

On Security Enhancement of Lightweight Encryption Employing Error Correction Coding and Simulators of Channels with Synchronization Errors

Miodrag MIHALJEVIC

**Mathematical Institute, Serbian Academy of Sciences and Arts
Belgrade, Serbia**

**- COST CRYPTACUS Workshop -
16-18 November 2017, Nijmegen - Netherlands**

Roadmap

- Introduction
- A Framework for Security Enhanced Encryption Based on Channels with Synchronization Errors
- Particular Instantiation
- Information-Theoretic Security Evaluation
- Computational-Complexity Security Evaluation
- Concluding Notes

1. Introduction

**Noisy Channels and Security
Enhancement of Encryption**

Security of Encryption and Implementation Complexity

- Mainly based on heuristic assumptions
- Particularly when the encryption is based on employment of finite state machines
- Lightweight encryption implies additional challenges ...
- Security enhancement appears as an interesting approach ...
- Asymmetric implementation complexity of encryption and decryption also appears as an interesting issue

Claude Shannon (1916-2001)



Two Key Papers

- **Information Theory**
- C. E. Shannon, “**A mathematical theory of communication**”. Bell System Technical Journal, vol. 27, pp. 379–423 and 623–656, July and October 1948.
- **Cryptology**
- C. E. Shannon “**Communication Theory of Secrecy Systems**”. Bell System Technical Journal, vol. 28 (4), pp. 656–715, 1949.

Noisy Channels

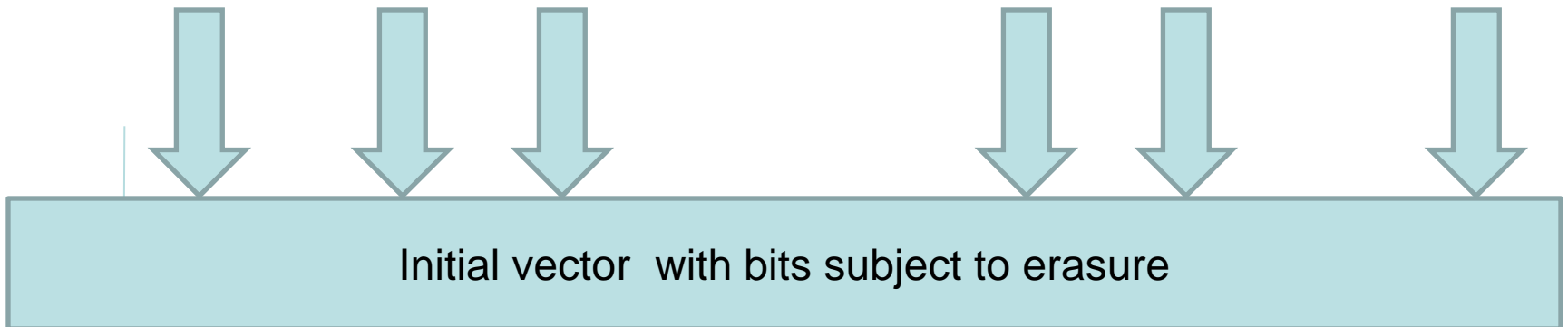
Channels with Additive Noise

- Erasure Channel
- Binary Symmetric Channel
- Gaussian Channel
- ...

Channels with Synchronization Noise

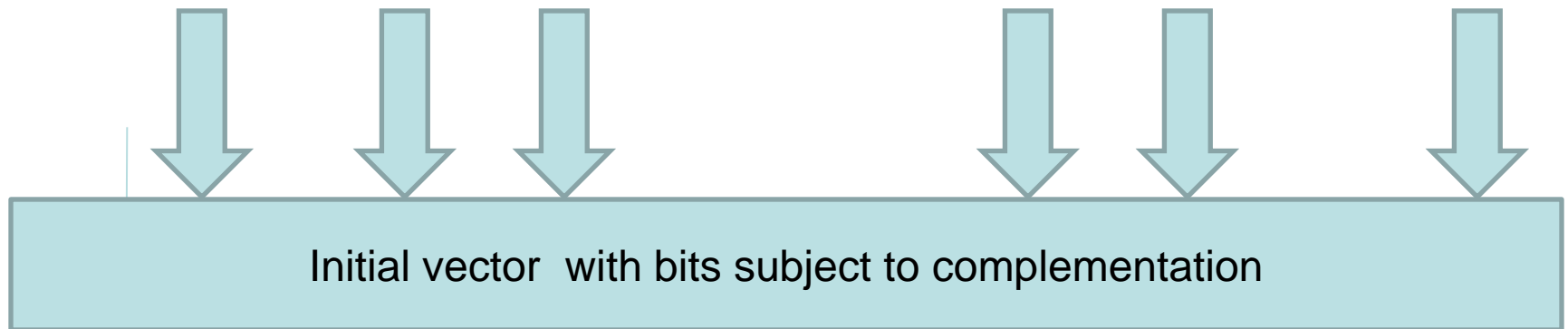
- Channels with insertion
- Channels with deletions
- Channels with Insertion, deletion and additive noise
- ...

Binary Channel with Random Erasures



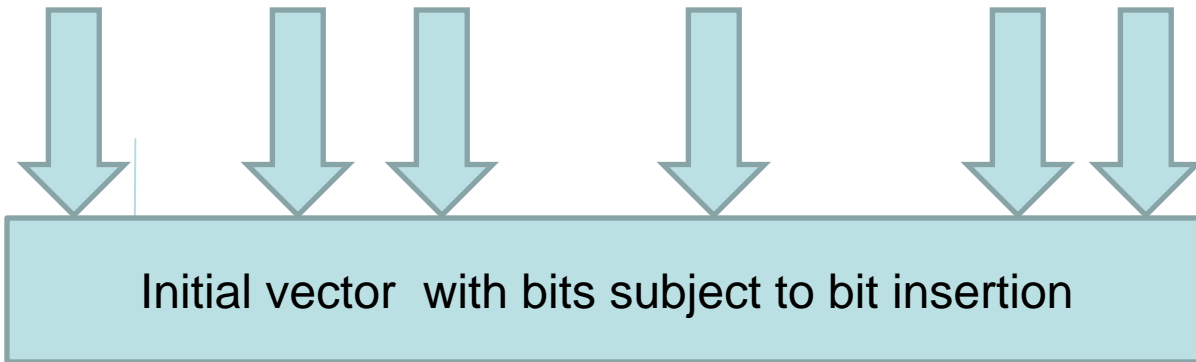
Erasure of bits is RANDOM – Positions of erased bits are KNOWN

Binary Channel with Random Bit Complementation (BSC Channel)



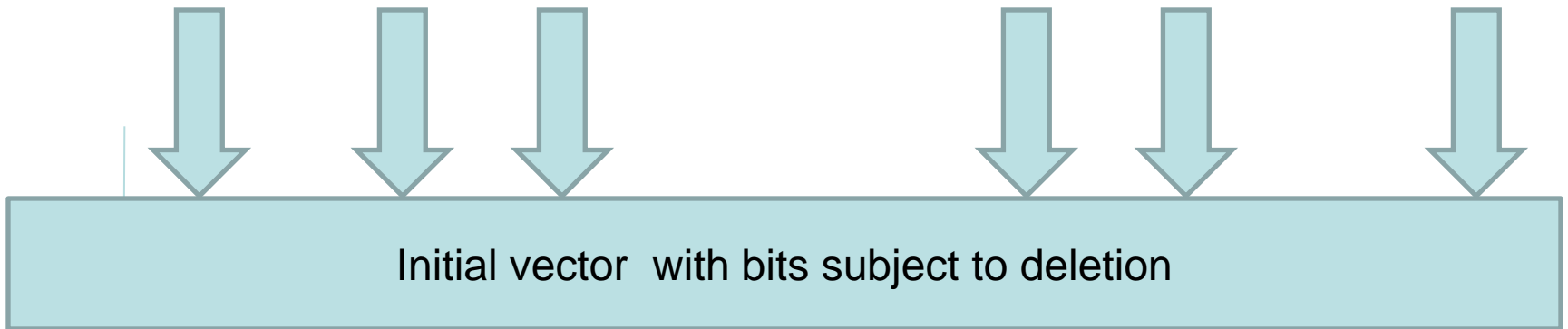
Complementation of bits is RANDOM – Positions of deleted bits are UNKNOWN

Binary Channel with Random Bit Insetion



Insertion of bits is RANDOM – Positions of insereted bits are UNKNOWN

Binary Channel with Random Bit Deletion

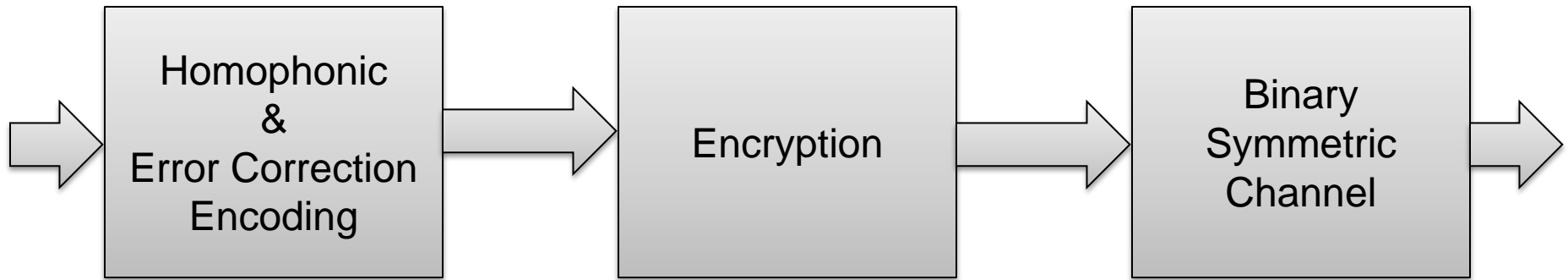


Shrunked vector after the chanel with random bits deletion

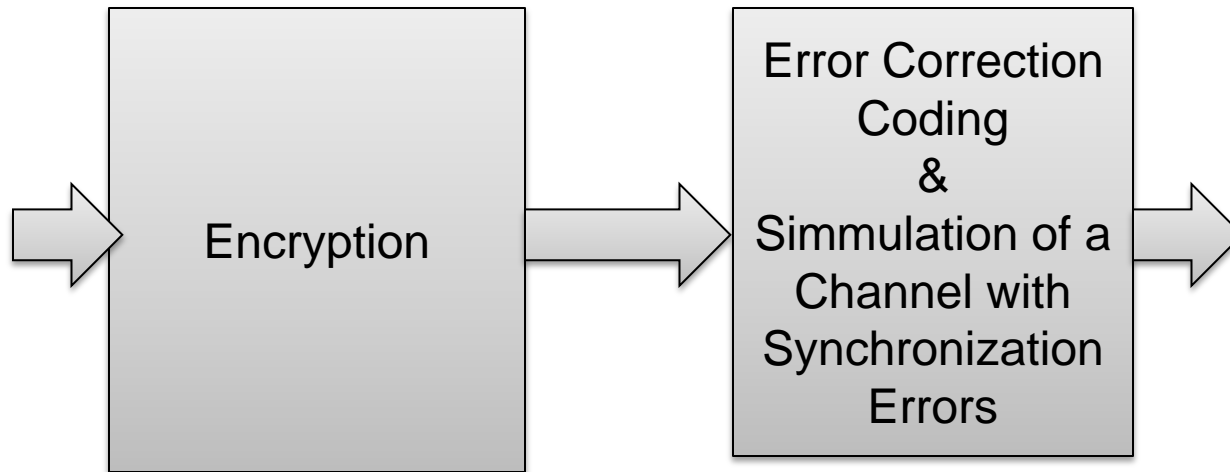
Deletion of bits is RANDOM – Positions of deleted bits are UNKNOWN

2. A Framework for Security Enhancement Based on the Channels with Synchronization Errors

Paradigm I

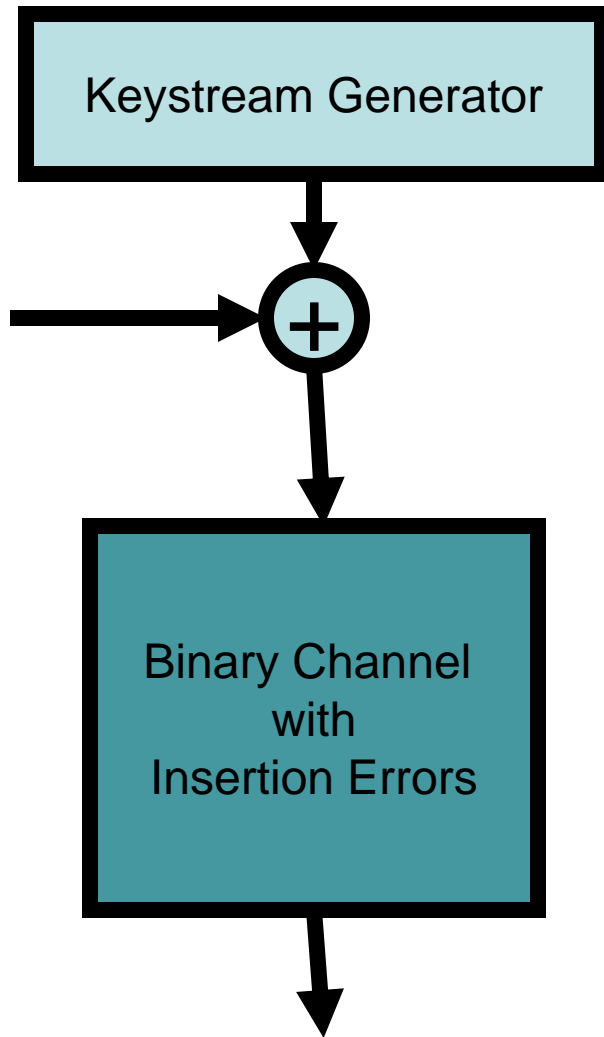


Paradigm II

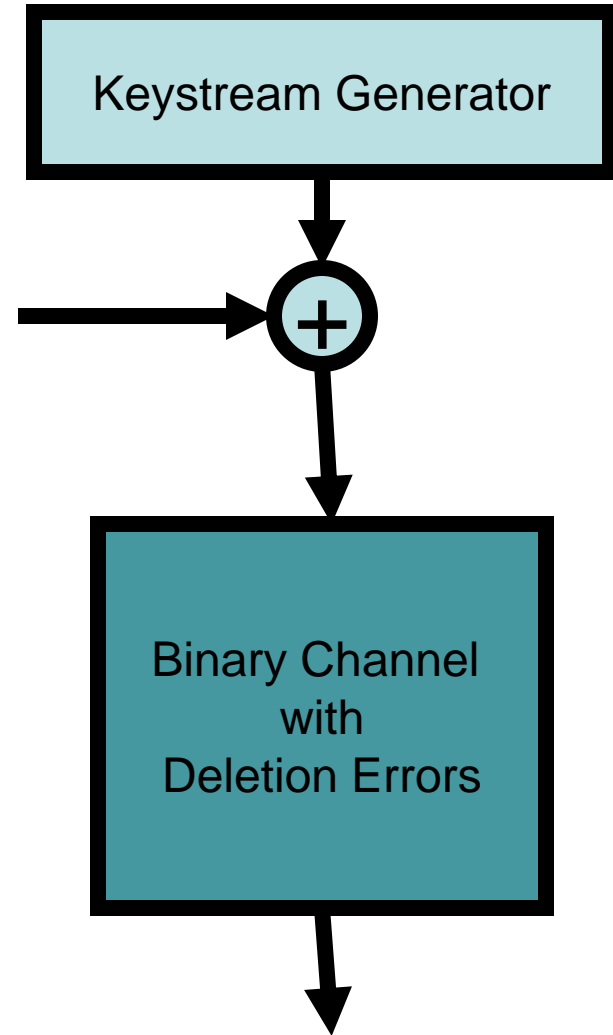


Desired Model of Encryption an Attacker Should Face

Encryption at Party I

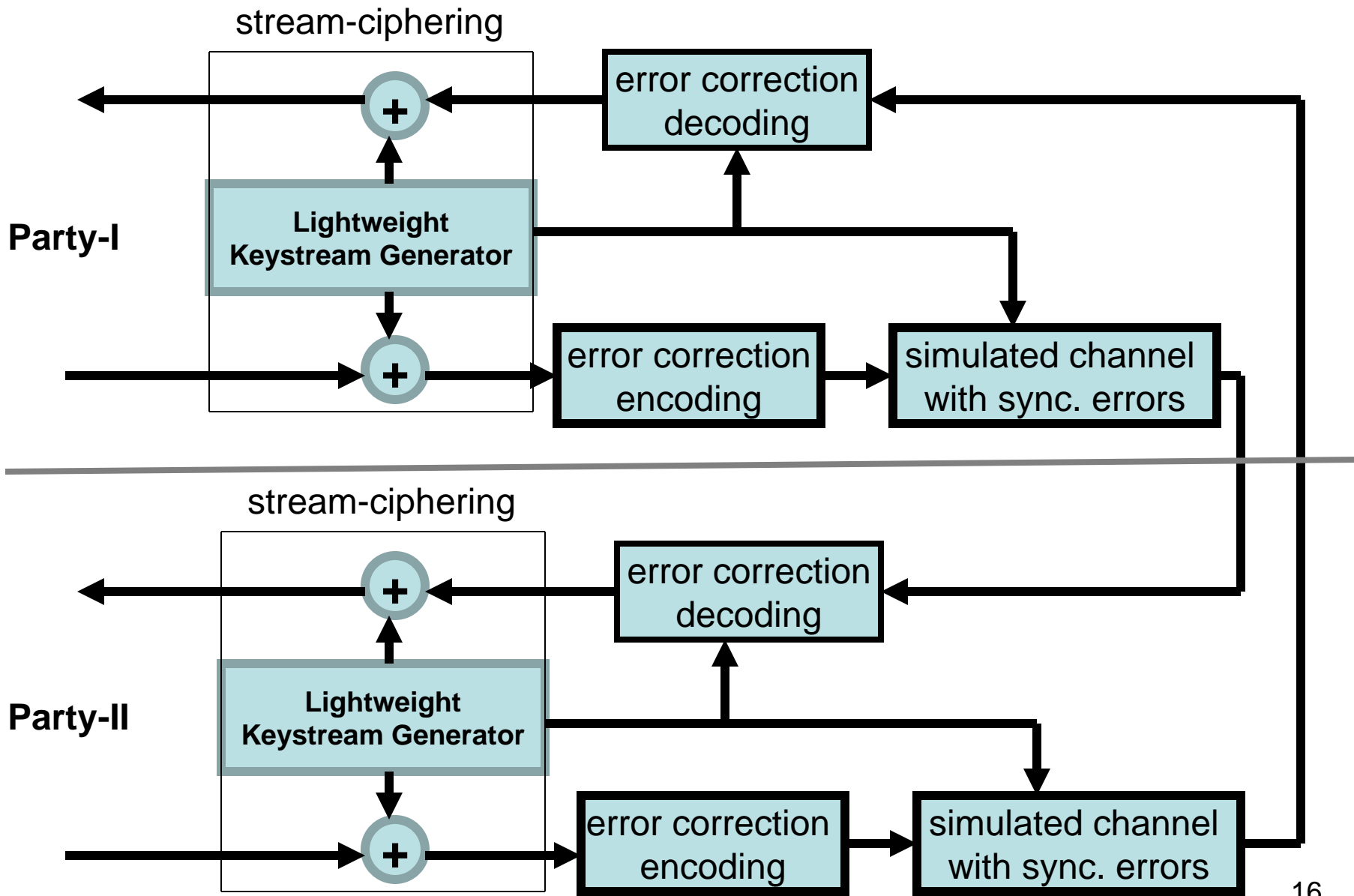


Encryption at Party II

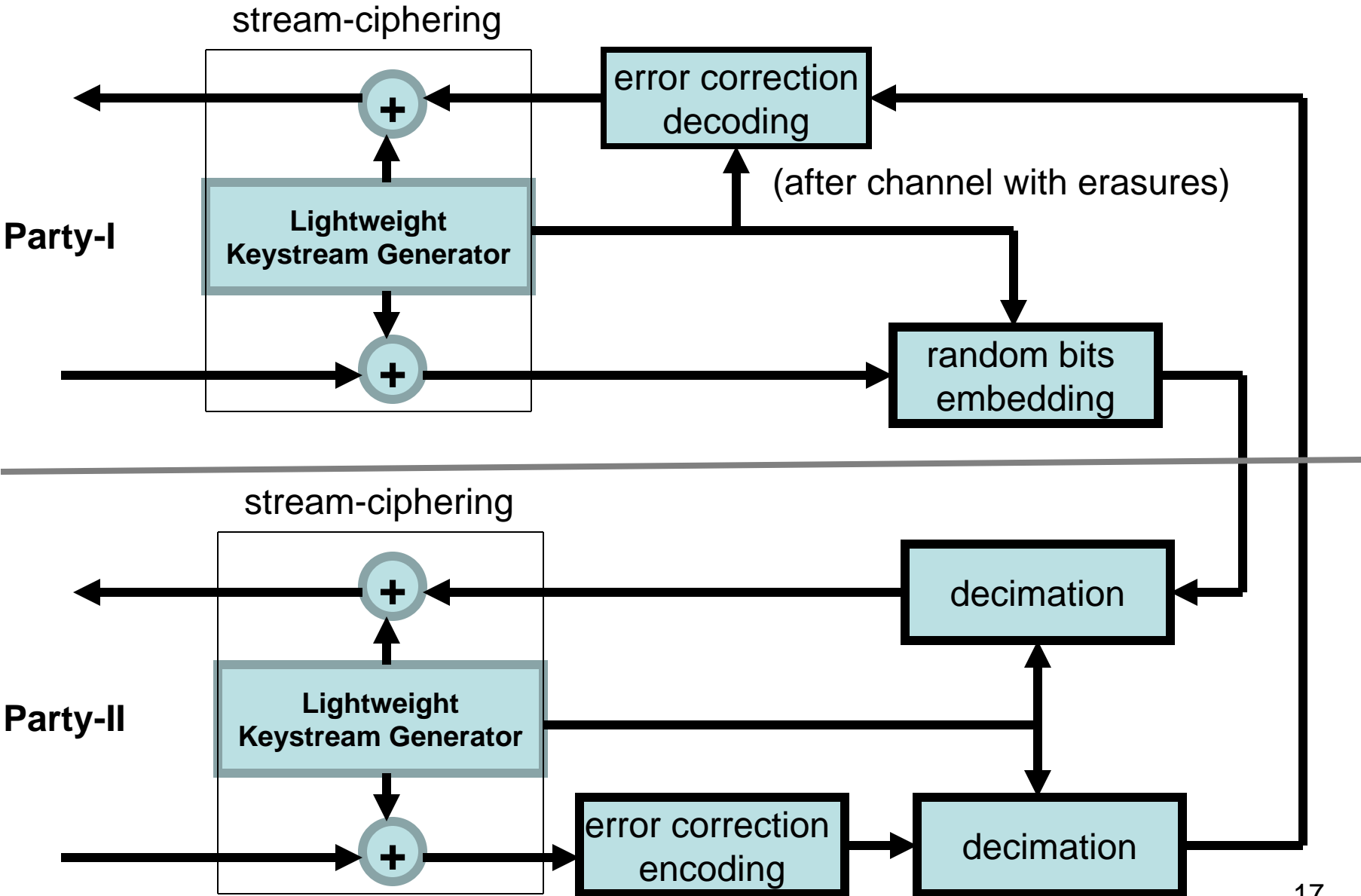


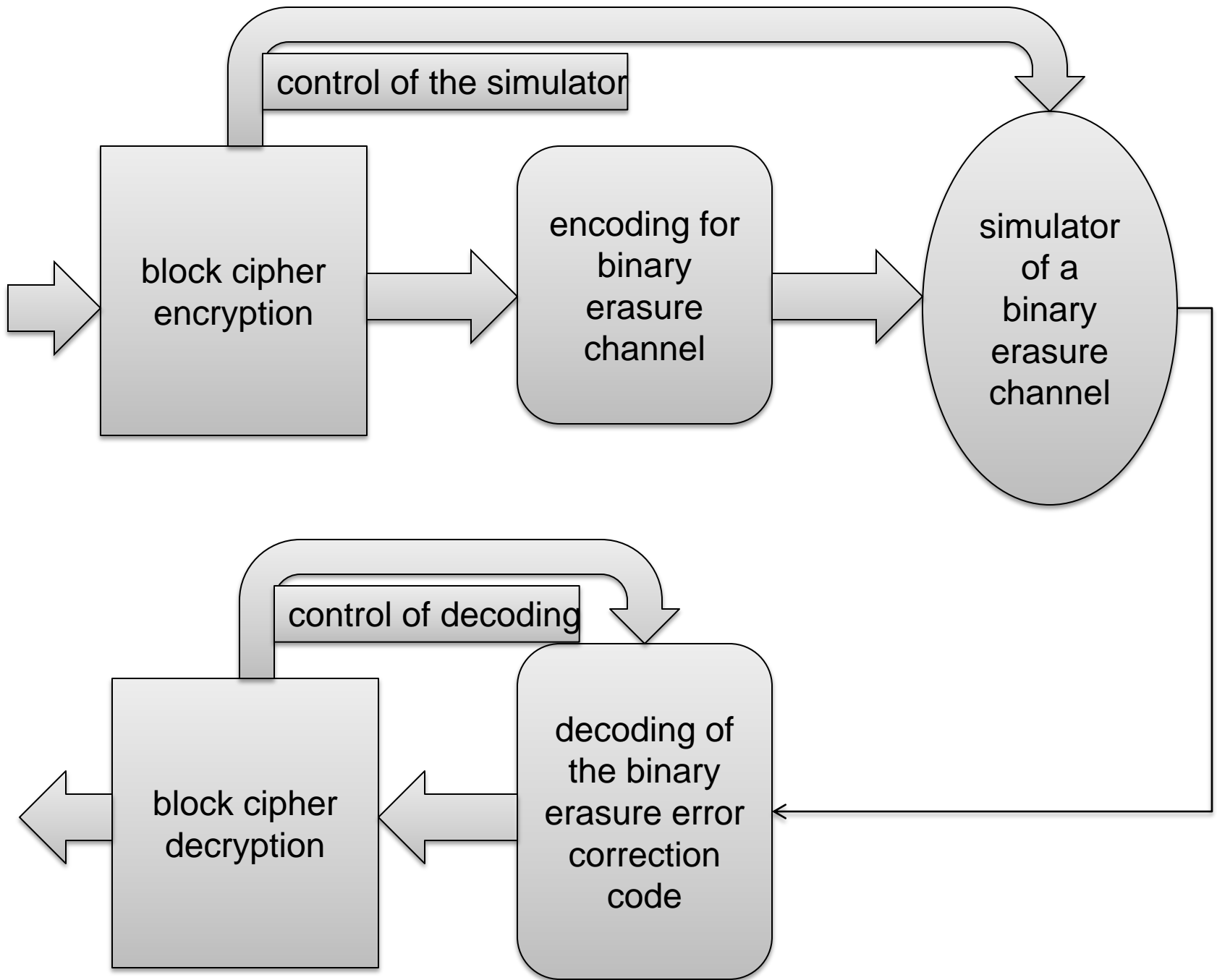
Attacker Side

A Framework for Encryption Based on Simulated Channels with Synchronization Errors

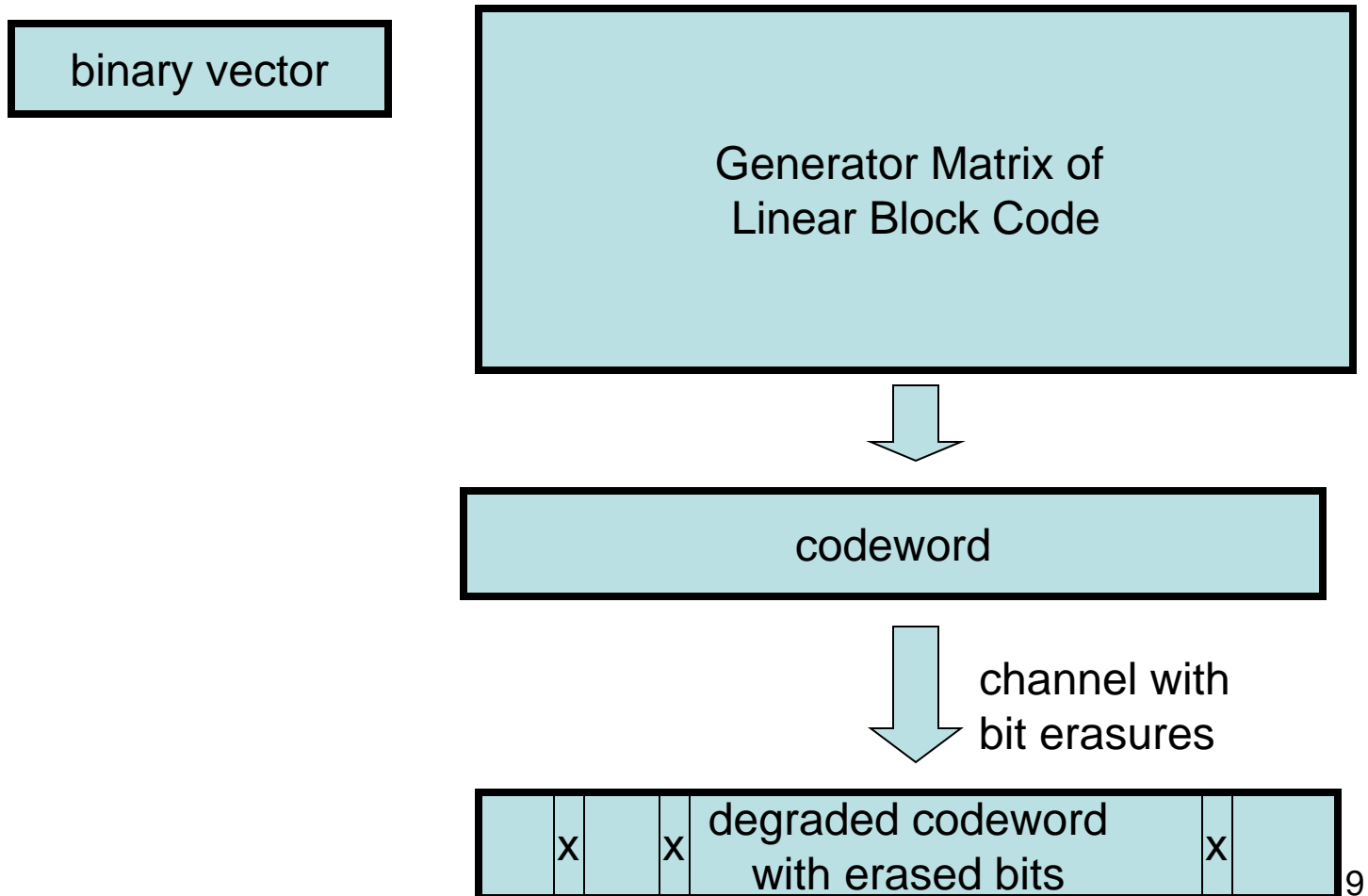


A Framework for Encryption with Asymmetric Implementation Complexity



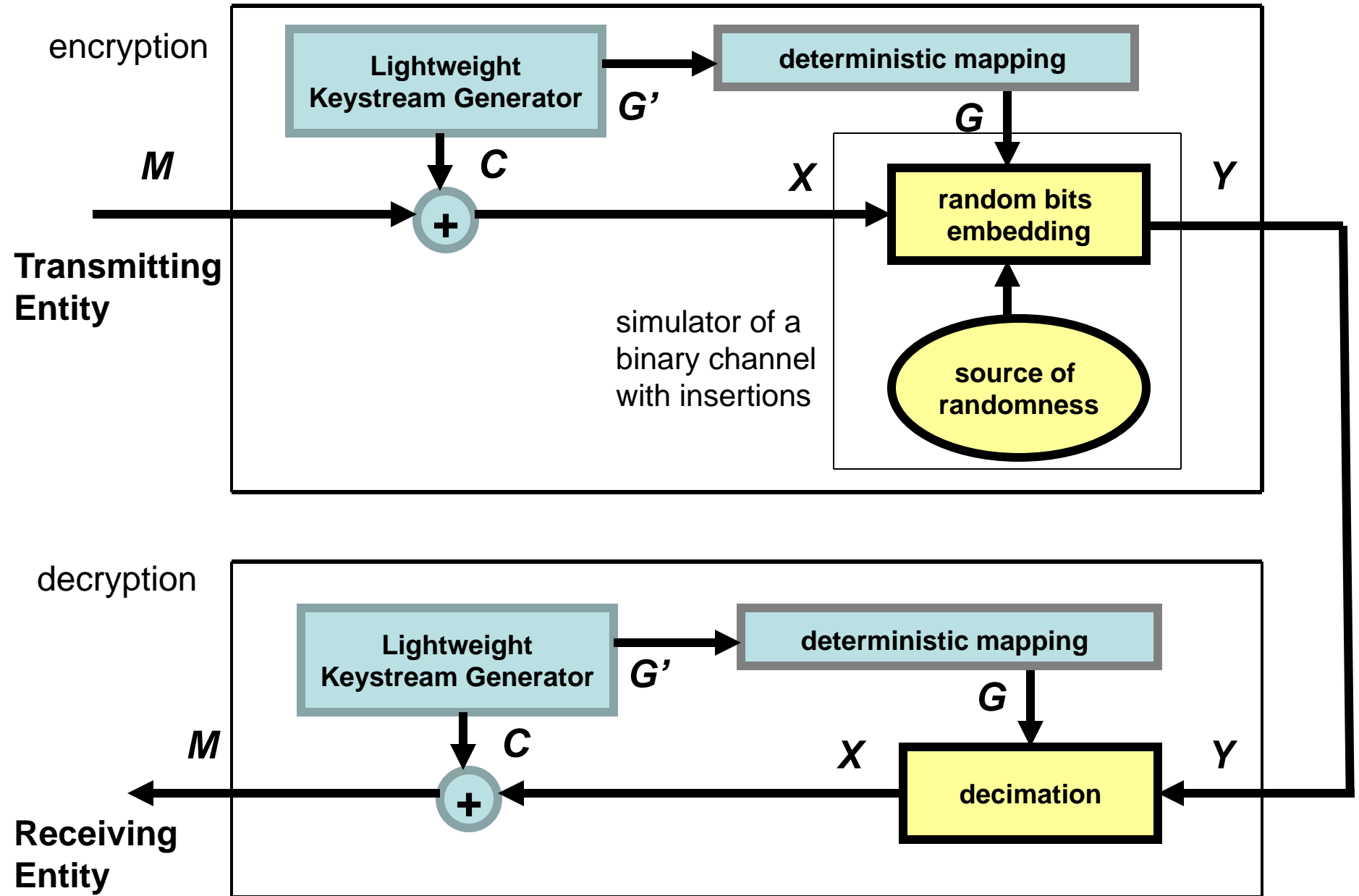


A Linear Binary Block Code Encoding Paradigm



3. Particular Instantiation Under Security Evaluation

A Framework for Encryption and Decryption with Asymmetric Implementation Complexity



4. Information-Theoretic Security Evaluation

Preliminaries

Eve (the eavesdropper) and Bob (the intended receiver) both receive the string $Y^{(n)}$ containing the randomly inserted symbols. The eavesdropper, not having access to the shared source of randomness G^n , cannot easily parse the string $Y^{(n)}$ to recover X^n . The intended receiver, on the other hand, has access to G^n , and since G_k represents the length of the inserted string between any two symbols X_k and X_{k+1} , the intended receiver (Bob) can easily remove the inserted symbols \underline{B}_k from $Y^{(n)}$ (i.e., decimate $Y^{(n)}$) to recover X^n . In other words, by sharing the source of randomness G^n .

The sequence C^n is a pseudo-random sequence, but for the purpose of computing information-theoretic quantities, we assume that C^n is modeled to be statistically indistinguishable from a sequence of iid Bernoulli- $\frac{1}{2}$ random variables.

The information-theoretic quantity of interest is the *iud information rate* defined as the information rate between X^n and $Y^{(n)}$ when the symbols X_k are independent and uniformly distributed (iud)

$$\mathcal{I}_{\text{iud}}(X; Y) \triangleq \lim_{n \rightarrow \infty} \frac{1}{n} I(X^n; Y^{(n)}) \Big|_{p(x^n) = 2^{-n}}.$$

The information rate $\mathcal{I}_{\text{iud}}(X; Y)$ represents the amount of information that the eavesdropper can “*learn*”, on average, about X after observing Y . The information rate $\mathcal{I}_{\text{iud}}(X; Y)$ is not computable in closed-form, but is attainable using Monte-Carlo techniques.

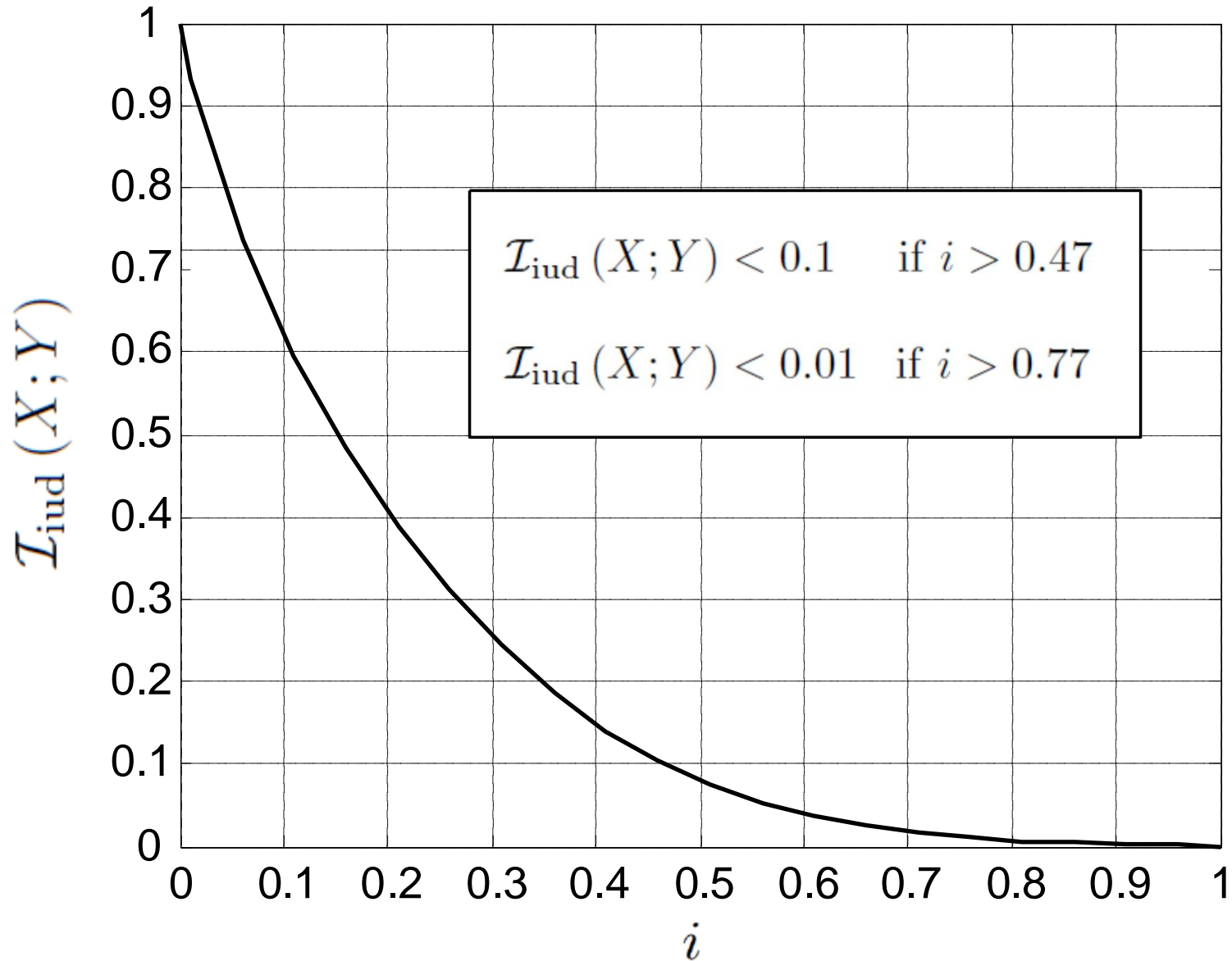
The information rate $\mathcal{I}_{\text{iud}}(X; Y)$ is not computable in closed-form, but is attainable using Monte-Carlo techniques. For example, known bounds are

$$\begin{aligned} \mathcal{I}_{\text{iud}}(X; Y) &\geq \frac{1}{n} I(X^n; Y^{(n)}) \Big|_{p(x^n)=2^{-n}} - \frac{1}{n} H(\mathcal{L}(Y^{(n)})) \\ \mathcal{I}_{\text{iud}}(X; Y) &\leq \frac{1}{n} I(X^n; Y^{(n)}) \Big|_{p(x^n)=2^{-n}}. \end{aligned} \quad (2)$$

For large n , the correction term $\frac{1}{n} H(\mathcal{L}(Y^{(n)}))$ equals

$$\frac{1}{n} H(\mathcal{L}(Y^{(n)})) = \frac{1}{2n} \log_2 \left(\frac{2\pi e \cdot i \cdot n}{(1-i)^2} \right) + O(n^{-2}). \quad (3)$$

Illustrative Numerical Example



We already established that learning X after observing Y is extremely unfavorable for the eavesdropper because the information rate $\mathcal{I}_{\text{iud}}(X; Y)$ is low for large insertion probabilities i . However, the eavesdropper may adopt a strategy in which she first attempts to *learn* the sequence G^n , and then attempt to crack X^n . To study the effects of this strategy, let us define the following quantities:

$$\begin{aligned} \mathcal{I}_{\text{iud}}(G; Y) &\triangleq \lim_{n \rightarrow \infty} \frac{1}{n} I(G^n; Y^{(n)}) \Big|_{p(x^n)=2^{-n}} \\ \mathcal{I}_{\text{iud}}(X, G; Y) &\triangleq \lim_{n \rightarrow \infty} \frac{1}{n} I(X^n, G^n; Y^{(n)}) \Big|_{p(x^n)=2^{-n}} \\ \mathcal{I}_{\text{iud}}(X; Y | G) &\triangleq \lim_{n \rightarrow \infty} \frac{1}{n} I(X^n; Y^{(n)} | G^n) \Big|_{p(x^n)=2^{-n}} \\ \mathcal{I}_{\text{iud}}(G; Y | X) &\triangleq \lim_{n \rightarrow \infty} \frac{1}{n} I(G^n; Y^{(n)} | X^n) \Big|_{p(x^n)=2^{-n}} \end{aligned}$$

Proposition 1:

$$\mathcal{I}_{\text{iud}}(G; Y) = 0 \quad (1)$$

$$\mathcal{I}_{\text{iud}}(X; Y | G) = 1 \quad (2)$$

$$\mathcal{I}_{\text{iud}}(X, G; Y) = 1 \quad (3)$$

$$\mathcal{I}_{\text{iud}}(G; Y | X) = 1 - \mathcal{I}_{\text{iud}}(X; Y). \quad (4)$$

5. Computational Complexity Security Evaluation

Definition 1: The Adversarial Indistinguishability Experiment consists of the following steps:

1. The adversary \mathcal{A} chooses a pair of messages $(\mathbf{m}_0; \mathbf{m}_1)$ of the same length n , and passes them on to the encryption system for encrypting.
2. A bit $b \in \{0,1\}$ is chosen uniformly at random, and only one of the two messages $(\mathbf{m}_0; \mathbf{m}_1)$, precisely \mathbf{m}_b , is encrypted into ciphertext $\text{Enc}(\mathbf{m}_b)$ and returned to \mathcal{A} ;
3. Upon observing $\text{Enc}(\mathbf{m}_b)$, and without knowledge of b , the adversary \mathcal{A} outputs a bit b_0 ;
4. The experiment output is defined to be 1 if $b_0 = b$, and 0 otherwise; if the experiment output is 1, denoted shortly as the event $(\mathcal{A} \rightarrow 1)$, we say that \mathcal{A} has succeeded.

Definition 2. An encryption scheme provides indistinguishable encryptions in the presence of an eavesdropper, if for all probabilistic polynomial-time adversaries \mathcal{A}

$$\Pr[\mathcal{A} \rightarrow 1 | \text{Enc}(\mathbf{m}_b)] \leq \frac{1}{2} + \epsilon ,$$

where $\epsilon = \text{negl}(n)$ is a negligibly small function.

Proposition 2: Let the encrypted mapping of M^n into X^n be such that $\frac{1}{2} + \epsilon$ equals the advantage of the adversary \mathcal{A} (specified by Definition 2) to win the indistinguishability game (specified by Definition 1), and let the mutual information $\mathcal{I}_{iud}(X; Y)$ be known. Under these assumptions, for large n ,

$$\Pr[\mathcal{A} \rightarrow 1 | Y^{(n)} = \mathbf{y}] = \frac{1}{2} + \epsilon \cdot \delta, \quad \text{where} \quad (1)$$

$$\delta \triangleq \Pr(X^n = \mathbf{x}_b | Y^{(n)} = \mathbf{y}) < \frac{1}{n} + \frac{1}{n} I(X^n, Y^{(n)}) \Big|_{p(x^n) = 2^{-n}} \quad (2)$$

Theorem 1: Let the encrypted mapping of M^n into X^n be such that $\frac{1}{2} + \epsilon$ equals the advantage of the adversary \mathcal{A} (specified by Definition 2) to win the indistinguishability game (specified by Definition 1), and let the mutual information $\mathcal{I}_{iud}(X; Y)$ be known. Under these assumptions, for large n ,

$$\Pr[\mathcal{A} \rightarrow 1 | Y^{(n)} = \mathbf{y}] = \frac{1}{2} + \epsilon \cdot \delta, \quad \text{where} \quad (1)$$

$$\delta < \mathcal{I}_{iud}(X; Y) + \frac{\log_2 \left[\frac{8\pi e \cdot i \cdot n}{(1-i)^2} \right]}{2n} + O(n^{-2}). \quad (2)$$

Y. Liron and M. Langberg, “A Characterization of the Number of Subsequences Obtained via the Deletion Channel”, *IEEE Transactions on Information Theory*, vol. 61, no. 5, pp. 2300-2312, May 2015.

Let $D_t(\mathbf{Z})$ be a set of subsequences of \mathbf{Z} that can be obtained from \mathbf{Z} after t deletions. A family of strings, named unbalanced strings has been defined. A string is called unbalanced, if all of the runs of symbols in the string are of length 1, except for one run. Let $U_{\ell,r}^{(i)}$ be a binary string of length ℓ with r runs, in which all runs are of length 1, except for the i -th run which is of length $\ell - r + 1$. Due to symmetry $|D_t(U_{\ell,r}^{(1)})| = |D_t(U_{\ell,r}^{(r)})|$, and consequently define

$$u(\ell, r, t) = |D_t(U_{\ell,r}^{(1)})| = |D_t(U_{\ell,r}^{(r)})|. \quad (1)$$

It has been shown that these extreme cases have the least number of subsequences among the unbalanced strings, as well as that they have the least amount of subsequences among all strings.

Y. Liron and M. Langberg, “A Characterization of the Number of Subsequences Obtained via the Deletion Channel”, *IEEE Transactions on Information Theory*, vol. 61, no. 5, pp. 2300-2312, May 2015.

Theorem (Closed-Form Formula for $u(\ell, r, t)$):

For all $t < \ell$, $2 < r \leq \ell$,

(i) when $r > t$:

$$u(\ell, r, t) = d(r, t) + \sum_{i=t+r-\ell-1}^{t-2} d(r-2, i), \quad (1)$$

(ii) when $r \leq t$:

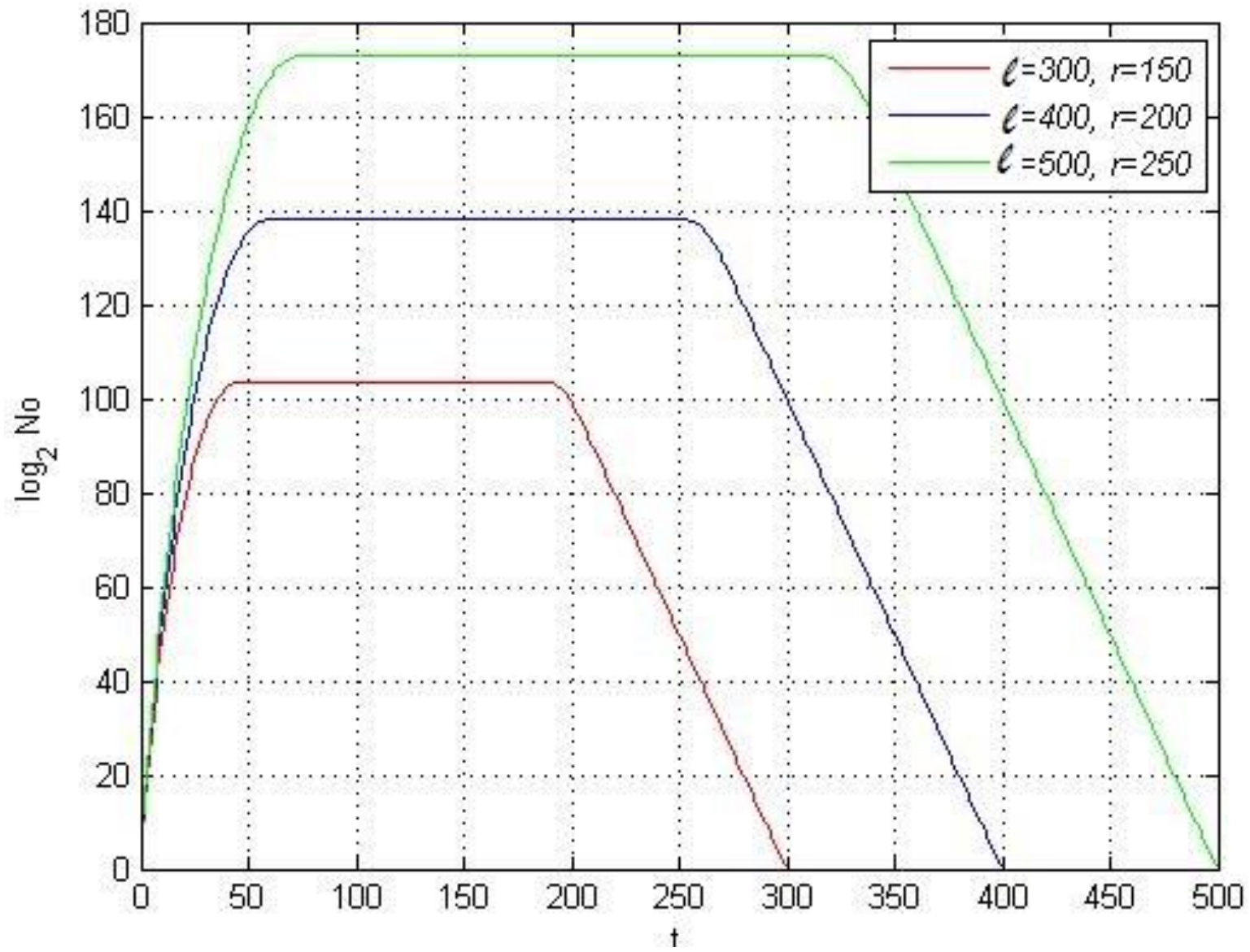
$$u(\ell, r, t) = 2 + \sum_{i=t+r-\ell-1}^{r-3} d(r-2, i), \quad (2)$$

where

$$d(r, i) = |D_i(\mathbf{Z}_r^C)| = \sum_{j=0}^i \binom{r-i}{j} \quad (3)$$

assuming that $d(r, 0) = 1$, and for $i < 0$, $d(r, i) = 0$, and that the following conventions are employed:

$$\sum_{i=j}^k a_i = 0 \text{ when } j > k, \quad (4)$$



Theorem 3. Assuming that the employed keystream generator is such that the following is valid:

$$I(\mathbf{M}; \mathbf{C}) = 0, \quad I(\mathbf{M}; \mathbf{G}) = 0, \quad I(\mathbf{C}; \mathbf{G}) = 0, \quad (1)$$

and

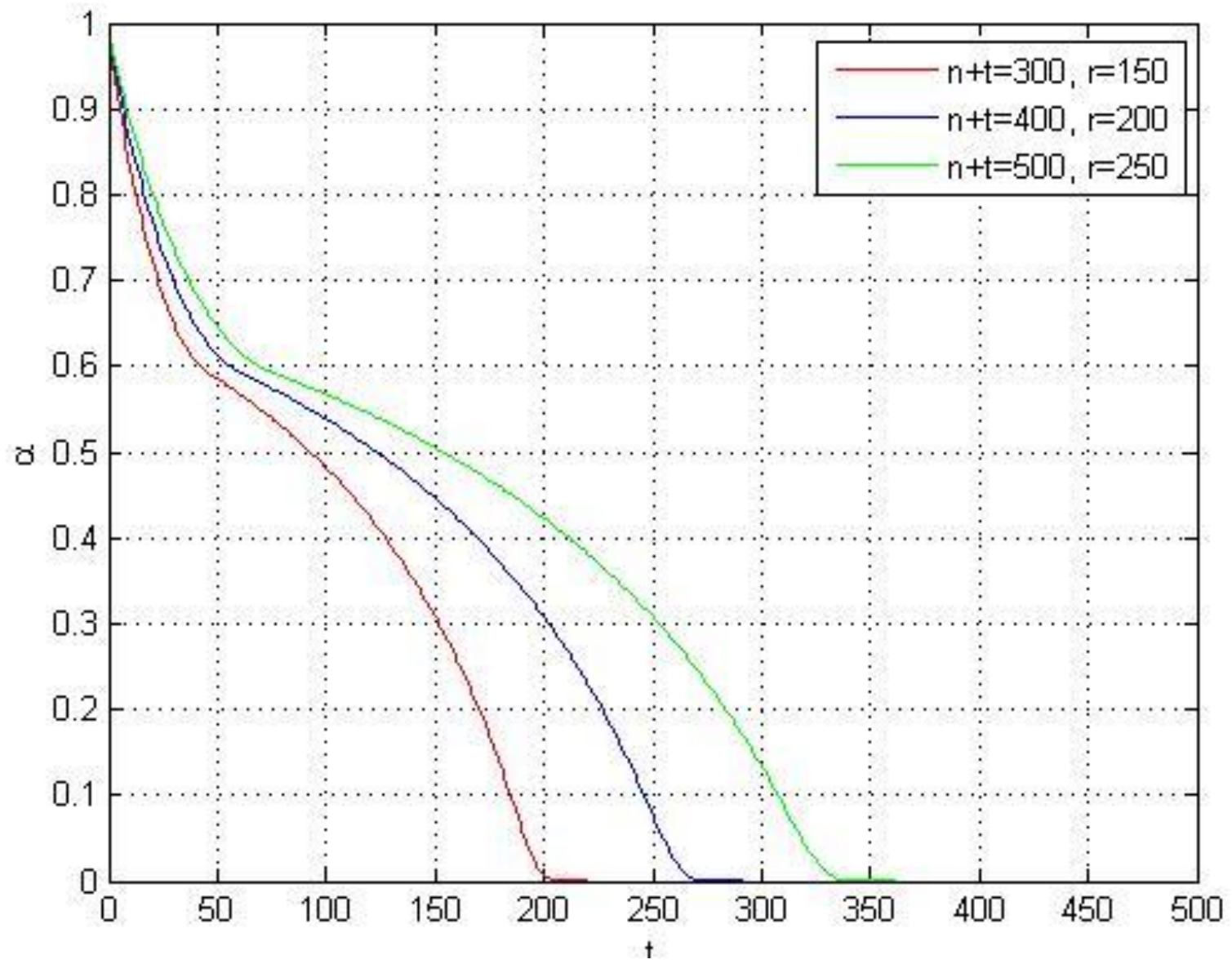
$$I(\mathbf{M}; \mathbf{X}) \leq \epsilon. \quad (2)$$

the simulator of binary channel with random insertions provides

$$\frac{1}{n} I(\mathbf{M}; \mathbf{Y}) \leq \frac{\alpha \cdot \epsilon}{n}, \quad (3)$$

$$\alpha = 1 - \frac{1}{n} \log_2(u(n + t, r, t)), \quad (4)$$

where $u(n + t, r, t)$ is number of certain equally likely subsequences.



Concluding Notes

Main Messages

- A paradigm for the security enhancement based on results regarding channels with synchronization errors has been proposed
- Evaluation of the security enhancement has been discussed from the information-theoretic and computational complexity points of view

Thank You Very Much for the
Attention,

and
QUESTIONS Please!