

# Context aware 6G security: The role of the physical layer

**Arsenia Chorti**

ETIS UMR8051, CYU, ENSEA, CNRS, FR and Barkhausen Institut gGmbH, DE

2022 Joint IEEE SPS and EURASIP Summer School on  
Defining 6G

1. Key ideas in cryptography
2. What changes in 6G, trust and trustworthiness
3. The role of physical layer security in future generations

# Fundamentals security concepts

Jean Luc Godard: *All you need to make a film is a girl and gun*

Crypto expert: *all you need to do cryptography is a XOR and a key...*

Info theorist: *all you need to do information theory is a log and a limit*

What do we want to achieve?

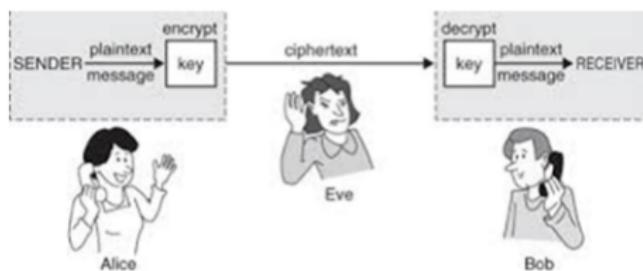
- What is the system model?
- What are the assumptions?
- What are the desirable properties?

## What do we aim to achieve?

- Data confidentiality; eavesdropping (passive) attacks
- Data integrity; active attacks
- User / device authentication (i.e., access control); active attacks
- Availability; active attacks

## What is the System Model?

- Basic model
- Extended model

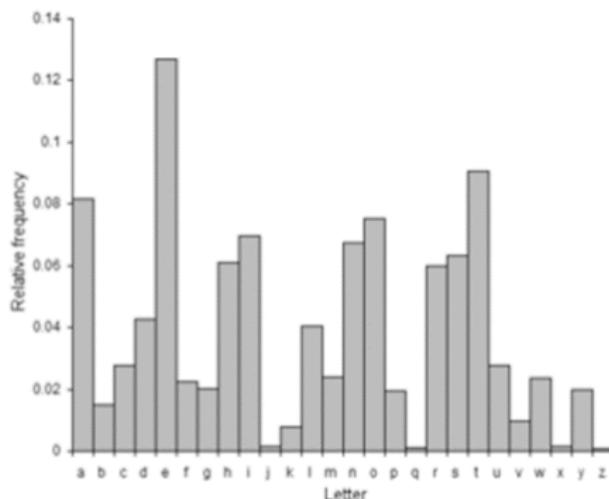


## Old example of symmetric key cipher

Substitution cipher: each letter is moved  $k$  positions to right (or left) in the alphabet

Brute force attack infeasible!

### Frequency of Letters in English Language



## One time pad (OTP)

A cipher  $E, D$  over  $\mathcal{M}, \mathcal{K}, \mathcal{C}$  has perfect secrecy if the entropy of the secret key is at least equal to the entropy of the message

$$c = k \oplus m, \quad c \in \mathcal{C}, k \in \mathcal{K}, m \in \mathcal{M}$$

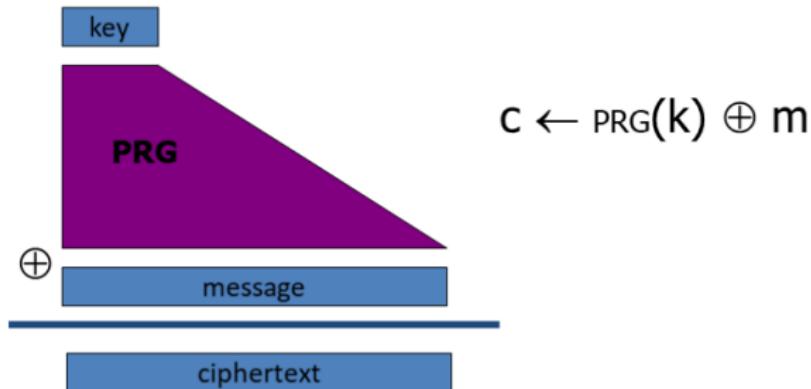
This means that the length of the key should be at least equal to the length of the message

The bad news: OTP is impractical

## Stream Ciphers: making OTP practical

Problem: OTP key is as long the message

Solution: Pseudo random key -- stream ciphers



credit: Dan Boneh, Cryptography 1

## Indistinguishability

The ensemble  $\{G(U_n)\}_{n \in N}$  is pseudorandom, iff for any probabilistic polynomial-time algorithm  $A$ , for any positive polynomial  $p$  and for all sufficiently large  $n$ ,

$$\left| \Pr(A(G(U_n)) = 1) - \Pr(A(U_{I(n)}) = 1) \right| < \frac{1}{p(n)} \quad (1)$$

## Next bit predictors

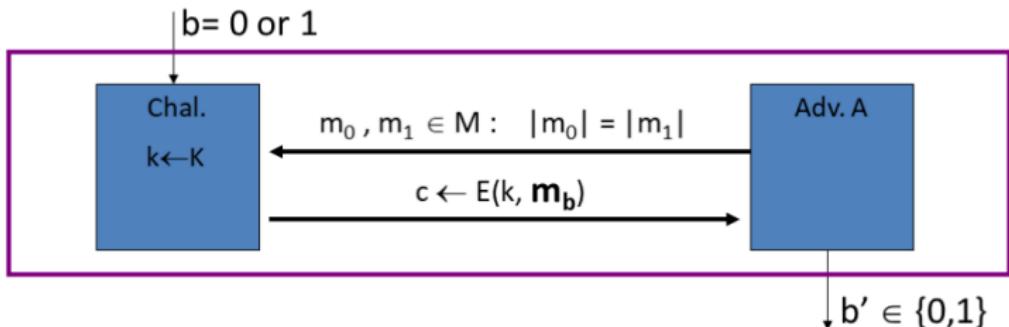
We say that  $G : \mathcal{K}\{0, 1\}^n$  is predictable if  $\exists$  an efficient algorithm  $A$  and an index  $i$  s.t.

$$\Pr \left[ A(G(k)|_{1, \dots, i}) = G(k)|_{i+1} \right] > \frac{1}{2} + \epsilon$$

for non-negligible  $\epsilon > \frac{1}{p(n)}$

# Semantic Security (one-time key)

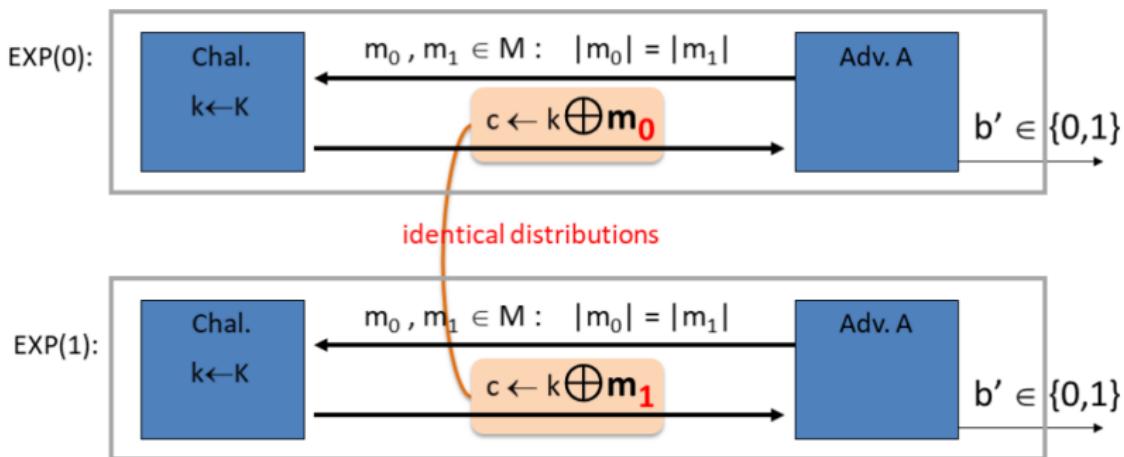
Define two experiments  $\text{EXP}(0)$  and  $\text{EXP}(1)$  as that correspond to messages  $m_0$  and  $m_1$



If there exists at least one experiment that would allow the adversary to predict which experiment was performed with non negligible probability then we don't have semantic security

credit: Dan Boneh, Cryptography 1

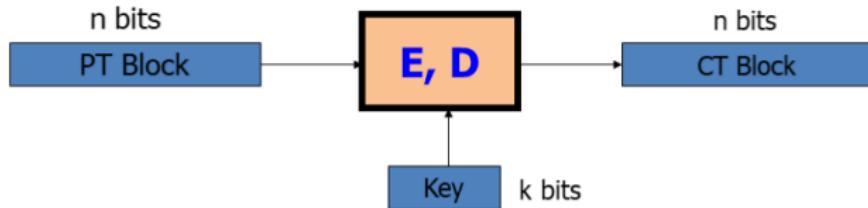
# OTP is semantically secure



credit: Dan Boneh, Cryptography 1

## Pseudorandom permutations and functions

### Modern Symmetric Block Ciphers



- Feistel Networks : DES, 3DES, CLEFIA (lightweight scheme)
- Substitution-permutation Networks: AES:  $n=128$  bits,  $k = 128, 192, 256$

credit: Dan Boneh, Cryptography 1

# Creating confusion using S-boxes

- S-box or a substitution box is a function producing a highly non-linear substitution of bits in a binary array
- S-boxes are normally represented as look-up tables

## S-box: Non-linearity

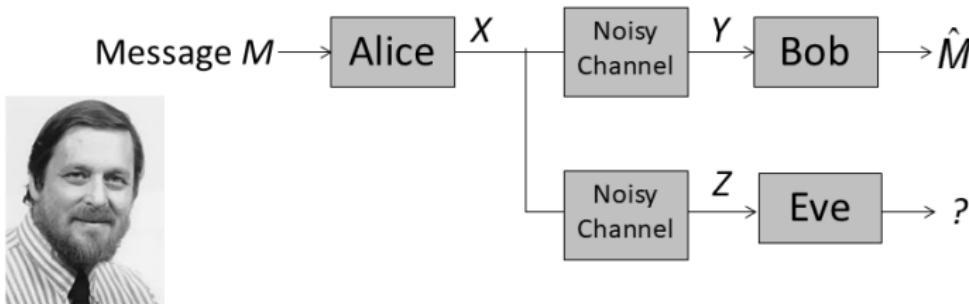
	*00*	*01*	*10*	*11*
0**0	01	00	11	10
0**1	11	10	01	00
1**0	00	10	01	11
1**1	11	01	00	11

The diagram illustrates an S-box with 8 input states and 8 output states. Inputs are grouped by their first two bits (0\*\* and 1\*\*) and outputs by their last bit (\*00\*, \*01\*, \*10\*, \*11\*). Curved arrows show the mapping from inputs to outputs. For example, 0\*\*0 maps to 00, 0\*\*1 maps to 00, 1\*\*0 maps to 11, and 1\*\*1 maps to 01.

For example, does the first bit of the output depend on the third bit of the input?  
Yes in a half of the cases, No in the other half of the cases

## Information Theoretic Security: Wyner's Model

“The Wiretap Channel”



- Tradeoff: reliable rate  $R$  to Bob vs. the equivocation  $H(M/Z)$  at Eve
- Secrecy capacity = maximum  $R$  such that  $R = H(M/Z)$
- Wyner (1975): Secrecy capacity  $> 0$  iff.  $Z$  is degraded relative to  $Y$

# Physical layer security (PLS)

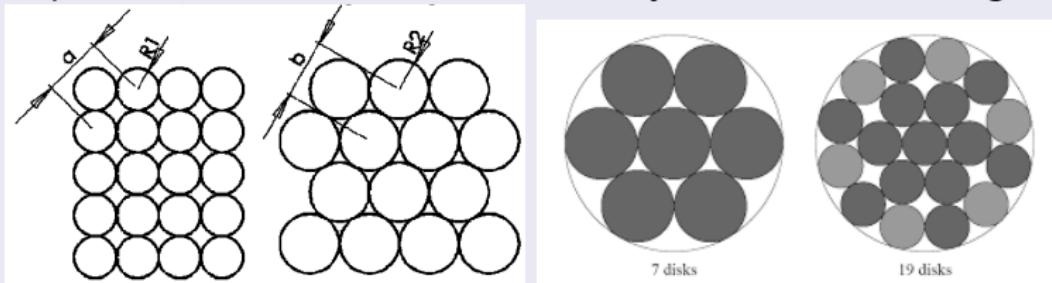
Physical properties of radio propagation (**diffusion, superposition and reciprocity**) provide opportunities for this

- Fading: provides natural degradedness over time
- Interference: allows active countermeasures to eavesdropping
- Spatial diversity (mMIMO, relays): creates “secrecy degrees of freedom”
- Randomness (small scale fading): sources of common randomness for key generation

# Secrecy capacity

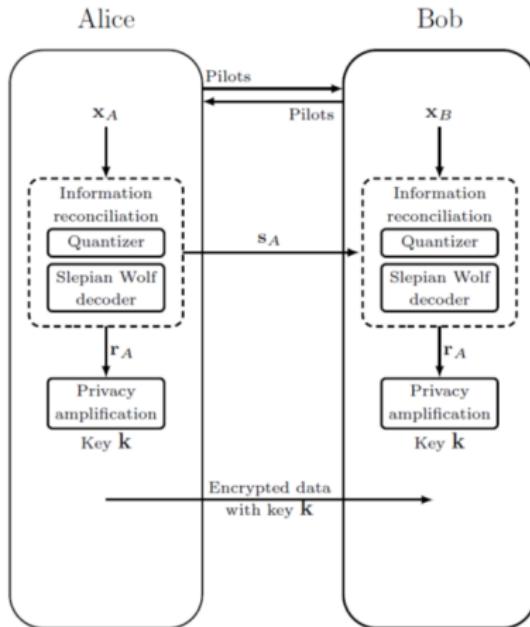
## Reminder on code construction for AWGN

For codelength  $K \rightarrow \infty$ , the observed vector in an AWGN with noise variance  $\sigma^2$  channel lies, with high probability, near the surface of a sphere with radius  $\sqrt{K}\sigma$ . For secrecy, use double binning.



## Secrecy capacity of broadcast channels

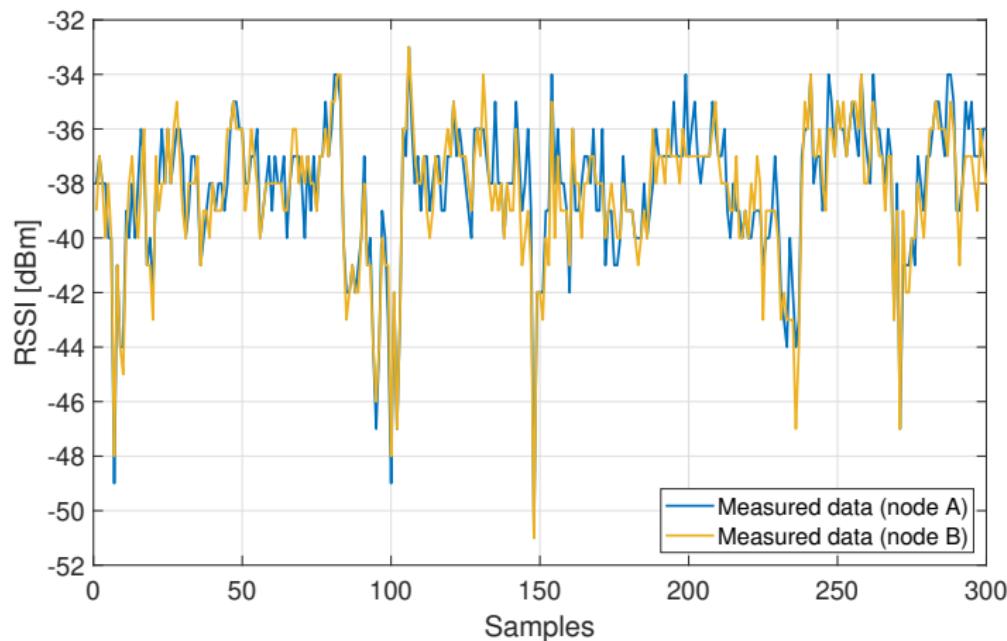
$$C_s = \max_{p(x)} [I(X; Y) - I(X; Z)]^+ \quad (2)$$



1. Quantizer outputs  $r_A = d \oplus e_A, r_B = d \oplus e_B$
2. Reconciliation: reconstruct  $r_A$  from  $r_B$  as  $r_A = r_B \oplus e_B \oplus e_A$ 
  - Temporal / spatial / frequency independence
3. Privacy amplification

# Secret key capacity

Alice and Bob observe a shared random event.



## Secret key capacity

$$C_k \leq I(X; Y|Z) \quad (3)$$

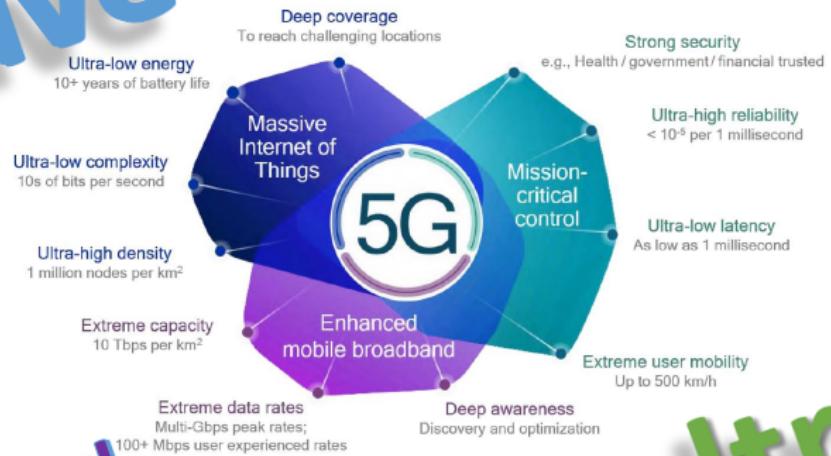
# What changes in 6G? (Will PLS be part of 6G?)

The need

The momentum

The technological roadmap

massive



enhanced

ultra<sup>2</sup>

- **Latency:**  $\sim 20$  msec to verify digital signatures on a vehicle [A. Teniou et al., Security and Privacy, 2018]
- **Scalability** in massive IoT: authentication and key distribution
- **Quantum computing:** post quantum cryptography
- **Constrained devices**, some deployed for 10+ years...
- **Artificial intelligence**, adversarial machine learning

Picture taken from BBC News (4th May 2021)



## Belgian farmer accidentally moves French border

| EUROPE

... "He made Belgium bigger and France smaller, it's not a good idea," David Lavaux, mayor of the Belgian village of Erquelinnes, told French TV channel TF1. That sort of move caused a headache between private landowners, he pointed out, let alone neighbouring states...

- **Fusion of digital and physical worlds**, autonomous systems making decisions

Νοῦς ὑγιὴς ἐν σώματι ὑγιεῖ

*A sound mind in a healthy body*

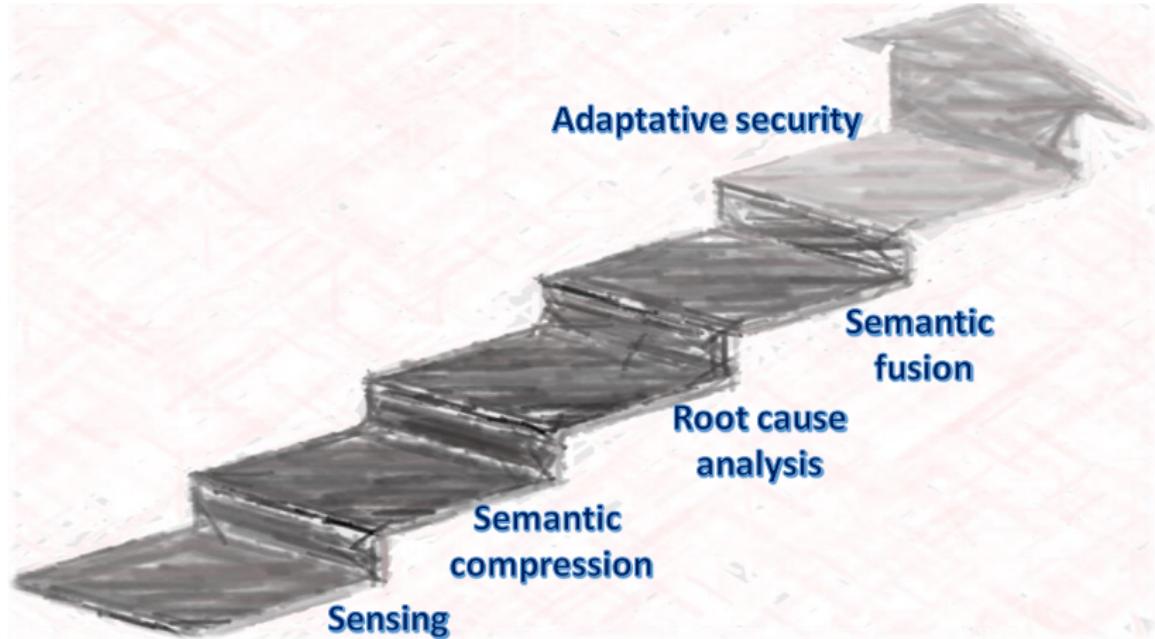
Until now trust has been studied for digital / cyber worlds

In 6G trust should include the physical world as well

**QoSec:** adaptive security levels, best effort security

- ① Define security / trust levels
- ② Build adaptive security controls
- ③ To deploy we need context awareness

# Context aware security: a sketch



## IEEE PLS Focus Group

Formed in September 2021, part of IEEE INGR Security

Steering Committee: A. Chorti (chair), A. Yener, M. Bloch, E. Hamad

More than 40 academic experts and industrial partners

- Main task: pave way for std WG on PLS for 6G
- Prepare white paper
- Engage with std organizations

4 working subgroups

- Use cases and applications (Stefano Tomasin)
- Threat model (Stefan Kopsell)
- Metrics and evaluation (Marco Baldi)
- Implementation aspects (Gunes Kurt)

# What physical layer security (PLS) can do for us

A spectrum of technologies with different TRLs:

- ① Counter-jamming at PHY (DoS attacks)
- ② Localization / RF fingerprinting as an authentication factor
- ③ Physical unclonable functions (PUFs) for device authentication
- ④ Secret key generation (SKG) from shared randomness
- ⑤ Anomaly detection monitoring hardware
- ⑥ Secrecy codes for wiretap channels, geofencing

Roadmap *or* Why is PLS not in 5G?

## Issue 1: the channel is not the one in the proof

Guarantees depend on **assumptions** for adversarial channel

$$\begin{aligned}C_s &= (C_m - C_e)^+ \\P_{out}^{(s)} &= \mathbb{E}[C_{s,k} < \tau]\end{aligned}$$

$$C_k = I(Z_A; Z_B | Z_E)$$

Solutions possible in 6G:

- Online, site specific learning of channel, let the data speak!
- Engineer transmission so that end-to-end it looks like in the proof
- Context awareness: taxonomy of PLS, which technology under what conditions

### Security guarantees in finite blocklength

Solutions possible in 6G:

- QoSec, privacy
- Recent important results

$$R_s(n, \epsilon, \delta) = C_s - \sqrt{\frac{V}{n}} Q^{-1} \left( \frac{\delta}{1 - \epsilon} \right) + \mathcal{O} \left( \frac{\log n}{n} \right) \quad (4)$$

$\delta$ : leakage and  $\epsilon$ : Bob's error probability [Yang et al., TIFS 2019]

- Analyses for stable throughput, scheduling, etc.

### PLS compromises transmission rates!

0.5 b/s/Hz not guaranteed even in subTHz when active attacks

[Ma et al., Nature, 2018]

Breaking the deadlock:

- Use PLS to distil / distribute keys
- Hybrid schemes, no need for high rates!
- Example: TLS AES 256 GCM SHA384 [GCM]: 96 bytes of key material for records of  $2^{14}$  bytes, i.e., 0.006 bits/sec
- Keys can be used over multiple records (e.g., up to  $2^{24}$ )

Adjust secrecy outages depending on QoSec level

$$P_{out}^{(s)} = \mathbb{E}[R_s < \tau] \quad (5)$$

$\tau$  : rate for key distribution

$P_{out}^{(s)}$  : select from QoSec

$R_s$  : finite blocklength analysis

transmission engineering

$\mathbb{E}[\cdot]$  : online learning of channel statistics

# Determinism & Randomness in Wireless Links for Authentication & Key Distillation

with Drs. Muralikrishnan Srinivasan, Sotiris Skaperas, Mahdi  
Shakiba Herfeh

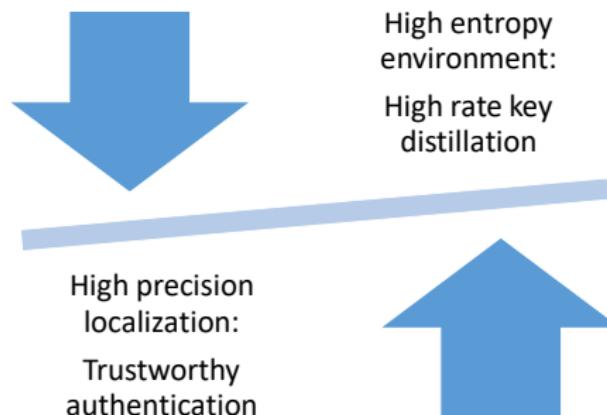


projects: SAFEST (DIM RFSI with Nokia Bell Labs)  
ELIOT (ANR PRCI with Univ. Sao Paolo and PUC Rio)  
eNiGMA (CYU INEX), PHEBE (CYU INEX)

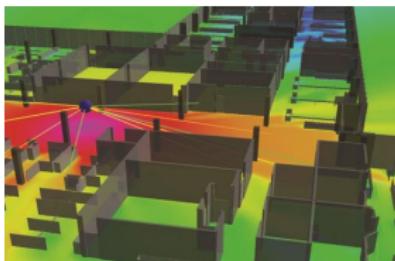
# The dual role of the wireless channel

Wireless coefficients can be used for:

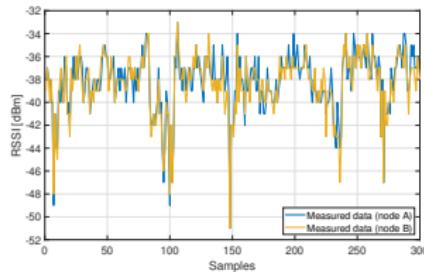
- ① Authentication through RF fingerprinting / localization
- ② Symmetric secret key generation (SKG)  $\implies$  Tx and Rx distil keys from "shared" randomness



# Deterministic and stochastic fading



Ray tracing



WiFi RSS at 9 m distance

- **Large scale fading** related to positioning
  - Largely predictable, useful for authentication
  - *Separability* is important (not necessarily decorrelation)
- 
- **Small scale fading** due to unpredictable variations in channel state information (CSI)
  - Source of entropy to distil keys from shared randomness
  - *Unpredictability* is important (resist next bit predictors)

## Pseudorandom number generator

Pseudorandomness of generator  $G$ : the ensemble  $\{G(U_n)\}_{n \in N}$  is pseudorandom, iff for any probabilistic polynomial-time algorithm  $A$ , for any positive polynomial  $p$ , and for all sufficiently large  $n$ ,

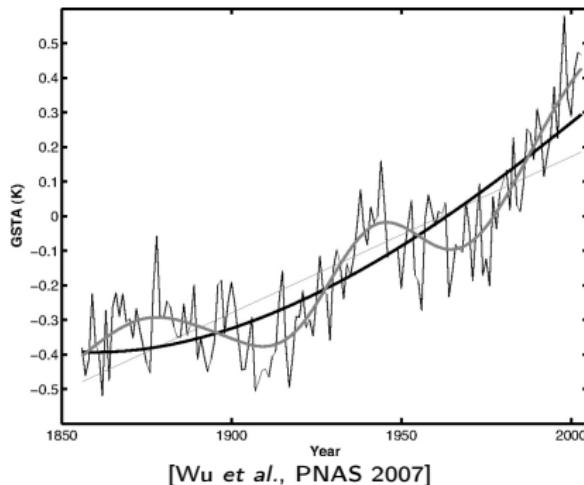
$$\left| \Pr(A(G(U_n); 1^{I(n)}) = 1) - \Pr(A(U_{I(n)}; 1^{I(n)}) = 1) \right| < \frac{1}{p(n)} \quad (6)$$

Sufficient condition of unpredictability is **independence**  
Achievable SKG rate

$$R_k \leq I(X; Y|Z) \quad (7)$$

**Wold's decomposition theorem:** every covariance-stationary time series can be written as the **sum** of two time series, one deterministic and one **stochastic**

Existence theorem, no guideline on how to do it



## How will we do it?

Key observation: separation more favourable in two domains,  
**frequency and power**

- ① Predictable, position dependent and separable components are probably slow varying and dominant in power
- ② Unpredictable, reciprocal and independent components are probably fast(er) varying, medium / low power

## Proposed pre-processing

In this work: focus on power domain separation

Employ unsupervised learning: PCA and autoencoders

Imply two different measures to capture separability and independence: TVD and HSIC

Work with synthetic *and* real data

## Proposed metric for separability

Total variation distance (TVD)

$$TVD(P, Q) = \frac{1}{2} \sum_{\omega \in \Omega} |P(\omega) - Q(\omega)| \leq \sqrt{\frac{1}{2} D_{KL}(P||Q)} \quad (8)$$

# Proposed metrics for unpredictability

## ① Pearson cross-correlation coefficient

$$\rho(\tilde{\mathbf{h}}_1, \tilde{\mathbf{h}}_2) = \frac{\mathbb{E} \left( \tilde{\mathbf{h}}_1 - \mathbb{E} \left( \tilde{\mathbf{h}}_1 \right) \right) \mathbb{E} \left( \tilde{\mathbf{h}}_2 - \mathbb{E} \left( \tilde{\mathbf{h}}_2 \right) \right)}{\sigma_{\tilde{\mathbf{h}}_1} \sigma_{\tilde{\mathbf{h}}_2}} \quad (9)$$

## ② dHSIC, kernel independence criterion based on Hilbert-Schmidt norm

$$dHSIC \left( \mathbb{P}^{(\tilde{\mathsf{H}}^1, \dots, \tilde{\mathsf{H}}^N)} \right) := \left\| \prod \left( \mathbb{P}^{\tilde{\mathsf{H}}^1} \otimes, \dots, \otimes \mathbb{P}^{\tilde{\mathsf{H}}^N} \right) - \prod \left( \mathbb{P}^{(\tilde{\mathsf{H}}^1, \dots, \tilde{\mathsf{H}}^N)} \right) \right\|_{\mathcal{H}}^2 \quad (10)$$

- $dHSIC = 0 \iff \mathbb{P}^{\tilde{\mathsf{H}}^1} \otimes, \dots, \otimes \mathbb{P}^{\tilde{\mathsf{H}}^N} = \mathbb{P}^{(\tilde{\mathsf{H}}^1, \dots, \tilde{\mathsf{H}}^N)}$
- $dHSIC$  is applied as a non parametric hypothesis test
  - $H_0$  :  $N$  – vectors are mutually (pairwise) independent
  - $H_1$  : not  $H_0$

## Asymptotic behavior under $H_0$ (critical value)

$dHSIC_M(\tilde{\mathbf{H}})$  can be approximated by a permutation test under  $H_0$

- Critical value

$$CV_\alpha = \left[ D^{dHSIC} \right]_{\lceil (B+1)(1-\alpha) \rceil + \sum_{i=1}^B 1_{\{dHSIC(\tilde{\mathbf{H}}) = dHSIC(\tilde{\mathbf{H}}_i)\}}}$$

- $D^{dHSIC}$  contains the  $B$  Monte-Carlo realisations of  $dHSIC(\tilde{\mathbf{H}})$  in an increasing order
- $dHSIC(\tilde{\mathbf{H}})$ ,  $\tilde{\mathbf{H}} = (r_1(\tilde{\mathbf{h}}_1), \dots, r_M(\tilde{\mathbf{h}}_M))$  is constructed by  $r_1, \dots, r_M$  random re-samplings without replacement
- $B$ : Monte-Carlo realizations •  $\lceil . \rceil$ : ceiling function
- $[.]_j$ :  $j$ -th element of a vector •  $1_{\{.\}}$ : indicator function

### Proposed normalized-metric

$$\overline{dHSIC} = \frac{dHSIC(\tilde{\mathbf{H}})}{CV_\alpha} 1_{\{dHSIC(\tilde{\mathbf{H}}) > CV_\alpha\}}$$

$\overline{dHSIC}$  near one  $\implies$  low dependence

- 2.180 GHz, 64 antennas, BS at height of 20 m, 50 OFDM subcarriers, 40 dBm (avg.) Tx power (7 dBi antenna gain)
- Each track has 45 subtracks, of 2000 timeshots of 64 complex channel measurements per subcarrier, subsampling by factor 5



Passive attacker (eavesdropper)

On the shoulder attack!

Active attacks addressed in other works:

- Jamming attacks: frequency hopping and energy harvesting
- Injection (impersonation): pilot randomization
- Tampering side information: authenticated encryption using SKG

# Statistical characterisation of the Nokia dataset (1)

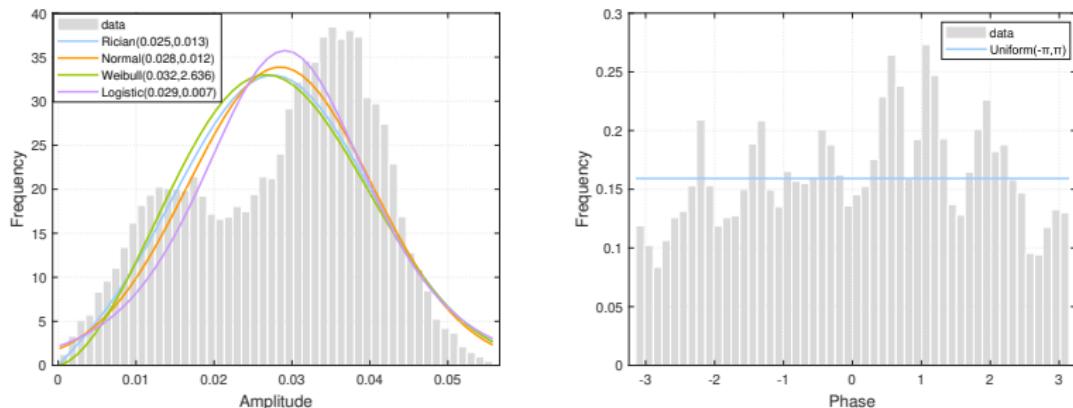


Figure: PDF fitting for the amplitude and the phase of measured CSIs on track 6

## Statistical characterisation of the Nokia dataset (2)

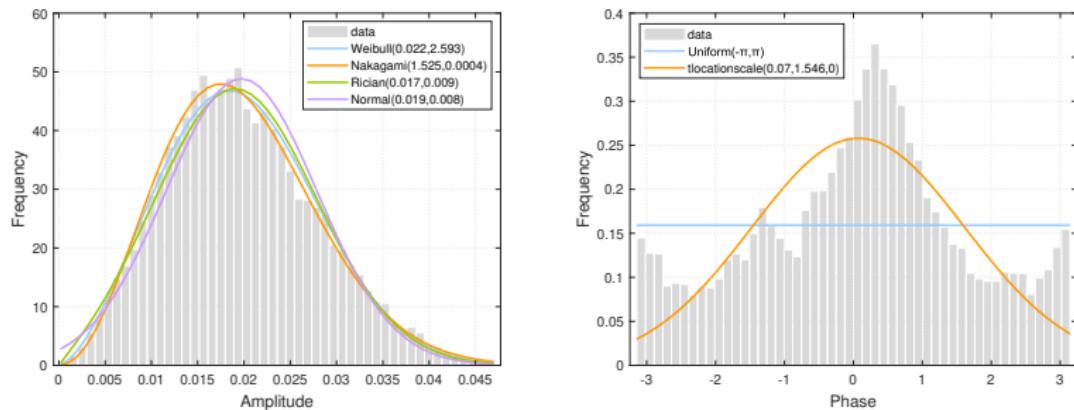


Figure: PDF fitting for the amplitude and the phase of measured CSIs on track 12

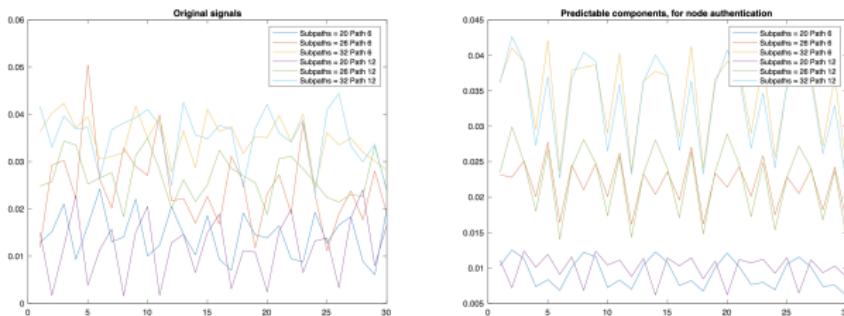
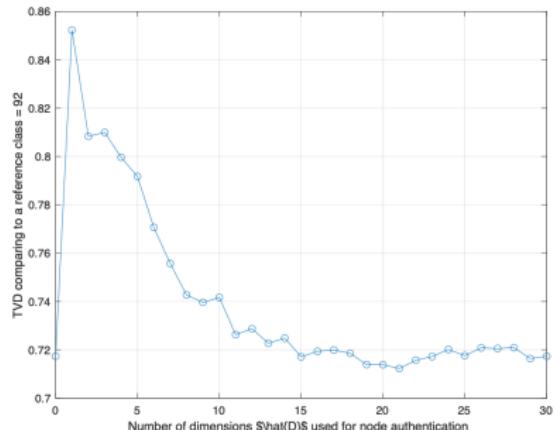
Conjecture: large scale fading dominant in power

- ① PCA of the CSI matrix to perform eigenvector decomposition
- ② Inverse PCA (invPCA) for *specific ranges* of eigenvectors (principal components – PCs)
- ③ **RF fingerprint:** invPCA over first  $\widehat{D}$  PCs
- ④ **Shared randomness:** invPCA over range  $\tilde{D}_1$  to  $\tilde{D}_1 + \tilde{D}_2$  PCs
- ⑤ **Denoising** by ignoring components after  $\widehat{D} + \tilde{D}_1 + \tilde{D}_2$

The triplet  $\{\widehat{D}, \tilde{D}_1, \tilde{D}_2\}$  chosen so that

- TVD between RF fingerprints is maximum
- Unpredictable components have low spatial, frequency (or time) correlation / dependence *while* being reciprocal

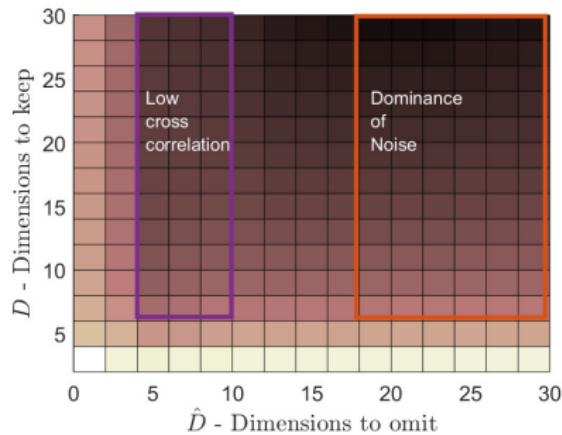
# PCA preprocessing for RF fingerprinting - Nokia



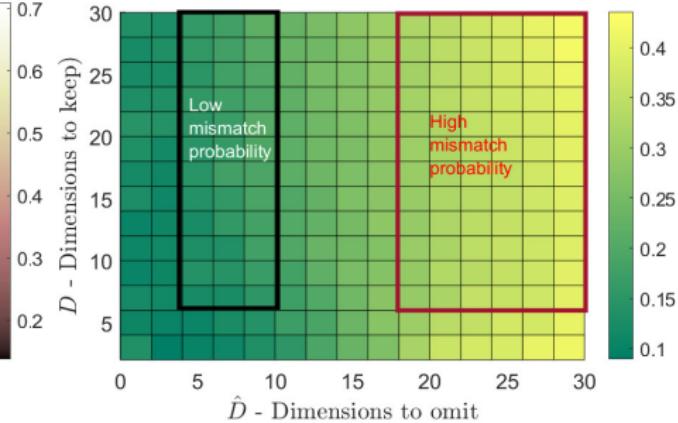
(a) Signals of 6 subtracks on tracks 6 and 12, SNR=20 dB

(b) First PC of 6 subtracks on tracks 6 and 12, SNR=20 dB

# PCA SKG pre-processing SNR=20 dB - Nokia



(a) Average CC (Original = 0.38)

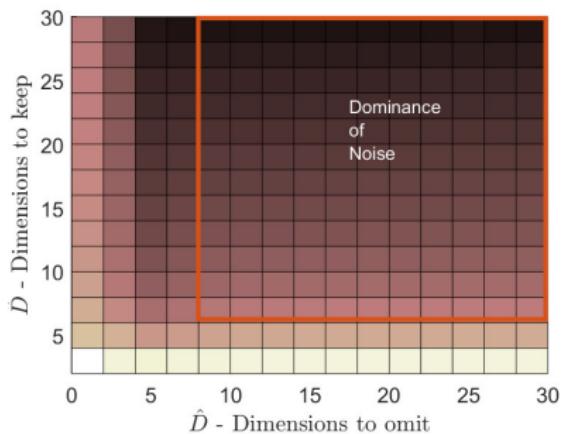


(b) Average MP

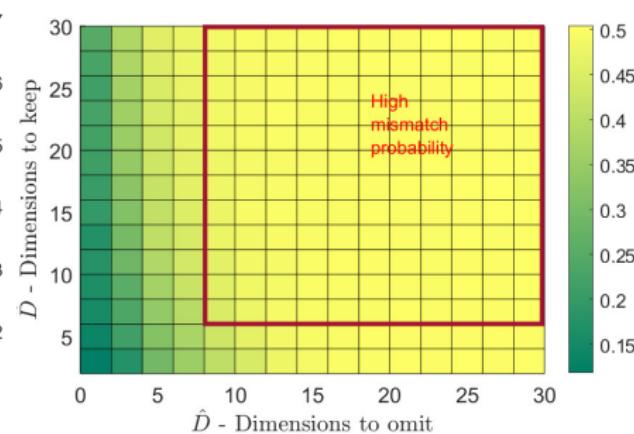
Figure: Trade-off for the Residual components for SNR = 20 dB for Nokia Data-set

For  $\tilde{D}_1 = 6$ ,  $\tilde{D}_2 = 30$ , CC drops to 0.14 with insignificant increase in MP

# PCA SKG pre-processing SNR=5 dB - Nokia



(a) Average CC (Original = 0.23)



(b) Average MP

Figure: Trade-off for the Residual components for SNR = 5 dB for Nokia dataset

# Evolution of $\overline{dHSIC}$

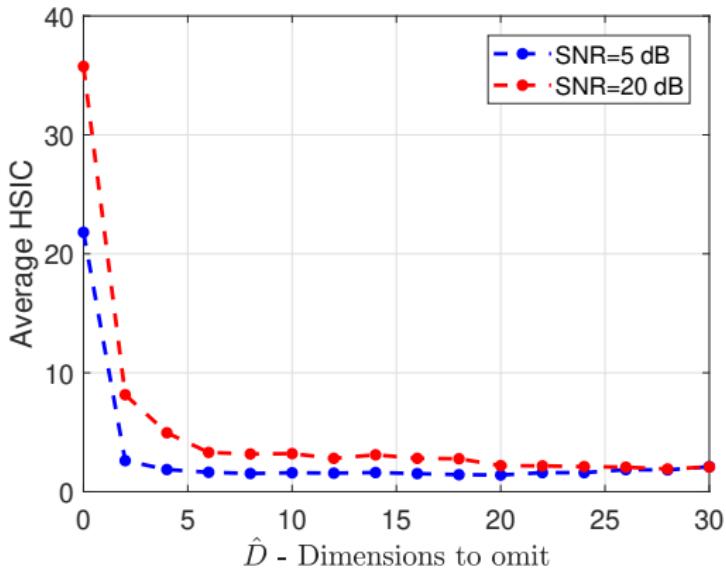
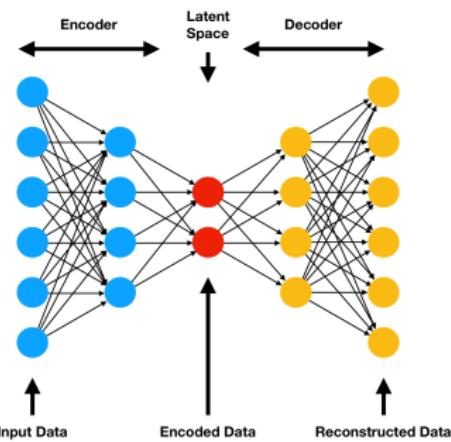


Figure: Nokia

Figure: Evolution of  $\overline{dHSIC}$  with  $\tilde{D}_1$  for  $\tilde{D}_2 = 30$  for Quadriga and Nokia datasets

# Autoencoder 1 - AE1



- Encoder maps  $M$ -d input matrix  $\mathbf{h}_{nu}$  to  $\tilde{D}$  dimensions
- Decoder maps back to an  $M$  dimensional output  $\hat{\mathbf{h}}_{nu}$
- Loss-function: MSE is minimized

$$E_1 = \frac{1}{N} \sum_{n=1}^N \|\mathbf{h}_{nu} - \hat{\mathbf{h}}_{nu}\|_2^2, \text{ for } u \in \{a, b\} \quad (11)$$

- Unpredictable component

$$\left\{ \tilde{\mathbf{h}}_{nu}(D) \right\}_{n=1}^N = \left\{ \mathbf{h}_{nu} - \hat{\mathbf{h}}_{nu} \right\}_{n=1}^N, \text{ for } u \in \{a, b\} \quad (12)$$

- Loss function includes an additional weighted correlation term
- Modified loss function

$$E_2 = \frac{1}{N} \sum_{\substack{n_1=1 \\ n_2 \in \mathcal{U}(n_1)}}^N \{\mathbf{h}_{n_1 u} - \hat{\mathbf{h}}_{n_1 u}\}^T \{\mathbf{h}_{n_2 u} - \hat{\mathbf{h}}_{n_2 u}\}, \quad (13)$$

for  $u \in \{a, b\}$

where  $\mathcal{U}(n_1)$  is the nearest neighbours of  $n_1$  (e.g., 8 by considering the first square grid around each point)

# AE architecture

[Studer et al., IEEE Access, 2018]

Layer	Dimensions	Activation
Input	200	Linear
1	100	tanh
2	50	softplus
3	20	tanh
Intermediate	$D$	linear
4	20	relu
5	50	softplus
6	100	tanh
Output	200	Linear

**Table:** The layers and activation function for AE1. For AE2 the only change is that the dimensions of the input and the output layers are 400

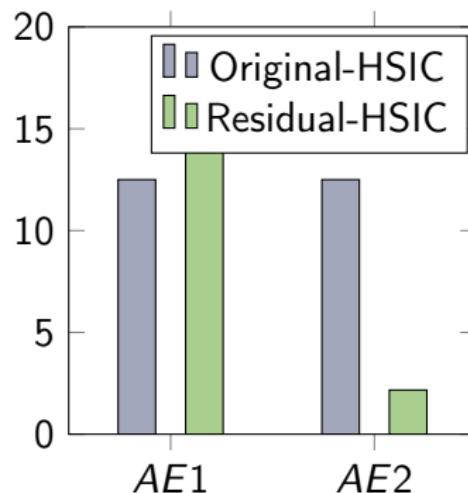
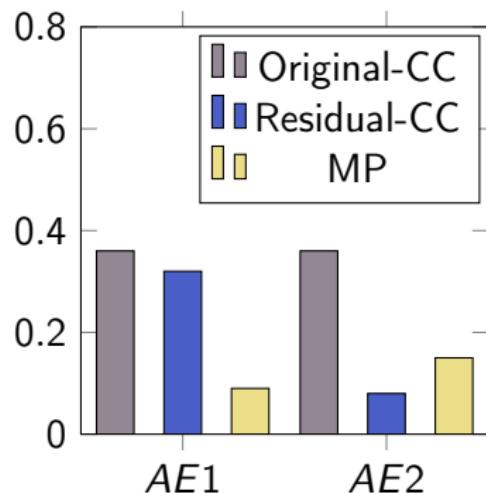
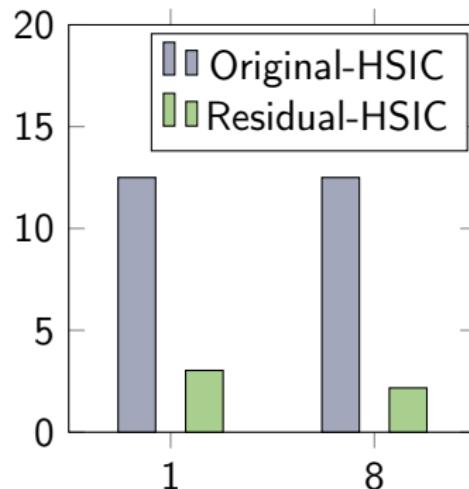
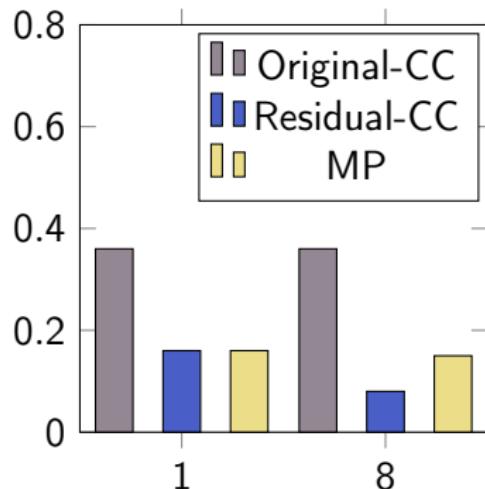


Figure: Key results for Nokia dataset for SNR= 20 dB and  $\tilde{D} = 8$

## Zoom in at AE2



**Figure:** Key results of AE2 for Nokia dataset for SNR= 20 dB to show the difference between  $\tilde{D} = 1$  and 8

- AE3 with composite loss function: weighted sum of  $\overline{dHSIC}$  and MP
- Online evaluation of channel statistics
- Trade-off between pre-processing and more hashing!

## Selected related papers

- M. Srinivasan, S. Skaperas, M.S. Herfah, A. Chorti, Joint Localization Based Node Authentication and Secret Key Generation, to appear in Proc. IEEE ICC 2022
- M. Srinivasan, S. Skaperas, and A. Chorti, "On the use of CSI for the generation of rf fingerprints and secret keys," in 25th Int. ITG Workshop on Smart Ant., 2021
- M. Mitev, A. Chorti, M.J. Reed and L. Musavian, Authenticated secret key generation in delay-constrained wireless systems, *Eurasip J. Wireless Com. Network*, 2020, vol. 122
- M. Shakiba-Herfah, A. Chorti, and H. Vincent Poor, Physical Layer Security: Authentication, Integrity, and Confidentiality. Springer Inter-national Publishing, 2021, pp. 129–150.
- M. Shakiba-Herfah and A. Chorti, "Comparison of short blocklength slepian-wolf coding for key reconciliation," in IEEE Stat. Signal Process. Workshop SSP 2021, 2021
- A. Chorti, A. N. Barreto, S. Kopsel, M. Zoli, M. Chafii, P. Sehier, G. Fettweis, and H. Poor, Context-aware security for 6G wireless, the role of physical layer security, *IEEE Communications Standards Mag.*, March 2022
- A. N. Barreto, S. Köpsell, A. Chorti, B. Poettering, J. Jelitto, J. Hesse, J. Boole, K. Rieck, M. Kountouris, D. Singele and Kumar Ashwinee, Towards Intelligent Context-Aware 6G Security, under review *IEEE Network Mgz*
- W. Njima, M. Chafii, A. Chorti, R. M. Shubair, and H. V. Poor, "Indoor localization using data augmentation via selective generative adversarial networks," *IEEE Access*, vol. 9, pp. 98 337–98 347, 2021
- G. Rezgui, E.V. Belmega, A. Chorti, "Mitigating Jamming Attacks Using Energy Harvesting", *IEEE Wireless Communications Letters*, 2019
- A. Chorti, "Optimal signalling strategies and power allocation for wireless secret key generation systems in the presence of a jammer," in 2017 ICC, IEEE, 2017
- E.V. Belmega, A. Chorti "Protecting Secret Key Generation Systems against Jamming: Energy Harvesting and Channel Hopping Approaches", *IEEE Trans. Inf. Forensics Security*, 2017
- G.A. Nunez Segura, A. Chorti, C. Borges Margi, "Centralized and Distributed Intrusion Detection for Resource Constrained Wireless SDN Networks", early access *IEEE IoT Journal*, 2021
- G.A. Nunez Segura, C. B. Margi, A. Chorti, "Understanding the Performance of Software Defined Wireless Sensor Networks Under Denial of Service Attack", *Open Journal of Internet of things (OJIOT)*, 2019

- Trust and trustworthiness in 6G: moving away from static security
- Securing RAN: first line of defence for 6G
- PLS enabled by sensing and AI: towards context aware security
- Preprocessing for RF fingerprinting and SKG with unsupervised learning: PCA and AE

Thank you for the invitation!

Questions? email: arsenia.chorti@ensea.fr