Syndrome-coding for the wiretap channel revisited

Gérard Cohen and Gilles Zémor
ENST and CNRS
46 rue Barrault, 75634 Paris 13,
FRANCE
{cohen,zemor}@enst.fr

Abstract — **To communicate an r-bit secret $\mathbf{s}$ through a wire-tap channel, the syndrome coding strategy consists of choosing a linear transformation h and transmitting an n-bit vector $\mathbf{x}$ such that $h(\mathbf{x}) = \mathbf{s}$. The receiver obtains a corrupted version of $\mathbf{x}$ and the eavesdropper an even more corrupted version of $\mathbf{x}$: the (syndrome) function h should be chosen in such a way as to minimize both the length n of the transmitted vector and the information leakage to the eavesdropper. We give a refined analysis of the information leakage that involves m-th moment methods.**

I. Introduction

The wire-tap channel was introduced by Wyner [11], as a special case of a broadcast channel defined by Cover [4](one sender, two receivers subjected to discrete memoriless channels). In this model Alice transmits a n-bit string $\mathbf{x}$ to Bob who receives a corrupted version $\mathbf{y}$ while the eavesdropper Eve receives an even more strongly corrupted string $\mathbf{z}$. Alice would like to transmit a secret string $\mathbf{s}$ of length r to Bob while ensuring that almost no information about $\mathbf{s}$ is leaked to Eve. In his original paper [11] Wyner solved the capacity problem when both Bob's and Eve's channels are discrete and memoriless. His method is existential and non-effective. The problem was then generalized by not requiring that Eve's reception be a degraded version of Bob's and solved in [6]. Wyner also introduced syndrome coding to solve the particular case when the main channel (between Alice and Bob) is noiseless and Eve receives $\mathbf{x}$ corrupted by a binary symmetric channel with transition probability p.

In short, let σ be the syndrome function of some linear code C. This just means that σ is the function

$$\begin{array}{rcl} \sigma : \{0,1\}^n & \longrightarrow & \{0,1\}^r \\ \mathbf{x} & \mapsto & \mathbf{H}\,{}^t\mathbf{x} \end{array}$$

for some $r \times n$ matrix $\mathbf{H}$ that can be thought of as a parity-check matrix of some linear code C.

Let $\mathbf{s} \in \{0,1\}^r$ be the secret message Alice wants to transmit to Bob. Alice sends over the channel a vector $\mathbf{x} \in \{0,1\}^n$, randomly chosen among all vectors such that $\sigma(\mathbf{x}) = \mathbf{s}$. Wyner shows by a non-constructive argument that, as long as the size r of the syndrome space is chosen to be smaller than the Shannon entropy of the binary symmetric channel, there exist codes that leak a vanishing *proportion* (with respect to the length r of the secret) of information bits on the secret to Eve.

In the present paper we shall be interested in more general *additive channels*: this means that if $\mathbf{x}$ is transmitted from Alice to Bob, Eve receives

$$\mathbf{z} = \mathbf{x} + \mathbf{b}$$

where $\mathbf{b} \in \{0,1\}^n$ is a random variable with any given probability distribution.

Wire-tap channels were rejuvenated during the 1990's under the name of *privacy amplification* [1]. The problem under study is a slight variation of the original wire-tap problem. In this setting Alice and Bob just want to share a common secret $\mathbf{s} \in \{0,1\}^r$, but they don't care what its actual value is: indeed, they want this value to be as random as possible because its purpose is to be a shared secret key for a classical cryptographic cipher, for example. The overall strategy in this case is similar, it is again for Alice to send Bob a message $\mathbf{x}$, and to decide with Bob that their shared secret will be $\mathbf{s} = h(\mathbf{x})$ where h is a properly chosen function. The goal is to minimize the quantity of information bits leaked to Eve, i.e. the quantity

$$H(h(\mathbf{x})\,|\mathbf{x} + \mathbf{b}). \tag{1}$$

The main difference between *transmitting* a secret (wire-tap channel) and *sharing* a secret (privacy amplification) is that in the latter case one has more leeway in choosing the function h. In particular it does not matter if we don't know how to find $\mathbf{x}$ such that $h(\mathbf{x})$ equals a given secret $\mathbf{s}$. We only need to ensure that randomly choosing $\mathbf{x}$ produces a uniformly distributed secret $\mathbf{s}$.

The paper [1] strengthens Wyner's result in the following sense: channels are more general and the estimate of the number of bits leaked to Eve is stronger. Specifically, in the setting of additive channels, it shows that if the size r of the secret space is taken to equal the *Renyi* entropy of $\mathbf{b}$, then the average *number* of information bits leaked to Eve (1) when h is randomly chosen from a family $\mathcal{H}$, is not more than 1 and can be made exponentially small in r, if r is taken to be smaller than the Renyi entropy of $\mathbf{b}$. For this to work, it is enough that $\mathcal{H}$ be a universal class of (hash) functions. Note that it applies in particular when $\mathcal{H}$ is the set of syndrome (i.e. linear) functions, so that these results are applicable to the wire-tap setting as well.

Finally let us mention yet another set of relevant results from another school, that of "extracting randomness", see e.g. [9, 10, 7]. The purpose is not necessarily motivated by cryptography and more generally is to transform a non-uniform random source $\mathbf{b}$ into an almost uniformly distributed random variable $h(\mathbf{b})$ by applying a randomly chosen function h from a class $\mathcal{H}$. The goal is usually not so much to obtain an ultra-fine measure of the closeness to the uniform distribution of $h(\mathbf{b})$, but to minimize the amount of randomness in the choosing of h by making the size of the class of functions $\mathcal{H}$ as small as possible. The source $\mathbf{b}$ this time can have any probability distribution, but the maximum length of the "secret" $h(\mathbf{b})$ is the *min-entropy* of $\mathbf{b}$ (rather than its Renyi entropy). The quality of the distribution D of $h(\mathbf{b})$ is measured by the average (over all possible functions h) of the L_1-distance between D and the uniform distribution.

The motivation for the present paper is the need for a stronger measure of the closeness to the uniform distribution of the functions $h(\mathbf{b})$. This is best illustrated with a real-life cryptographic example. Take the example of syndrome coding for the wiretap channel, so that if Eve computes $\sigma(\mathbf{z}) = \mathbf{s} + \sigma(\mathbf{b})$ from the received vector, she gets the secret $\mathbf{s}$ corrupted by the random quantity $\sigma(\mathbf{b})$. Suppose the length of the secret $\mathbf{s}$ is that of a standard secret-key cryptosystem, e.g. $r = 128$ bits. Even if the difference in Shannon entropy (or for that matter the L_1-distance) between the probability distribution of $\sigma(\mathbf{b})$ and that of the uniform distribution is only a fraction of a bit, say ϵ, this does not necessarily rule out the existence of nasty cryptanalytic attacks. For example, take the probability measure P on $\{0,1\}^r$ such that $P(v) = 1/1000$, and $P(\mathbf{s}) = (1 - 1/1000)/(2^r - 1)$ for $\mathbf{s} \neq v$. Then the difference of P to the uniform distribution, measured both in Shannon entropy or in L_1-distance, is about one tenth of a bit. However, the attacker will bet on the most probable syndrome value v, and be right (and therefore discover the secret key $\mathbf{s}$) on average once in a thousand. This is in practice an unacceptable level of security.

We see therefore that the measures of the randomness of $\sigma(\mathbf{b})$ highlighted above are not always of sufficient quality to defend against the existence of this sort of attack. This is made worse by the fact that results that rely upon universal hashing are averaged on the choice of the function h (in our case the syndrome function σ). If one wants a fixed hash function h, how does one choose it ? If a randomly chosen h leaks on average ϵ bits, we know that there must exist an h that leaks not more than ϵ bits. However this is highly non-constructive, we have no way of making sure a given h is good enough. We can avoid this difficulty with probability estimates, but since we hardly have anything else at our disposal besides Markov inequality, this degrades the randomness estimate (we can guarantee that with probability $1 - 1/100$, h will leak not more than 100ϵ bits).

To counteract the most-likely-syndrome attack, one must show that the most likely syndrome $\sigma(\mathbf{b})$ does not have too high a probability of occurence. To this end we shall look for a lower bound on the min-entropy of the distribution of $\sigma(\mathbf{b})$, i.e. $-\log_2(\max_{v\in\{0,1\}^r} P(\sigma(\mathbf{b}) = v)$. We shall take a renewed look at the syndrome coding strategy for the wire-tap channel: our main result is the following theorem.

Theorem 1 *Let $\mathbf{b}$ be a random binary vector of length n with a fixed probability distribution, with min-entropy $H_\infty(\mathbf{b}) = r$. Let $\mathbf{H}$ be a uniformly randomly chosen $r \times n$ binary matrix, and let σ be the associated syndrome function. The probability, over the choice of $\mathbf{H}$, that $H_\infty(\sigma(\mathbf{b})) < r - \log_2(1 + 2^m)$ is not more than $2^{-3m^2/4+m+r}$.*

To illustrate this result, take up again the above numerical example: suppose $\mathbf{b}$ is any random source with min-entropy at least equal to the secret size of $r = 128$ bits. Set $m = 2(r)^{1/2}$. Then, by choosing the linear function σ at random, we can guarantee that with probability at least $1 - 2^{-233}$, every syndrome value has a probability of occurence less than 2^{-105}.

As another illustration, pick a m such that
$m = o(r) = o(m^2)$; e.g., $m = r/\log r$.
Then $P(H_\infty(\sigma(\mathbf{b})) < r - o(r)) \leq 2^{-3r^2/4(1-o(1))}$.

II. Informational and coding tools

We shall use the following notions (see [5] for details):

- $H(X)$: (Shannon) entropy of a random variable X. This is the usual entropy in communications, source and channel coding.

$$H(X) := \sum_x P(X = x) \log_2(1/P(X = x)).$$

- $H(R|T) = E_T[H(R|T = t)]$ is the *conditional entropy* or *equivocation* of T about R.

- $R(X)$: Renyi entropy (of order two). Denote by $P_c(X) = \Sigma_x P(X = x)^2$ the (collision) probability that X takes the same value twice after two random independent experiments. Then $R(X) := -\log_2(P_c(X))$.

  Renyi entropy is used to measure randomness produced by universal hashing (see, e.g., [1]).

- $H_\infty(X)$: min-entropy of X.

  $H_\infty(X) := \text{Max } \{j : \forall x : Pr\{X = x\} \leq 2^{-j}\}$,

  Equivalently, $H_\infty(X) = -\log_2(\max_x P(X = x))$.

  $H_\infty(X)$ measures the minimum amount of information conveyed by a realization of X; it is also the minimum work factor for an adversarial guessing strategy (namely, bet on the most probable outcome).

By noting that

$$(\max_i\{p_i\})^2 \leq \Sigma p_i^2 \leq (\max_i\{p_i\})(\Sigma p_i) = \max_i\{p_i\},$$

we get: $H_\infty(X) \leq R(X) \leq 2H_\infty(X)$.

It is also easy to check that $R(X) \leq H(X)$.

We also need some coding terminology (see [2] for an account centered on coverings): A code C has parameters $[n,k]$ if it is a linear subspace of dimension k of the n-dimensional binary Hamming space (hypercube). A *parity-check matrix* $\mathbf{H}$ is formed by writing as rows a basis of the dual code. This means that if the syndrome function associated to the matrix $\mathbf{H}$ is defined by:

$$\begin{array}{rcl} \sigma : \{0,1\}^n & \longrightarrow & \{0,1\}^r \\ \mathbf{x} & \mapsto & \mathbf{H}^t\mathbf{x} \end{array}$$

then the code C is the set of vectors $\mathbf{x}$ such that $\sigma(\mathbf{x}) = 0$.

III. The coset-coding scheme

A. Description

Let C denote a binary linear code of length n together with an $r \times n$ parity-check matrix $\mathbf{H}$.

Let the secret $\mathbf{s}$ be a given vector in the syndrome space $\{0,1\}^r$.

Let the vector $\mathbf{x}$ be chosen *uniformly* among the vectors of syndrome $\mathbf{s}$ (i.e. in a given coset of C). Note that this is constructive:

1. Pick an "easy" vector $\mathbf{y}$ with syndrome $\mathbf{s}$ (For example, if $\mathbf{H}$ is in *systematic* form,

   $\mathbf{H} = [\mathbf{I}_r \mid \mathbf{P}]$, where $\mathbf{I}_r$ is the identity matrix of order r:

   $\mathbf{y} = \Sigma_{i \in supp(\mathbf{s})} e^i$, with $\{e^i\}$ the natural basis.)
2. Add a *random* $\mathbf{c} \in C$, i.e. a random combination of $n-r$ generating codewords, to $\mathbf{y}$;
3. Transmit $\mathbf{x} = \mathbf{y} + \mathbf{c}$.

B. Eavesdropper's uncertainty

Given $\mathbf{z} = \mathbf{x} + \mathbf{b}$, Eve can compute $\sigma(\mathbf{z}) = \mathbf{s} + \sigma(\mathbf{b})$, which we can state informally by saying that the eavesdropper is submitted to a one-time pad in the syndrome space. In the syndrome space, we see therefore that Eve's equivocation on the secret $\mathbf{s}$ is directly linked to the closeness to the uniform distribution of $\mathbf{s}$. However, since Eve also has $\mathbf{x} + \mathbf{b}$, one might be object that Eve does not necessarily have to be bound by the syndrome space: but a little thought shows that this is indeed the case. In other words, whatever Eve can do with $\mathbf{x}+\mathbf{b}$, she can do with $\mathbf{s}+\sigma(\mathbf{b})$ alone.

More formally, start by noticing that

$$H(\mathbf{s} \mid \mathbf{x}+\mathbf{b}) - H(\mathbf{s} \mid \mathbf{s}+\mathbf{H}^t\mathbf{b}) \leq 0.$$

This is because, since $\mathbf{s}+\mathbf{H}^t\mathbf{b}$ is a function of $\mathbf{x}+\mathbf{b}$, knowledge of $\mathbf{x}+\mathbf{b}$ can only yield more knowledge (and less uncertainty) than $\mathbf{s}+\mathbf{H}^t\mathbf{b}$. Let us now prove the reverse inequality: we have $H(\mathbf{s} \mid \mathbf{x}+\mathbf{b}) - H(\mathbf{s} \mid \mathbf{s}+\mathbf{H}^t\mathbf{b})$

$$\begin{array}{rl} = & H(\mathbf{s}, \mathbf{x}+\mathbf{b}) - H(\mathbf{x}+\mathbf{b}) \\ & -H(\mathbf{s}, \mathbf{s}+\mathbf{H}^t\mathbf{b}) + H(\mathbf{s}+\mathbf{H}^t\mathbf{b}) \\ = & H(\mathbf{s}, \mathbf{x}+\mathbf{b}, \mathbf{s}+\mathbf{H}^t\mathbf{b}) - H(\mathbf{s}, \mathbf{s}+\mathbf{H}^t\mathbf{b}) \\ & -[H(\mathbf{x}+\mathbf{b}, \mathbf{s}+\mathbf{H}^t\mathbf{b}) - H(\mathbf{s}+\mathbf{H}^t\mathbf{b})] \\ = & H(\mathbf{x}+\mathbf{b} \mid \mathbf{s}, \mathbf{s}+\mathbf{H}^t\mathbf{b}) - H(\mathbf{x}+\mathbf{b} \mid \mathbf{s}+\mathbf{H}^t\mathbf{b}) \\ \geq & H(\mathbf{x} \mid \mathbf{s}, \mathbf{b}) - H(\mathbf{x}+\mathbf{b} \mid \mathbf{s}+\mathbf{H}^t\mathbf{b}) \\ = & H(\mathbf{x} \mid \mathbf{s}) - H(\mathbf{x}+\mathbf{b} \mid \mathbf{s}+\mathbf{H}^t\mathbf{b}) \\ \geq & 0, \end{array}$$

where the last inequality is due to $\mathbf{x}$ being uniformly distributed among vectors with syndrome $\mathbf{s}$, hence the maximality of $H(\mathbf{x} \mid \mathbf{s})$.

We have therefore proved that $H(\mathbf{s} \mid \mathbf{x}+\mathbf{b}) = H(\mathbf{s} \mid \mathbf{s}+\mathbf{H}^t\mathbf{b})$, meaning that there is no advantage for the eavesdropper in possessing $\mathbf{x}+\mathbf{b}$ on top of its syndrome. Note that we did not need to suppose *anything* on the distribution of $\mathbf{s}$.

IV. The m-th moment method

This section is devoted to the proof of Theorem 1. We shall treat the noise $\mathbf{b}$ as a random variable with probability distribution P, i.e. for any $\mathbf{x} \in \{0,1\}^n$, we have $P(\mathbf{b} = \mathbf{x}) = P(\mathbf{x})$. We start with the immediate

Lemma 1 *For $\mathbf{x} \neq 0, \mathbf{s}$ fixed, $\mathbf{H}$ uniformly distributed: $Pr\{\mathbf{H}^t\mathbf{x} = \mathbf{s}\} = 2^{-r}$.*

For any given $\mathbf{x} \neq 0$ and $\mathbf{s}$ define the Bernoulli random variable $X_{\mathbf{x},\mathbf{s}} = 1$ if $\mathbf{H}^t\mathbf{x} = \mathbf{s}$, $X_{\mathbf{x},\mathbf{s}} = 0$ otherwise; Lemma 1 translates as:

$$E[X_{\mathbf{x},\mathbf{s}}] = 2^{-r}.$$

Now define the random variable $X_{\mathbf{s}} = \Sigma_{\mathbf{x} \in \{0,1\}^n, \mathbf{x} \neq 0} P(\mathbf{x}) X_{\mathbf{x},\mathbf{s}}$. For $\mathbf{s} \neq 0$, this quantity equals the probability that $\sigma(\mathbf{b})$ equals $\mathbf{s}$, viewed as a random quantity over the space of matrices $\mathbf{H}$. The probability that $\sigma(\mathbf{b})$ equals the zero syndrome is:

$$P(0) + \Sigma_{\mathbf{x} \neq 0} P(\mathbf{x}) X_{\mathbf{x},0}.$$

The core result is the following:

Lemma 2 *If $r \leq H_\infty(\mathbf{b})$, then, for any $\mathbf{s} \in \{0,1\}^r$ and for any integer $m \leq r$, $E[X_{\mathbf{s}}^m] \leq 2^{m+m^2/4-mr}$.*

Proof of Lemma 2: Denote $V = \{0,1\}^n$. Let rk denote the linear rank function: the mth moment $E[X_{\mathbf{s}}^m]$ satisfies (2) $-$ $-$(7) (see next page). To obtain (4), consider that to generate all m-tuples of rank j, one may first choose j coordinates among m, then choose a full-rank j-tuple on these coordinates, and fill in the remaining coordinates. Furthermore, if $\mathbf{x}_{j+1}, \ldots, \mathbf{x}_m$ are linear combinations of $\mathbf{x}_1, \ldots \mathbf{x}_j$, then the (Bernoulli) random variable $X_{\mathbf{x}_1,\mathbf{s}} \ldots X_{\mathbf{x}_m,\mathbf{s}}$ either equals $X_{\mathbf{x}_1,\mathbf{s}} \ldots X_{\mathbf{x}_j,\mathbf{s}}$, or

$$
\begin{aligned}
E[X_{\mathbf{s}}^m] &= \sum_{(\mathbf{x}_1,\dots\mathbf{x}_m)\in(V\setminus\{0\})^m} P(\mathbf{x}_1)\dots P(\mathbf{x}_m)E[X_{\mathbf{x}_1,\mathbf{s}}\dots X_{\mathbf{x}_m,\mathbf{s}}] && (2)\\
&= \sum_{j=1}^{m} \sum_{rk(\mathbf{x}_1,\dots\mathbf{x}_m)=j} P(\mathbf{x}_1)\dots P(\mathbf{x}_m)E[X_{\mathbf{x}_1,\mathbf{s}}\dots X_{\mathbf{x}_m,\mathbf{s}}] && (3)\\
&\le \sum_{j=1}^{m} \binom{m}{j} \sum_{rk(\mathbf{x}_1,\dots\mathbf{x}_m)=rk(\mathbf{x}_1,\dots\mathbf{x}_j)=j} P(\mathbf{x}_1)\dots P(\mathbf{x}_j)\dots P(\mathbf{x}_m)E[X_{\mathbf{x}_1,\mathbf{s}}\dots X_{\mathbf{x}_j,\mathbf{s}}] && (4)\\
&\le \sum_{j=1}^{m} \binom{m}{j} \sum_{(\mathbf{x}_1,\dots\mathbf{x}_j)\in(V\setminus\{0\})^j} 2^{j(m-j)}P(\mathbf{x}_1)\dots P(\mathbf{x}_j)(\max_{u\in V} P(u))^{m-j}E[X_{\mathbf{x}_1,\mathbf{s}}]^j && (5)\\
&\le \sum_{j=1}^{m} \binom{m}{j} 2^{m^2/4}2^{-r(m-j)} \sum_{(\mathbf{x}_1,\dots\mathbf{x}_j)\in(V\setminus\{0\})^j} P(\mathbf{x}_1)\dots P(\mathbf{x}_j)2^{-jr} && (6)\\
&\le 2^{m^2/4+m}2^{-mr}. && (7)
\end{aligned}
$$

equals zero (if $\mathbf{s}\neq 0$ and some $\mathbf{x}_i$, $i>j$, is an even-weight linear combination of $\mathbf{x}_1,\dots\mathbf{x}_j$).

To obtain (5), use the fact that $\mathbf{x}_1,\dots,\mathbf{x}_j$ are linearly independent implies that the random variables $X_{\mathbf{x}_1,\mathbf{s}}\dots X_{\mathbf{x}_j,\mathbf{s}}$ are independent.

To obtain (6), recall that the hypothesis $r\le H_\infty(\mathbf{b})$ means that $P(\mathbf{u})\le 2^{-r}$ for any $\mathbf{u}\in V$. Bound from above all terms $2^{j(m-j)}$ by $2^{m^2/4}$ and apply Lemma 1 to $E[X_{\mathbf{x}_1,\mathbf{s}}]$.

Finally, since P is a probability measure, the sums $\sum P(\mathbf{x}_1)\dots P(\mathbf{x}_j)$ equal $(1-P(0))^j$ and are upper bounded by 1. ∎

Proof of Theorem 1: We invoke the "Markov Inequality of order m", stating that for a positive random variable Y and real number λ: $P(Y>\lambda\}\le E[Y^m]/\lambda^m$.

We apply it to $Y=X_{\mathbf{s}},\lambda=2^{m-r}$, which yields: $P(X_{\mathbf{s}}>2^{m-r})\le 2^{-3m^2/4+m}$. Apply the union bound to obtain $P(\exists\mathbf{s}\mid X_{\mathbf{s}}>2^{m-r})\le 2^{-3m^2/4+m+r}$. This means exactly that with probability $\ge 1-2^{-3m^2/4+m+r}$ the most likely value of $\sigma(\mathbf{b})$ occurs with probability not more than $P(0)+2^{m-r}\le 2^{-r}(1+2^m)$. ∎

V. Concluding remarks

Note that, in exchange for a stronger requirement on $\mathbf{b}$ than the one in Theorem 3 of [1] (namely, in terms of min-entropy instead of Renyi's), we obtain a stronger result in some respects:

- a lower bound on $H_\infty(\sigma(\mathbf{b}))$, implying a fortiori one on $R(\sigma(\mathbf{b}))$;
- *very* strong concentration behavior for $H_\infty(\sigma(\mathbf{b}))$, that cannot be obtained by averaging arguments (Markov inequalities of order 1) alone.

Furthermore, Theorem 1 can be seen to extend to the case when Bob is subjected himself to a noisy channel. This has a natural application to biometry [3], where the "biometric noise" $\mathbf{b}$ represents in fact the traits of the user: in this context, the quantities $\mathbf{x}+\mathbf{b}$ for each user are stored in some database, and it should be impossible to extract from it any information on the secret $\mathbf{s}$ without explicit knowledge of the user's biometric data. The only natural statistical assumption on this sort of noise is additivity. Thus our results, valid irrespectively of the distribution of syndromes, can be put to use. One last point: by allowing some leeway in the choice of r, i.e. picking r such that $H(\mathbf{b})=r(1+\gamma)$ for some positive γ, we can refine the upper bounding of $E[X_{\mathbf{s}}^m]$ in Lemma 2 and get an improved bound on $H_\infty(\sigma(\mathbf{b}))$.

References

[1] C.H. Bennett, G. Brassard, C. Crepeau and U.M. Maurer, "Generalized privacy amplification", IEEE Trans. Inform.Th., vol. 41, 1915-1923 (1995).

[2] G. Cohen, I. Honkala, S. Litsyn and A. Lobstein, "Covering codes", North-Holland Mathematical Library 54 (1997).

[3] G. Cohen and G. Zémor, "Generalized coset schemes for the wire-tap channel: application to biometrics", ISIT 2004, Chicago, June 2004.

[4] T.M. Cover, "Broadcast channels", IEEE Trans. Inform.Th., vol. 18, 2-14 (1972).

[5] I. Csiszár and J. Körner "Information Theory", Academic Press (1982).

[6] I. Csiszár and J. Körner, "Broadcast channels with confidential messages", IEEE Trans. Inform.Th., vol. 24, 339-348 (1978).

[7] Y. Dodis, L. Reyzin and A. Smith "Fuzzy extractors and cryptography, or how to use your fingerprints", Eurocrypt 2004, LNCS 3027, 523-540 (2004).

[8] A. Juels and M. Wattenberg, "A fuzzy commitment scheme", in 6th ACM Conference on Computer and Communications Security, pp. 28–36, ACM Press, 1999.

[9] N. Nisan and A. Ta-Shma, Extracting randomness: a survey and new constructions, J. Computer and System Sciences 58(1), 148–173 (1999).

[10] N. Nisan and D. Zuckerman "Randomness is linear in space", J. Computer and System Sciences 52, 43-52 (1996).

[11] A. Wyner "The wire-tap channel", BSTJ 54 , 1355-1387 (1975).