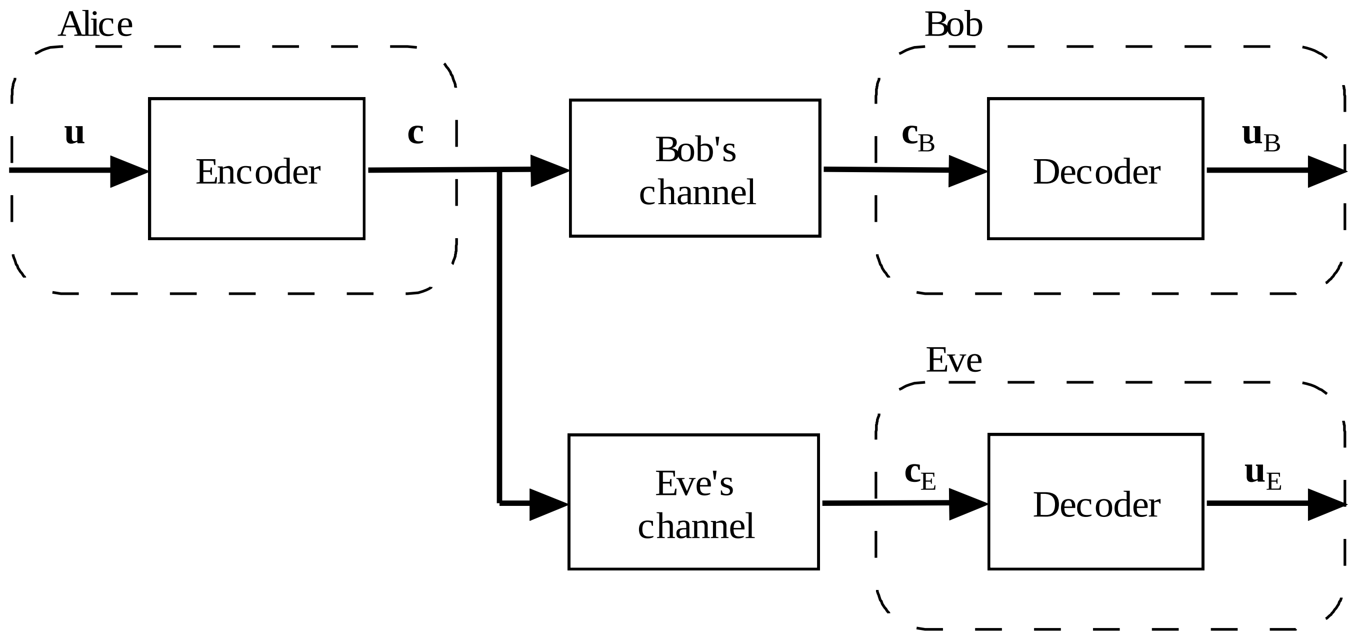
Achieving Semantic Security Without Keys Through Coding And All-Or-Nothing Transforms Over Wireless Channels

Marco Baldi, **Linda Senigagliesi**, Franco Chiaraluce
DII, Università Politecnica delle Marche,
Ancona, Italy
Email: l.senigagliesi@pm.univpm.it

Physical Layer Security



➢ Wiretap channel [1]

- Setting which exploits the differences between Bob’s and Eve’s channel
- Attacker: passive eavesdropper

[1] A. D. Wyner, “The wire-tap channel,” *Bell Syst. Tech. J.*, vol. 54, no. 8, pp. 1355–1387, Oct. 1975.

Security Metrics

Practical systems

Physical Layer Security ⟺ **?** ⟺ **Computational Security**

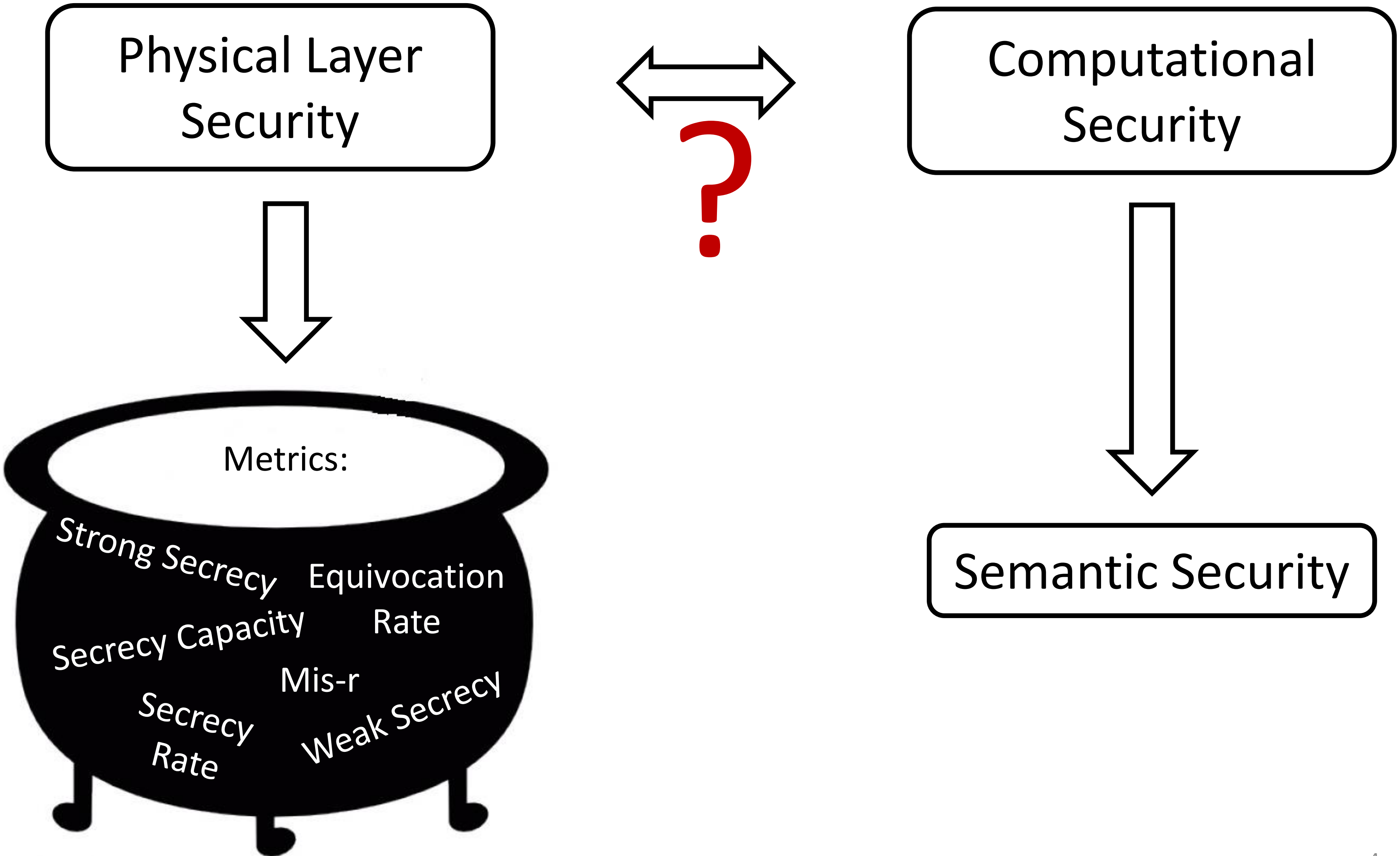
Physical Layer Security

- The wiretap channel
- Key distillation
- …

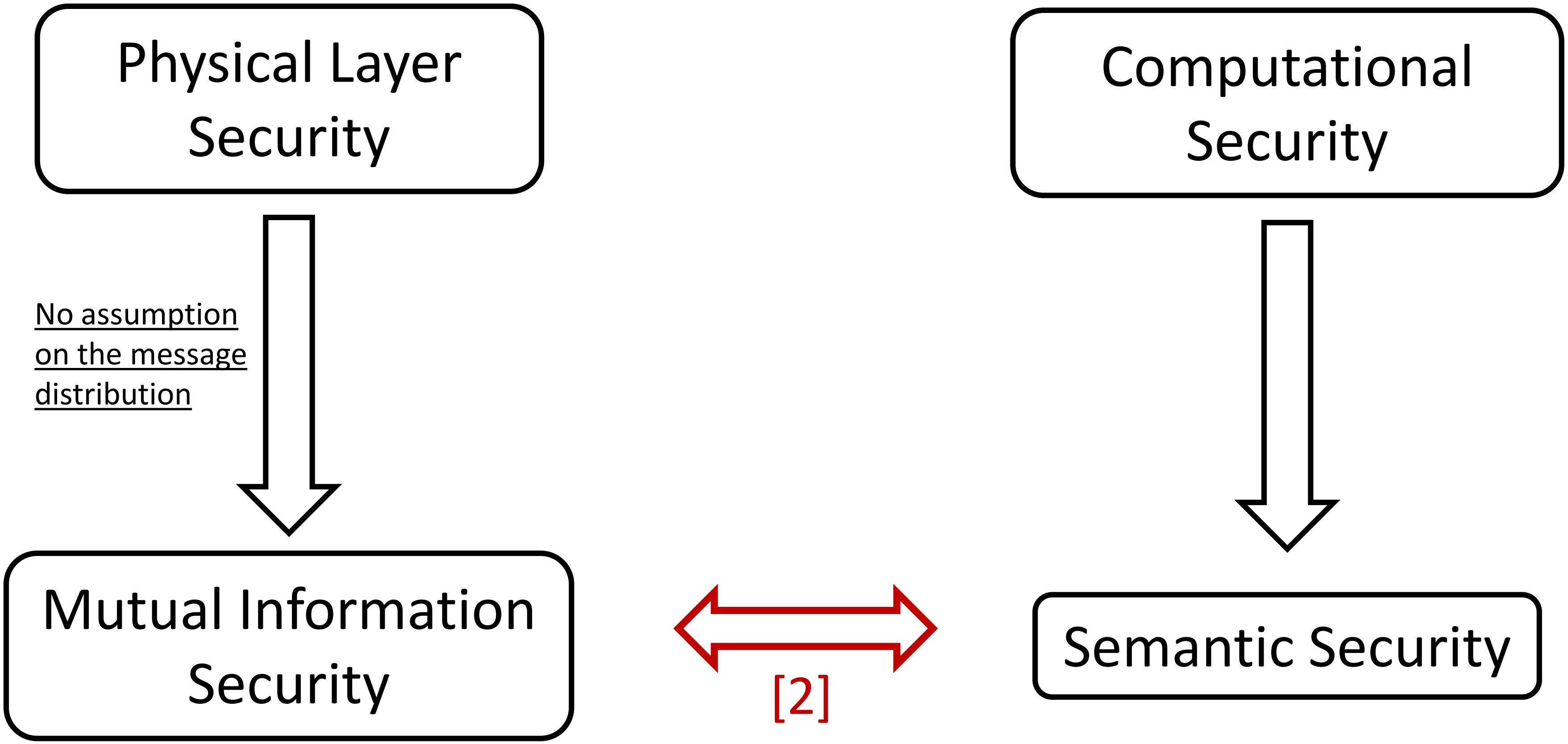
Computational Security

- Cryptographic techniques

Security Metrics



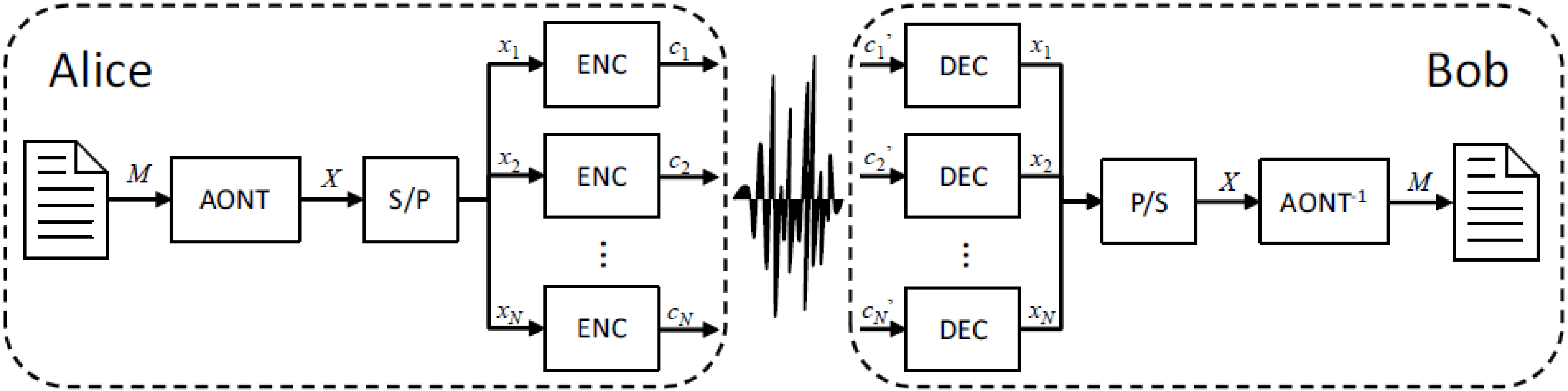
Security Metrics



[2] M. Bellare, S. Tessaro, and A. Vardy, “Semantic security for the wiretap channel,” in *Advances in Cryptology - CRYPTO 2012*, R. Savafi-Naini and R Canetti, Eds. 2012, vol. 7417 of *LNCS*, pp. 294–311, Springer Berlin Heidelberg.

The Protocol

Alice wishes to securely transmit a document M to Bob over a wireless channel:



1. **Encryption**: Alice transforms M into X through an AONT;
   M is padded with random bits before entering the AONT until the length of the output X reaches L (which must take into account the expansion due to the AONT and be a multiple of k);
2. **Slicing**: X is split into N blocks of k bits each;
3. **Coding**: Each block is then encoded through a binary linear block code $C(n, k)$, where n denotes the code length and k is the code dimension.

→ We use short codes and high order modulations

All-Or-Nothing Transform[3]

- Random-like transformation infeasible to invert, even in part, unless the transformed data is completely available

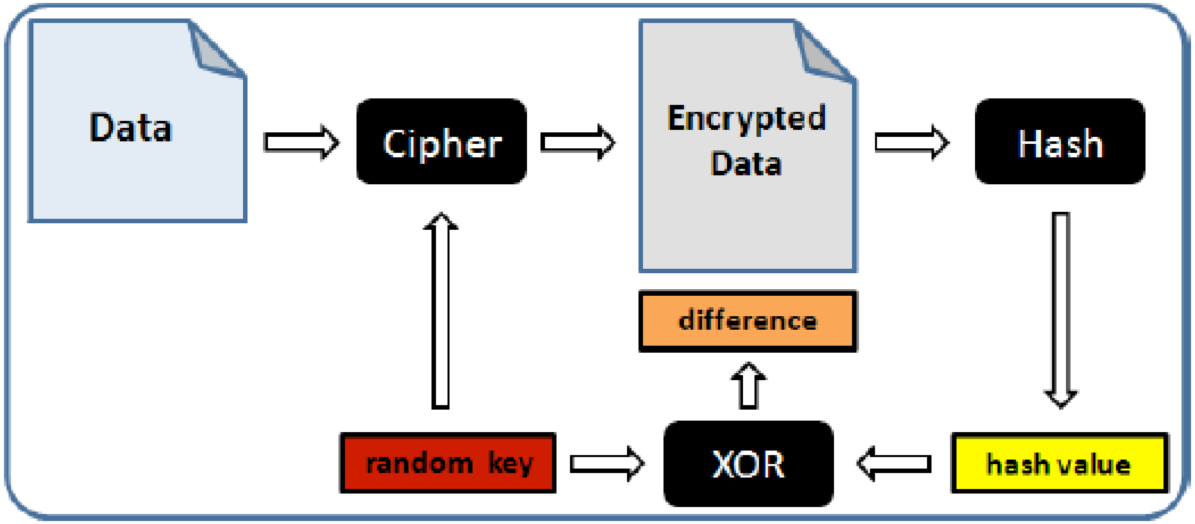
- An AONT-processed message cannot be recovered if some part of it is lost during transmission

⇒ Unconditional Security [4]

[3] R. L. Rivest, "All-or-nothing encryption and the package transform," in *Fast Software Encryption*. 1997, vol. 1267 of *LNCS*, pp. 210–218, Springer Berlin Heidelberg.
[4] D. R. Stinson, "Something about all or nothing (transforms)," *Designs, Codes and Cryptography*, vol. 22, pp. 133–138, Mar. 2001.

All-Or-Nothing Transform

- The message is divided into blocks
- A random-generated key K is embedded with the data exploiting a key-based encryption algorithm (for example AES-256)
  - No need of pre-shared keys
- The final codeword is computed as the XOR between K and the hash digest of the other codewords, and it is appended to the message

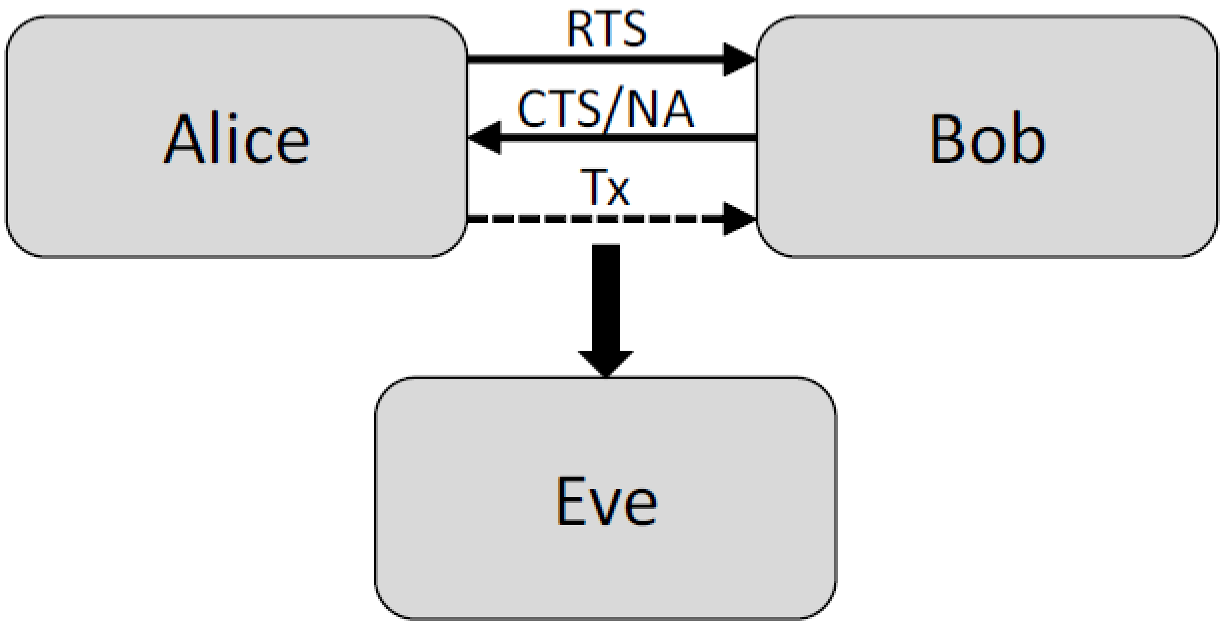


[4]

[4] J. Resch and J. Plank, "AONT-RS: Blending security and performance in dispersed storage systems," in *Proc. 9th USENIX Conference on File and Storage Technologies (FAST)*, San Jose, USA, Feb. 2011.

The Protocol (2)

Three-way communication:



1) RTS (Request To Send)
2) CTS (Clear To Send) or NA (Not Available), depending on channel quality
3) Tx (Transmission), only upon CTS

Hp: The channel does not vary during any exchange of a pair of RTS-CTS/NA messages between Alice and Bob and the possible subsequent transmission of a codeword.

Wiretapper's Equivocation

- **Equivocation as a metric:** it allows to obtain a lower bound on the size of a list that Eve can reliably limit the message to.

- Definition: $s = H(\mathbf{c}|\mathbf{c}_E) = H(\mathbf{c}) - I(\mathbf{c};\mathbf{c}_E)$

$$H(\mathbf{c}) = k' \leq k$$

$$I(\mathbf{c};\mathbf{c}_E) \leq \frac{n}{q} C_E$$

no assumption on the message distribution

Eve's equivocation depends on the message entropy and on the mutual information between the message and Eve's observation.

- Lower bound:

$$s = n\bar{R}_e \geq n\left[R_h - \frac{C_E}{q}\right]^+ = \tilde{s}$$

Equivocation rate: $R_e = \frac{s}{n}$	Source entropy rate: $R_h = \frac{k'}{n} \leq \frac{k}{n} = R_c$
Eve's channel capacity: C_E	Number of bits per transmitted symbol: q

Wiretapper's Equivocation

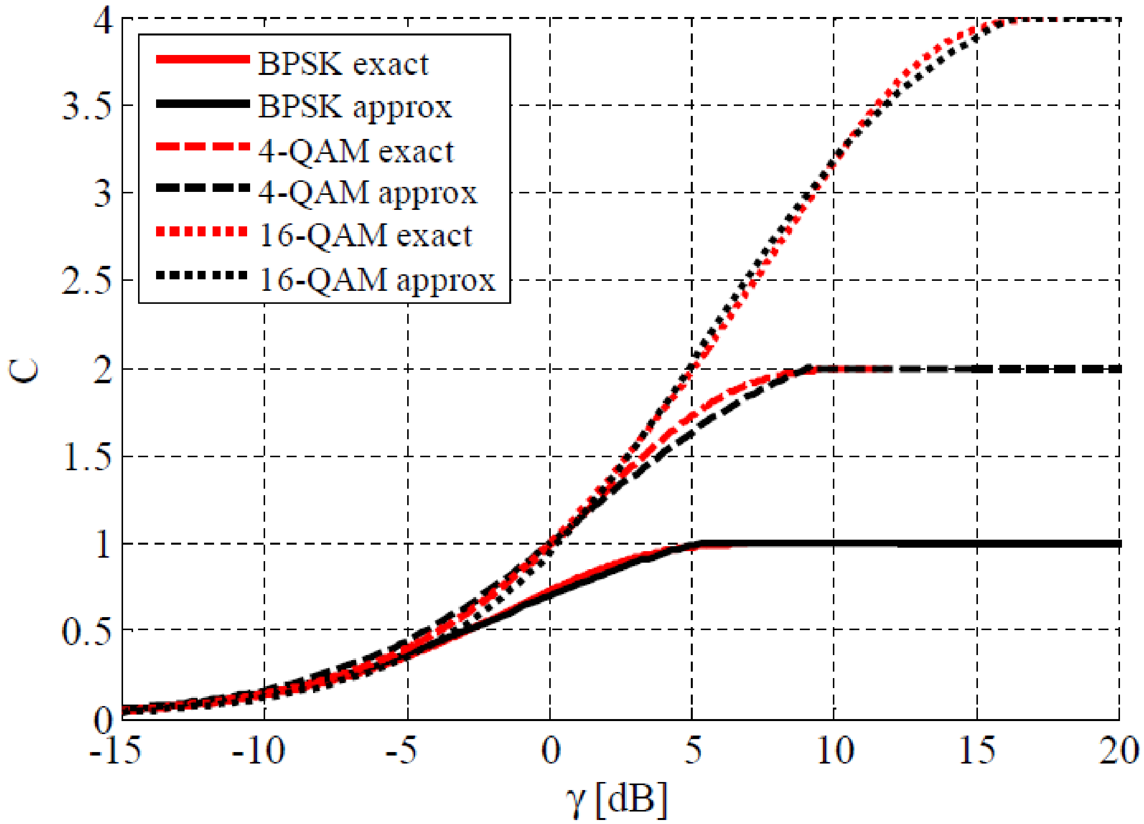
- Perfect secrecy:

$$s = k'$$

- If $0 < s < k'$, perfect secrecy is not achievable, but Eve still needs to perform 2^s attempts on average in order to correctly decode $\boldsymbol{c}$ from $\boldsymbol{c}_E$

Approximate Input-Constrained Capacity

$$C \approx \begin{cases} \alpha_1 \log_2\left(\dfrac{1+\alpha_2\gamma}{1+\alpha_3\gamma}\right), & \text{for } \gamma \le \gamma_{\max} \\ q, & \text{for } \gamma > \gamma_{\max} \end{cases}$$



Fading Channels

- We consider a Rayleigh model, in order to compute:
  - p.d.f. of the approximated input-constrained capacity

$$p_C(C) = \begin{cases} \beta e^{-\frac{\gamma_f(C)}{\overline{\gamma}}} + e^{-\frac{\gamma_{\max}}{\overline{\gamma}}} \delta(C-q), & 0 \le C \le q \\ 0, & \text{otherwise} \end{cases}$$

where $\beta = \frac{\ln(2)\left(1+\alpha_2\gamma_f(C)\right)\left(1+\alpha_3\gamma_f(C)\right)}{\overline{\gamma}\alpha_1(\alpha_2-\alpha_3)}$ and $\gamma_f(C) = \frac{2^{C/\alpha_1}-1}{\alpha_2-\alpha_3 2^{C/\alpha_1}}$;

  - p.d.f. of the lower bound on wiretapper's equivocation

$$p_{\tilde{s}}(\tilde{s}) = \begin{cases} \frac{q}{n} p_{C_E}(\tau) + \varphi\delta(\tilde{s}), & 0 \le \tilde{s} \le k' \\ 0, & \text{otherwise} \end{cases}$$

where $\tau = q\left(R_h - \frac{\tilde{s}}{n}\right)$ and $\varphi = \Pr\{C_E > qR_h\}$.

Wiretapper's Equivocation Under Outage Constraints

- Equivocation Outage Probability: probability that $\tilde{s}$ falls below some specified lower threshold $\tilde{s}_{min}$

$$P_O = \int_0^{\tilde{s}_{min}} p_{\tilde{s}}(\tilde{s})d\tilde{s} = 1 - \int_{\tilde{s}_{min}}^{k'} p_{\tilde{s}}(\tilde{s})d\tilde{s}$$

- We fix a level of semantic security equal to $\tilde{s}_{min}$; such level is achieved unless outage occurs

- In order to preserve such a security level, we must impose:

$$\frac{1}{P_O} \geq 2^{\tilde{s}_{min}}$$

Reliability and Security Conditions

- The minimum value for Eve's average SNR to achieve these conditions is:

$$\bar{\gamma}_E \leq \frac{\eta}{\tilde{s}_{min}\ln(2)} = \bar{\gamma}_E^*$$

$\tilde{s}_{min}$-bit semantic security over a single codeword

- We want semantic security over all the transmitted message, composed by *N* codewords

- SNR gap:

$$S_g = \frac{\gamma_B^*}{\bar{\gamma}_E^*}$$

Threshold values

RELIABILITY: we fix the maximum decoding error probability experienced by Bob

GOAL: Find *N* in such a way as to reach the level of semantic security we wish to achieve, as a function of the SNR gap

Example: WiMax Links

Setting:

- WiMax standard LDPC codes
  - $n = 2304;$
  - Rate $^1/_2, ^2/_3, ^3/_4, ^5/_6$
- Modulations
  - BPSK
  - 4-QAM
  - 16-QAM
- Reliability requirement
  - Decoding error probability $\leq 10^{-4}$ for Bob ⟹ $\gamma_B \geq \gamma_B^*$

R_c	1/2	1/2	1/2	2/3	2/3	2/3
Mod.	BPSK	4-QAM	16-QAM	BPSK	4-QAM	16-QAM
γ_B^*	-1.26	1.75	7.16	0.58	3.59	9.55

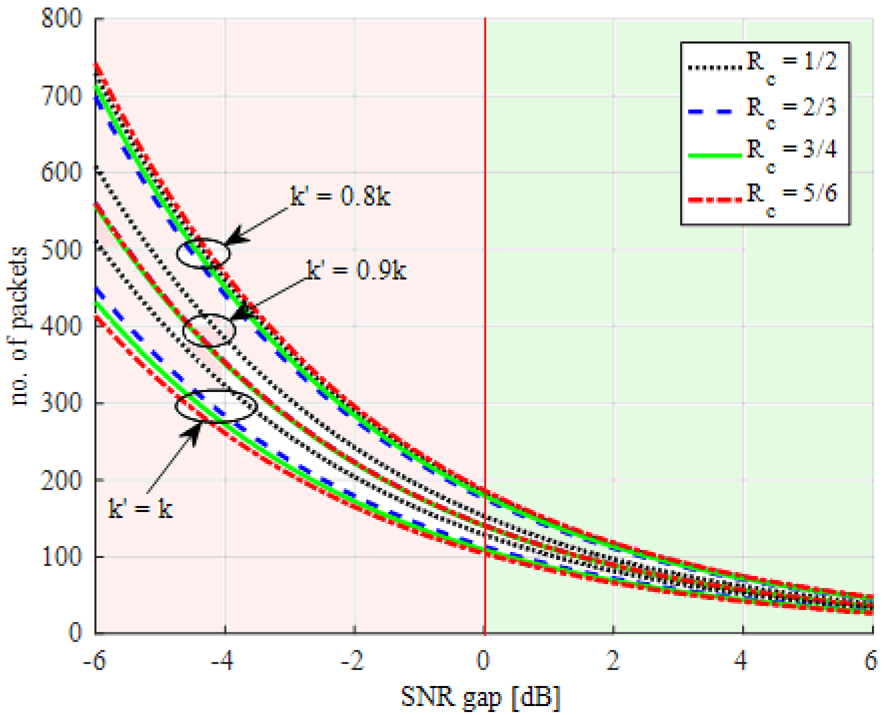
R_c	3/4	3/4	3/4	5/6	5/6	5/6
Mod.	BPSK	4-QAM	16-QAM	BPSK	4-QAM	16-QAM
γ_B^*	1.63	4.64	10.74	2.76	5.77	12.12

Results for 128-bit security

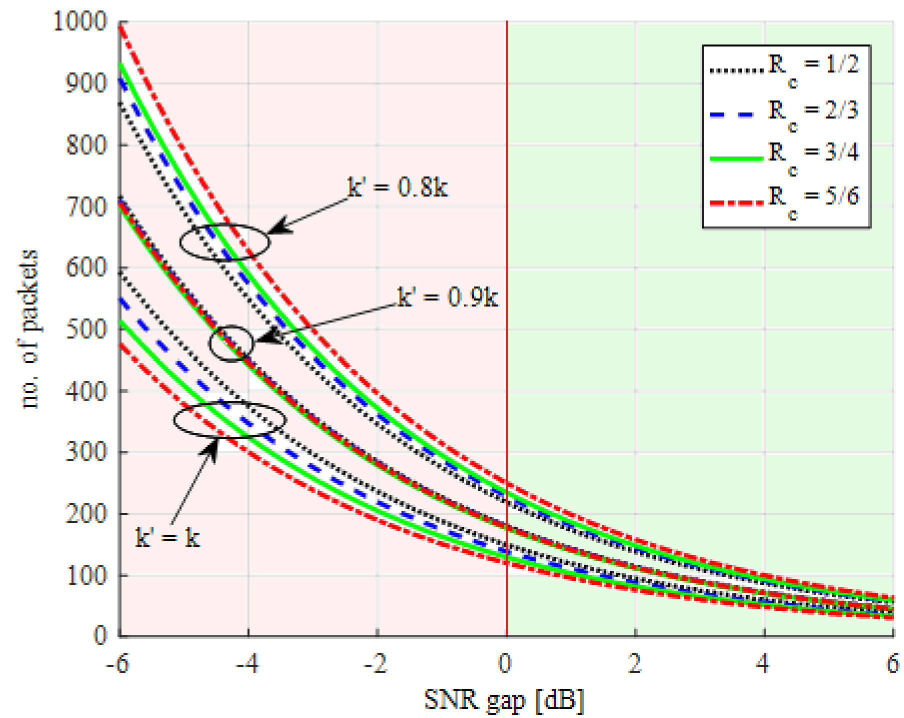
Eve's channel better than Bob's channel

Eve's channel worse than Bob's channel

BPSK



16-QAM



Number of packets needed to achieve 128-bit semantic security versus SNR gap with WiMax LDPC codes having length $n = 2304$, rate $R_c = {}^1/_2, {}^2/_3, {}^3/_4, {}^5/_6$, for the cases of $k' = 0.8k$, $k' = 0.9k$ and $k' = k$

Conclusions

- **Semantic security is achievable even in disadvantage conditions**, i.e. when the average SNR of Eve's channel is considerably larger than that of Bob's channel, although this is obviously paid in terms of an increasing number of packets

- Varying the **code rate has not great influence** on the required number of packets

- Using **high order modulations** is **not beneficial** from the number of packets standpoint, but they may be needed to ensure that the channel remains static during each three way communication between Bob and Alice

Future Work

- Introduction of transmission of fake packets when Bob’s channel is under a suitable threshold

- Generalization of the fading model, e.g., by exploiting the Nakagami distribution