it is possible under the assumption of a noisy channel. This opens the possibility of perhaps *having* bit commitment under realistic conditions where one can ensure that all available channels are noisy.

## 7 Discussion

We have considered bit–string commitment by using a noisy channel and have characterised the exact capacity for this task by a single–letter formula. This implies a lower bound on the coin tossing capacity of that channel by the same formula, which in fact we conjecture to be an equality.

Satisfactory as this result is, it has to be noted that we are not able in general to provide an explicit protocol: our proof is based on the random coding technique and shows only existence. What is more, even if one finds a good code it will most likely be inefficient: the codebook is just the list of  $\xi_{a\mu}$ . In this connection we conjecture that the commitment capacity can be achieved by random linear codes (compare the situation for channel coding!). It is in any case an open problem to find efficient good codes, even for the binary symmetric channel. Note that we only demand efficient *encoding* — there is no decoding of errors in our scheme, only an easily performed test.

Our scheme is a block–coding method: Alice has to know the whole of her message, say a bit string, before she can encode. One might want to use our result as a building block in other protocols which involve committing to bits at various stages — then the natural question arises whether there is an “online” version which would allow Alice to encode and send bits as she goes along.

In the same direction of better applicability it would be desirable to extend our results to a more robust notion of channel: compare the work of [15] where a cheater is granted partial control over the channel characteristics. Still, the fixed channel is not beyond application: note that it can be simulated by pre-distributed data from a trusted party via a “noisy one-time pad” (compare [3] and [24]).

Another open question of interest is to determine the reliability function, i.e., the optimal asymptotic rate of the error  $\epsilon + \delta$  (note that implicit in our proposition 8 is a lower bound): it is especially interesting at  $R = 0$ , because there the rate tells exactly how secure single-bit commitment can be made.

Finally, we have outlined that a class of quantum channels also allows bit commitment: they even have a commitment capacity of the same form as the classical result. This opens up the possibility of unconditionally secure bit commitment for other noisy quantum channels.

We hope that our work will stimulate the search for optimal rates of other cryptographic primitives, some of which are possible based on noise, e.g. oblivious transfer.

## Acknowledgements

We thank Ning Cai and Raymond W. Yeung for sharing their ideas on the quantum wiretap channel, and making available to us their manuscript [10]. We also acknowledge interesting discussions with J. Müller-Quade and P. Tuyls at an early stage of this project.

AW is supported by the U.K. Engineering and Physical Sciences Research Council. ACAN and HI are supported by the project “Research and Development of Quantum Cryptography” of the Telecommunications Advancement Organisation as part of the programme “Research and Development on Quantum Communication Technology” of the Ministry of Public Management, Home Affairs, Posts and Telecommunications, Japan.

## A Typical sequences

This appendix collects some facts about typical sequences used in the main body of the text. We follow largely the book of Csiszár and Körner [14].

The fundamental fact we shall use is the following large deviation version of the law of large numbers:

**Lemma 13 (Chernoff [11])** *For i.i.d. random variables  $X_1, \dots, X_N$ , with  $0 \leq$*

$X_n \leq 1$  and with expectation  $EX_n = p$ :

$$\Pr \left\{ \frac{1}{N} \sum_{n=1}^N X_n \geq (1 + \eta)p \right\} \leq \exp \left( -N \frac{p\eta^2}{2 \ln 2} \right),$$

$$\Pr \left\{ \frac{1}{N} \sum_{n=1}^N X_n \leq (1 - \eta)p \right\} \leq \exp \left( -N \frac{p\eta^2}{2 \ln 2} \right).$$

□

For a probability distribution  $P$  on  $\mathcal{X}$  and  $\epsilon > 0$  define the set of  $\epsilon$ -typical sequences:

$$\mathcal{T}_{P,\epsilon}^n = \left\{ x^n : \forall x \mid N(x|x^n) - P(x)n \mid \leq \epsilon n \ \& \ P(x) = 0 \Rightarrow N(x|x^n) = 0 \right\},$$

with the number  $N(x|x^n)$  denoting the number of letters  $x$  in the word  $x^n$ . The probability distribution  $P_{x^n}(x) = \frac{1}{n} N(x|x^n)$  is called the *type* of  $x^n$ . Note that  $x^n \in \mathcal{T}_{P,\epsilon}^n$  is equivalent to  $|P_{x^n}(x) - P(x)| \leq \epsilon$  for all  $x$ .

These are the properties of typical sequences we shall need:

$$P^{\otimes n}(\mathcal{T}_{P,\epsilon}^n) \geq 1 - 2|\mathcal{X}| \exp(-n\epsilon^2/2). \quad (9)$$

This is an easy consequence of the Chernoff bound, lemma 13, applied to the indicator variables  $X_k$  of the letter  $x$  in position  $k$  in  $X^n$ , with  $\eta = \epsilon P(x)^{-1}$ .

$$\forall x^n \in \mathcal{T}_{P,\epsilon}^n \quad \begin{cases} P^{\otimes n}(x^n) \leq \exp(-nH(P) + n\epsilon D), \\ P^{\otimes n}(x^n) \geq \exp(-nH(P) - n\epsilon D), \end{cases} \quad (10)$$

with the constant  $D = \sum_{x:P(x) \neq 0} -\log P(x)$ . See [14].

$$|\mathcal{T}_{P,\epsilon}^n| \leq \exp(nH(P) + n\epsilon D), \quad (11)$$

$$|\mathcal{T}_{P,\epsilon}^n| \geq \left( 1 - 2|\mathcal{X}| \exp(-n\epsilon^2/2) \right) \exp(nH(P) - n\epsilon D). \quad (12)$$

This follows from eq. (10). These estimates also allow to lower bound the size of sets with large probability: assume  $P^{\otimes n}(\mathcal{C}) \geq \eta$ , then

$$|\mathcal{C}| \geq \left( \eta - 2|\mathcal{X}| \exp(-n\epsilon^2/2) \right) \exp(nH(P) - n\epsilon D). \quad (13)$$

We also use these notions in the “non-stationary” case: consider a channel  $W : \mathcal{X} \rightarrow \mathcal{Z}$ , and an input string  $x^n \in \mathcal{X}^n$ . Then define, with  $\epsilon > 0$ , the set of conditional  $\epsilon$ -typical sequences:

$$\begin{aligned} \mathcal{T}_{W,\epsilon}^n(x^n) &= \left\{ z^n : \forall x, z \mid N(xz|x^n z^n) - nW(z|x)P_{x^n}(x) \mid \leq \epsilon n \right. \\ &\quad \left. \& \ W(z|x) = 0 \Rightarrow N(xz|x^n z^n) = 0 \right\} \\ &= \prod_x \mathcal{T}_{W_x, \epsilon P_{x^n}(x)^{-1}}^{\mathcal{I}_x} \end{aligned}$$

with the sets  $\mathcal{I}_x$  of positions in the word  $x^n$  where  $x_k = x$ . The latter product representation allows to easily transport all of the above relations for typical sequences to conditional typical sequences:

$$W_{x^n}^n (\mathcal{T}_{W,\epsilon}^n(x^n)) \geq 1 - 2|\mathcal{X}||\mathcal{Z}| \exp(-n\epsilon^2/2). \quad (14)$$

$$\forall x^n \in \mathcal{T}_{W,\epsilon}^n(x^n) \quad \begin{cases} W_{x^n}^n(x^n) \leq \exp(-nH(W|P_{x^n}) + n\epsilon E), \\ W_{x^n}^n(x^n) \geq \exp(-nH(W|P_{x^n}) - n\epsilon E), \end{cases} \quad (15)$$

with  $E = \max_x \sum_{z:W_x(z) \neq 0} -\log W_x(z)$  and the conditional entropy  $H(W|P) = \sum_x P(x)H(W_x)$ .

$$|\mathcal{T}_{W,\epsilon}^n(x^n)| \leq \exp(nH(W|P_{x^n}) + n\epsilon E), \quad (16)$$

$$|\mathcal{T}_{W,\epsilon}^n(x^n)| \geq \left(1 - 2|\mathcal{X}||\mathcal{Z}| \exp(-n\epsilon^2/2)\right) \exp(nH(W|P_{x^n}) - n\epsilon E). \quad (17)$$

A last elementary property: for  $x^n$  of type  $P$  and output distribution  $Q$ , with  $Q(z) = \sum_x P(x)W_x(z)$ ,

$$\mathcal{T}_{W,\epsilon}^n(x^n) \subset \mathcal{T}_{Q,\epsilon}^n|\mathcal{X}|. \quad (18)$$

As an application, let us prove the following lemma:

**Lemma 14** *For words  $x^n$  and  $y^n$  with  $d_H(x^n, y^n) \geq \sigma n$ , and a channel  $W$  such that*

$$\forall x \in \mathcal{X}, P \text{ p.d. with } P(x) = 0 \quad \left\| W_x - \sum_y P(y)W_y \right\|_1 \geq \eta,$$

one has, with  $\epsilon = \frac{\sigma^2 \eta}{2|\mathcal{X}|^2|\mathcal{Z}|}$ ,

$$W_{y^n}^n (\mathcal{T}_{W,\epsilon}^n(x^n)) \leq 2 \exp(-n\epsilon^4/2)$$

*Proof.* There exists an  $x$  such that the word  $x^{\mathcal{I}_x}$  (composed of letters  $x$  only) has distance at least  $\frac{1}{|\mathcal{X}|}\sigma n$  from  $y^{\mathcal{I}_x}$ . In particular,  $N_x := N(x|x^n) = |\mathcal{I}_x| \geq \frac{1}{|\mathcal{X}|}\sigma n$ .

This implies also, by assumption on the channel,

$$\left\| \frac{1}{N_x} \sum_{k \in \mathcal{I}_x} W_{y_k} - W_x \right\|_1 \geq \frac{1}{|\mathcal{X}|}\sigma \eta.$$

Hence there must be a  $z \in \mathcal{Z}$  with

$$\left| \frac{1}{N_x} \sum_{k \in \mathcal{I}_x} W_{y_k}(z) - W_x(z) \right| \geq \frac{1}{|\mathcal{X}||\mathcal{Z}|}\sigma \eta.$$

By definition, this in turn implies that for all  $z^n \in \mathcal{T}_{W,\epsilon}^n(x^n)$ ,

$$\left| N(z|z^{\mathcal{I}_x}) - \sum_{k \in \mathcal{I}_x} W_{y_k}(z) \right| \geq \frac{1}{2|\mathcal{X}||\mathcal{Z}|}\sigma \eta N_x.$$

Introducing the sets  $\mathcal{J}_{xy} = \{k \in \mathcal{I}_x : y_k = y\}$ , with cardinalities  $N_{xy} = |\mathcal{I}_{yx}|$ , there is a  $y$  such that (still for all  $z^n \in \mathcal{T}_{W,\epsilon}^n(x^n)$ ),

$$\begin{aligned} |N(z|z^{\mathcal{J}_{xy}}) - N_{xy}W_y(z)| &\geq \frac{1}{2|\mathcal{X}|^2|\mathcal{Z}|}\sigma\eta N_x \\ &\geq \frac{1}{2|\mathcal{X}|^2|\mathcal{Z}|}\sigma\eta N_{xy}. \end{aligned}$$

This implies

$$N_{xy} \geq \frac{1}{4|\mathcal{X}|^2|\mathcal{Z}|}\sigma\eta N_x \geq \frac{1}{4|\mathcal{X}|^3|\mathcal{Z}|}\sigma^2\eta n,$$

and with lemma 13 we obtain the claim.  $\square$

## References

- [1] R. Ahlswede, I. Csiszár, “Common Randomness in Information Theory and Cryptography – Part I: Secret Sharing”, *IEEE Trans. Inf. Theory*, vol. 39, no. 4, pp. 1121–1132, 1993.
- [2] R. Ahlswede, A. Winter, “Strong converse for identification via quantum channels”, *IEEE Trans. Inf. Theory*, vol. 48, no. 3, pp. 569–579, 2002. Addendum *ibid.*, vol. 49, no. 1, p. 346, 2003.
- [3] D. Beaver, “Commodity–Based Cryptography” (Extended Abstract), *Proc. 29<sup>th</sup> Annual ACM Symposium on the Theory of Computing (El Paso, TX, 4–6 May 1997)*, pp. 446–455, ACM, 1997.
- [4] C. H. Bennett, G. Brassard, “Quantum Cryptography: Public Key Distribution and Coin Tossing”, *Proc. IEEE Int. Conf. on Computers Systems and Signal Processing, Bangalore (India)*, pp. 175–179, 1984.
- [5] C. H. Bennett, G. Brassard, C. Crépeau, U. Maurer, “Generalized Privacy Amplification”, *IEEE Trans. Inf. Theory*, vol. 41, no. 6, pp. 1915–1923, 1995.
- [6] C. H. Bennett, P. W. Shor, “Quantum Information Theory”, *IEEE Trans. Inf. Theory*, vol. 44, no. 6, pp. 2724–2742, 1998.
- [7] M. Blum, “Coin flipping by telephone: a protocol for solving impossible problems”, *Proc. IEEE Computer Conference*, pp. 133–137, 1982.
- [8] G. Brassard, D. Chaum, C. Crépeau, “Minimum disclosure proofs of knowledge”, *J. Computer Syst. Sci.*, vol. 37, pp. 156–189, 1988.
- [9] G. Brassard, C. Crépeau, M. Yung, “Constant–round perfect zero–knowledge computationally convincing protocols”, *Theoretical Computer Science*, vol. 84, pp. 23–52, 1991.

- [10] N. Cai, R. W. Yeung, “Quantum Privacy and Quantum Wiretap Channels”, manuscript, 2003.
- [11] H. Chernoff, “A measure of asymptotic efficiency for tests of a hypothesis based on the sum of observations”, *Ann. Math. Statistics*, vol. 23, pp. 493–507, 1952.
- [12] C. Crépeau, “Efficient Cryptographic Protocols Based on Noisy Channels”, *Advances in Cryptology: Proc. EUROCRYPT 1997*, pp. 306–317, Springer 1997.
- [13] C. Crépeau, J. Kilian, “Achieving oblivious transfer using weakened security assumptions”, *Proc. 29<sup>th</sup> FOCS*, pp. 42–52. IEEE, 1988.
- [14] I. Csiszár, J. Körner, *Information Theory: Coding Theorems for Discrete Memoryless Channels*, Academic Press, NY 1981.
- [15] I. B. Damgård, J. Kilian, L. Salvail, “On the (Im)possibility of Basing Oblivious Transfer and Bit Commitment on Weakened Security Assumptions”, *Advances in Cryptology: EUROCRYPT 1999*, pp. 56–73, Springer 1999.
- [16] S. Halevi, “Efficient commitment schemes with bounded sender and unbounded receiver”, *Proc. CRYPTO 1995*, pp. 84–96. LNCS 963, Springer Verlag, 1995.
- [17] S. Halevi, S. Micali, “Practical and Provably-Secure Commitment Schemes from Collision Free Hashing”, *Advances in Cryptology: CRYPTO 1996*, pp. 201–215, LNCS 1109, Springer Verlag, 1996.
- [18] A. S. Holevo, “Bounds for the quantity of information transmitted by a quantum channel”, *Probl. Inf. Transm.*, vol. 9, no. 3, pp. 177–183, 1973.
- [19] U. Maurer, “Protocols for Secret Key Agreement by Public Discussion Based on Common Information”, *Advances in Cryptology: CRYPTO 1992*, pp. 461–470, Springer 1993. “Secret Key Agreement by Public Discussion”, *IEEE Trans. Inf. Theory*, vol. 39, no. 3, pp. 733–742, 1993.
- [20] D. Mayers, “Unconditionally secure quantum bit commitment is impossible”, *Phys. Rev. Letters*, vol. 78, no. 17, pp. 3414–3417, 1997.
- [21] M. Naor, “Bit commitment using pseudo-randomness”, *J. Cryptology*, vol. 2, no. 2, pp. 151–158, 1991.
- [22] M. A. Nielsen, I. L. Chuang, *Quantum Computation and Quantum Information*, Cambridge University Press, 2000.
- [23] R. Ostrovsky, R. Venkatesan, M. Yung, “Secure commitments against a powerful adversary”, *Proc. STACS 1992*, pp. 439–448, LNCS 577, Springer Verlag, 1992.

- [24] R. L. Rivest, “Unconditionally Secure Commitment and Oblivious Transfer Schemes Using Private Channels and a Trusted Initializer”, unpublished manuscript, 1999.
- [25] C. E. Shannon, “A mathematical theory of communication”, Bell System Tech. Journal, vol. 27, pp. 379–423 and 623–656, 1948.
- [26] A. Wyner, “The Wire Tap Channel”, Bell System Tech. Journal, vol. 54, pp. 1355–1387, 1975.