NOKIA Bell Labs

The Road Towards an AI-Native Air Interface (AI-AI) for 6G

Jakob Hoydis

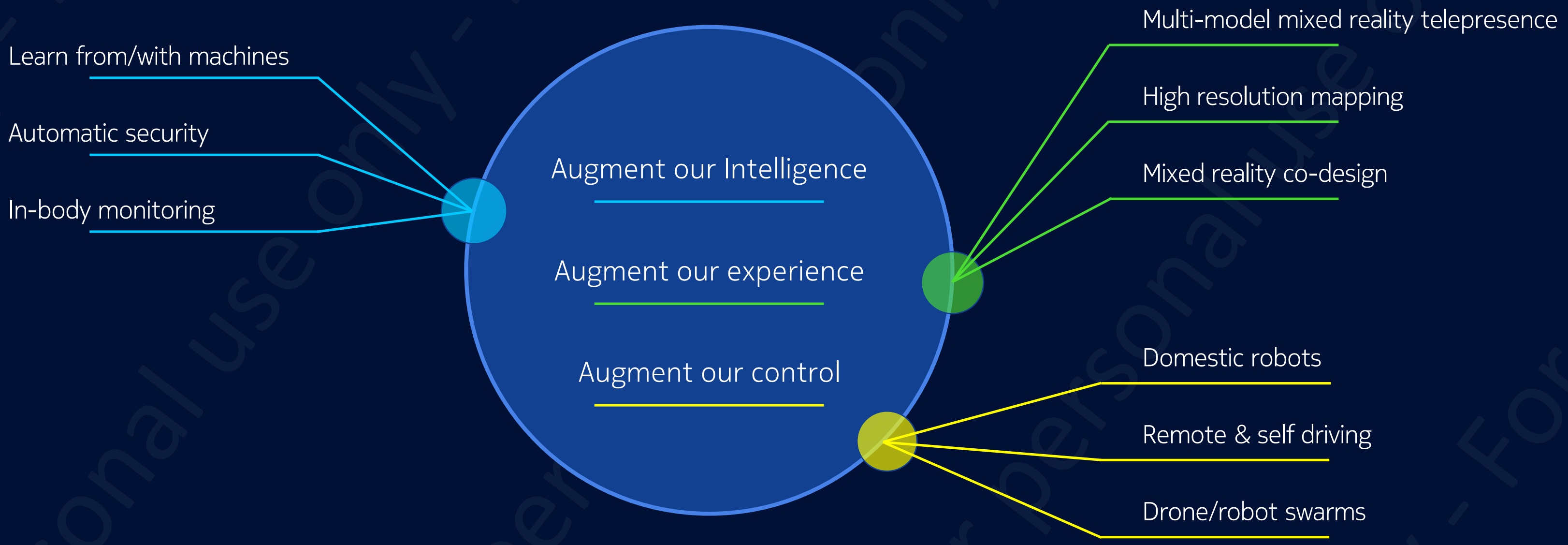
Radio Systems Research & AI

Nokia Bell Labs, France

jakob.hoydis@nokia-bell-labs.com

What will future communications look like in 2030?

Creating the 'augmented human'

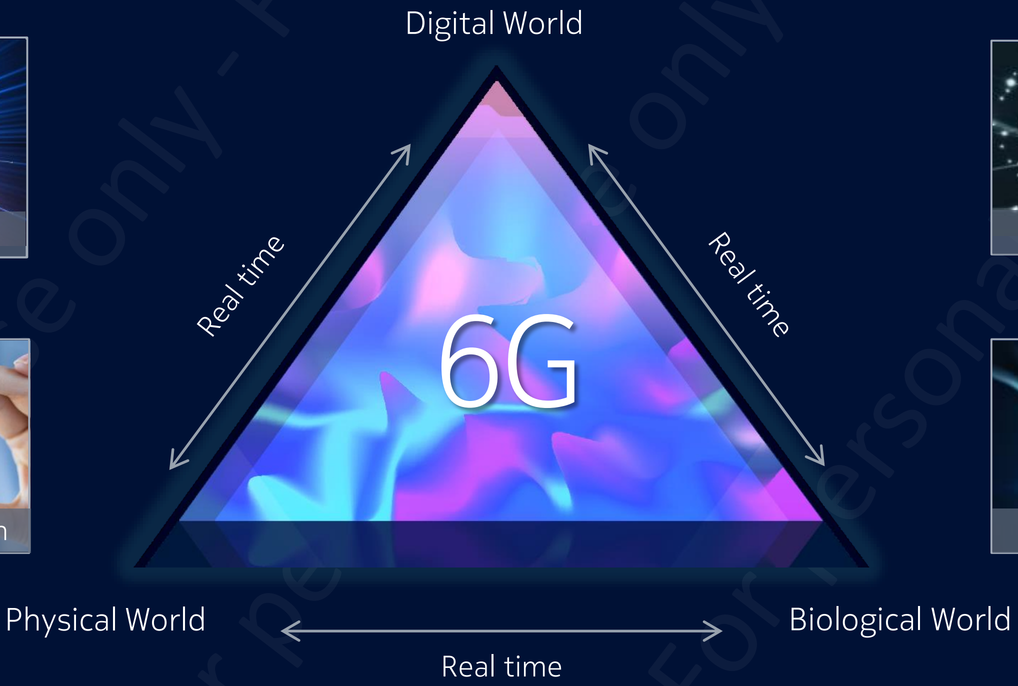


6G to enable a new lifestyle at scale

The enabling foundation for that future…


Ubiquitous Compute


Precision Sensing & Actuation




Knowledge systems


Human Machine Interface

H. Viswanathan and P. Mogensen, “Communications in the 6G Era”, IEEE Access, 2020

Six key technologies for 6G

AI/ML Air-Interface

New Spectrum Technologies

Network as a sensor

RAN-Core Convergence & Specialization

Extreme Connectivity

Security and Trust

Role of ML for 5G

Slide orchestration
VM management
Anomaly detection

Random Access detection
Symbol demapping
MIMO detection
Channel estimation
MIMO User Pairing

Deployment optimization
Load balancing
Carrier Aggregation
Handover prediction
Beam Prediction
Interference Management

User localization
Trajectory prediction

Complexity | Algorithms not optimal | Lack of accurate models

No component of 5G has been designed by ML

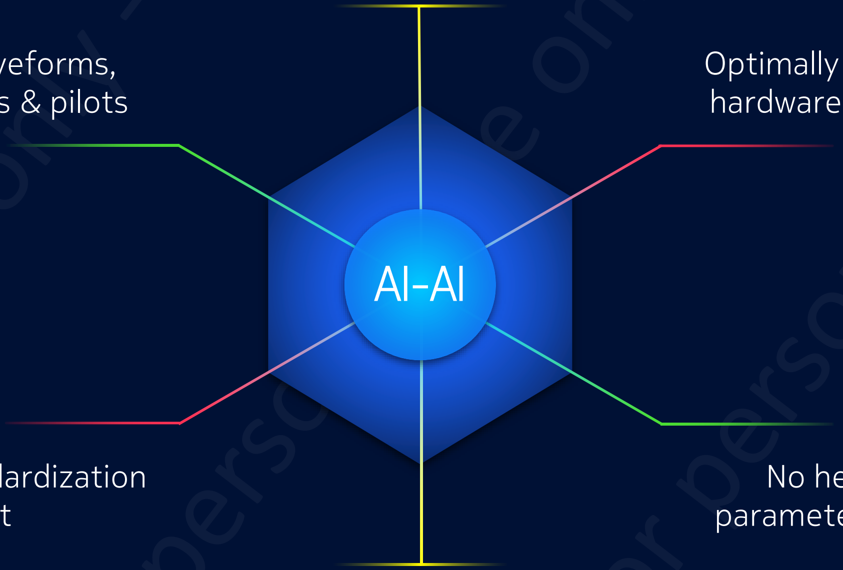
What if 6G was built so that ML could optimize parts of the PHY & MAC if needed?

Possible benefits

Custom signaling and access schemes

Bespoke waveforms, constellations & pilots

Optimally adapted to hardware limitations

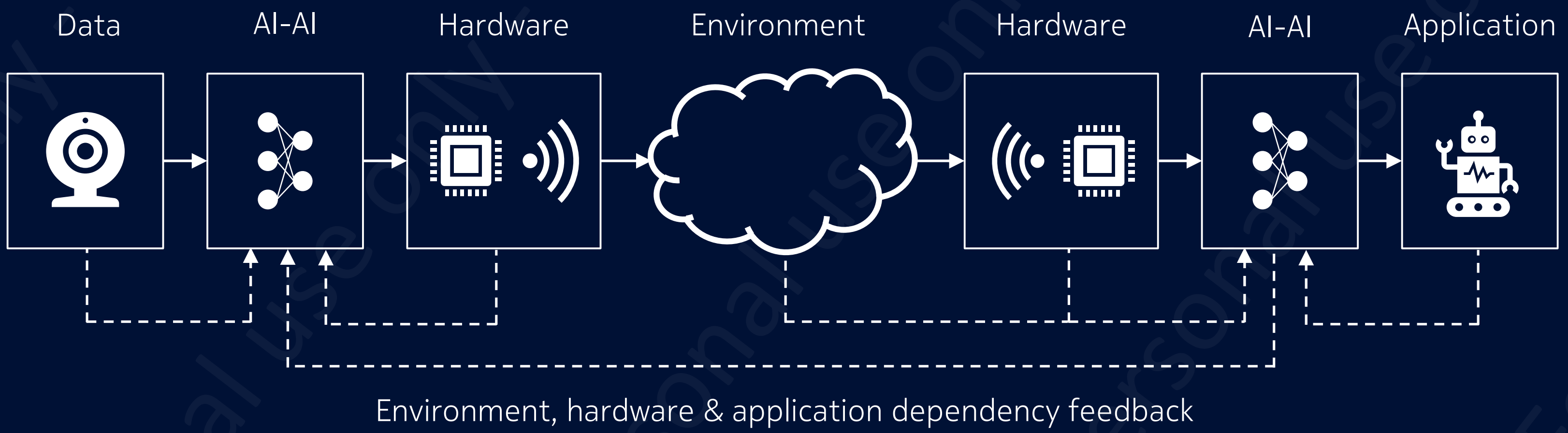


Reduced standardization effort

No heuristic parameter settings

Faster development & deployment time

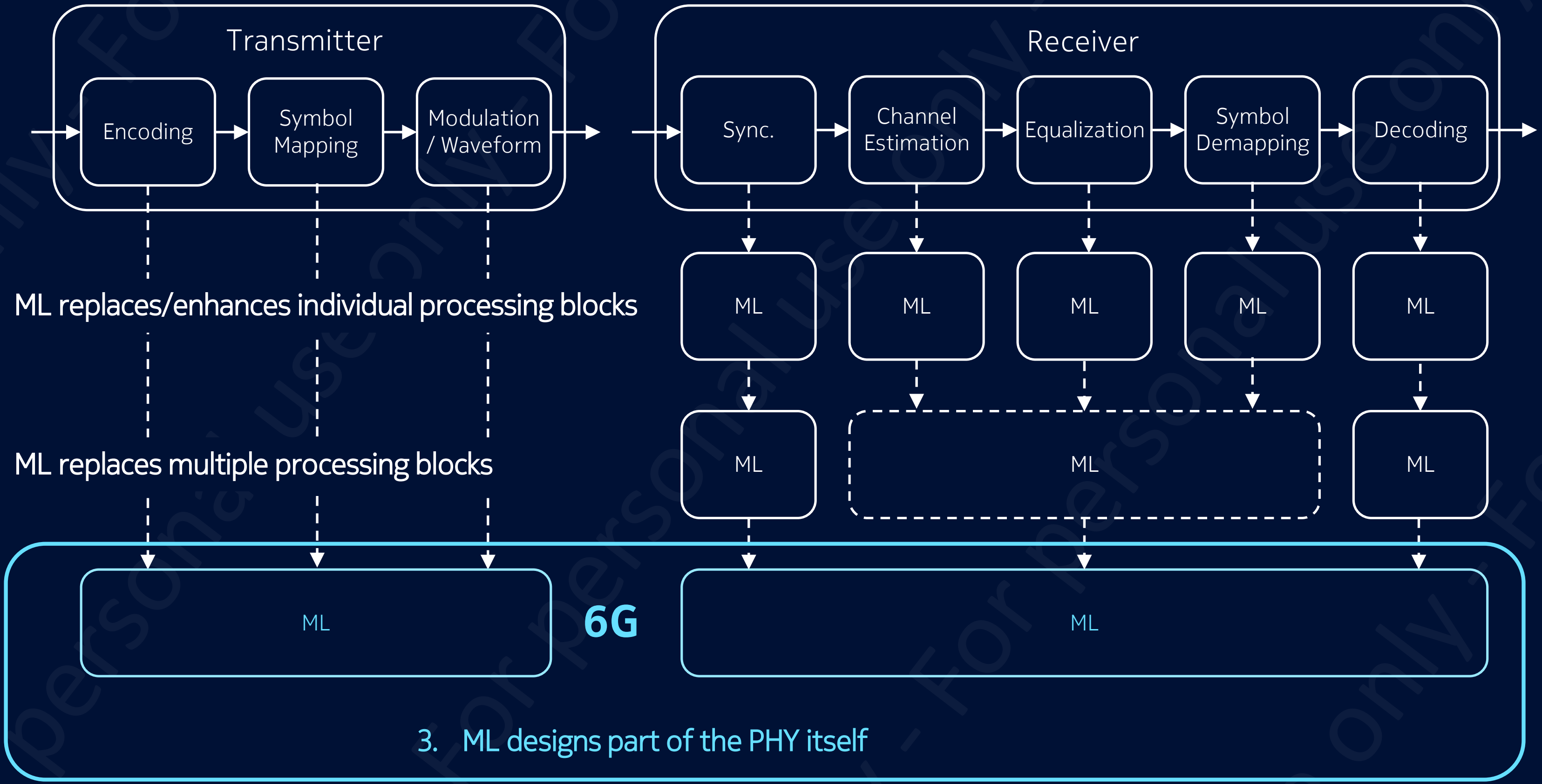
AI-Native Air Interface (AI-AI) for 6G



"Post Shannon": Not about reliably transmitting bits anymore, but rather <u>serving an application with data in an optimal way</u>

The AI-AI optimally adapts to different environments, hardware, data, and applications

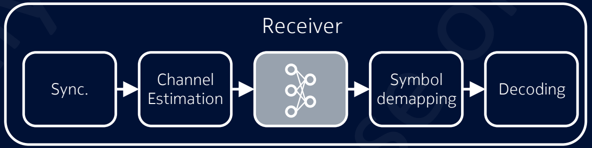
A roadmap to an AI-Native Air Interface for 6G



The importance of each step in the transition

1. ML replaces/enhances individual processing blocks

- Paradigm change in transceiver design & deployment
- Online training & transfer/federated learning
- No new signaling

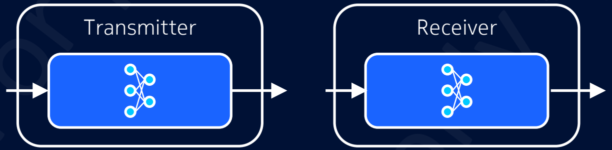


2. ML replaces multiple processing blocks

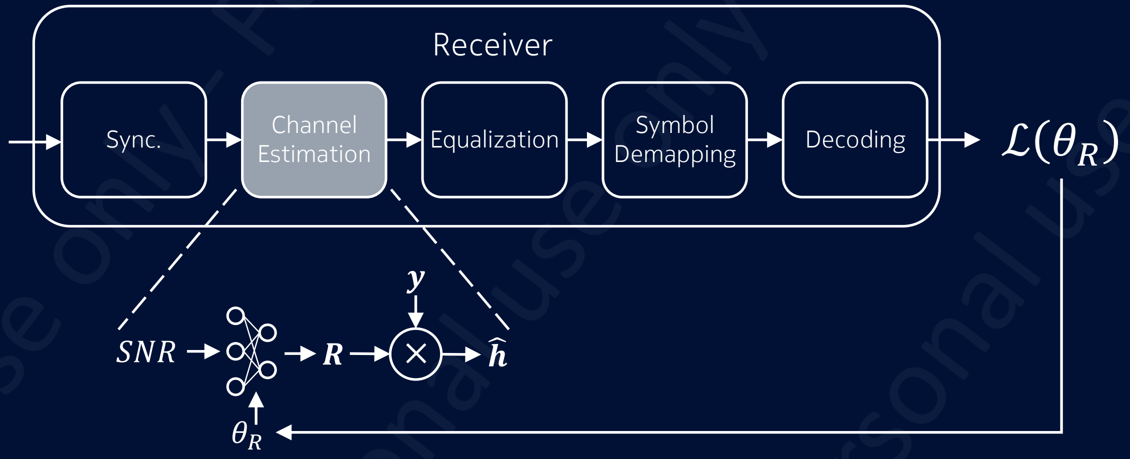
- Enables qualitative new features (e.g., no cyclic prefix, less pilots)
- HW acceleration essential
- ML-first approach without backup solution

3. ML designs part of the PHY itself

- Paradigm change in communication systems design
- Distributed & end-to-end learning
- New signaling & procedures



Differentiable algorithms enable hybrid "ML-Expert" systems



Training on end-to-end loss :

- No need for block-wise ground-truth
- Each block optimized for end-to-end performance

Fully differentiable transceivers allows simple integration of trainable components

A universal end-to-end design goal

Estimated posterior bit probabilities

Binary cross-entropy

$$\mathcal{L}(\theta_T, \theta_R) = \sum_{m=1}^{M} -\mathbb{E}_{b_i,y}\left[\log_2\left(\tilde{p}_\theta(b_i|y)\right)\right]$$

Achievable rate with a mismatched bit-metric decoder

$$= M - \left(\sum_{m=1}^{M} I(b_i; y) - \sum_{m=1}^{M} \mathbb{E}_y\left[\mathrm{D}_{\mathrm{KL}}\left(p(b_i|y) \| \tilde{p}(b_i|y)\right)\right] \right)$$

Number of transmitted bits

Bit-metric decoding rate (depends on transmitter)

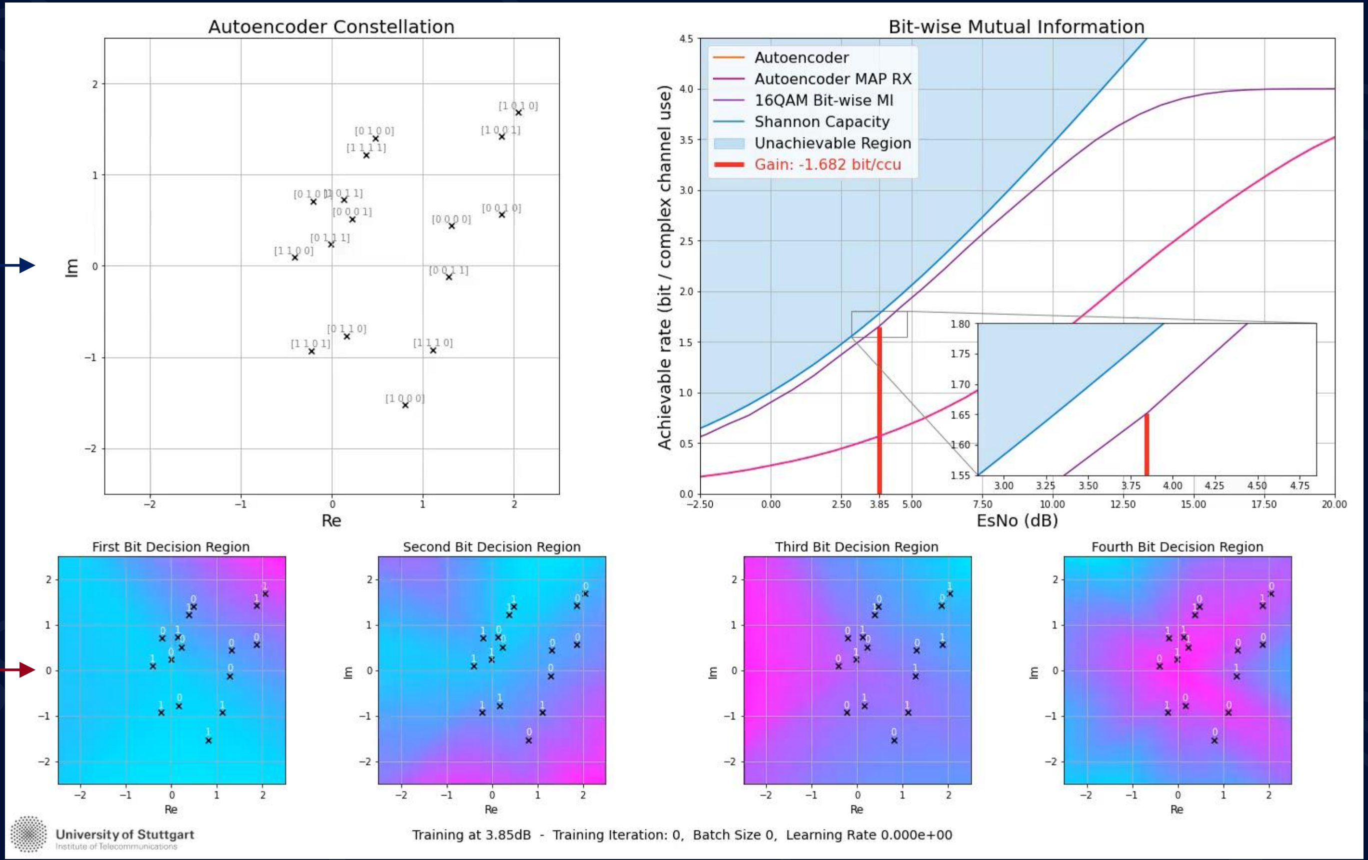
Loss due to imperfect receiver (depends on receiver)

Minimizing the binary cross-entropy maximizes an achievable rate with a practical decoder

Understanding end-to-end learning on an AWGN channel

Constellation determines bit-metric decoding rate

Decision regions determine loss w.r.t. MAP



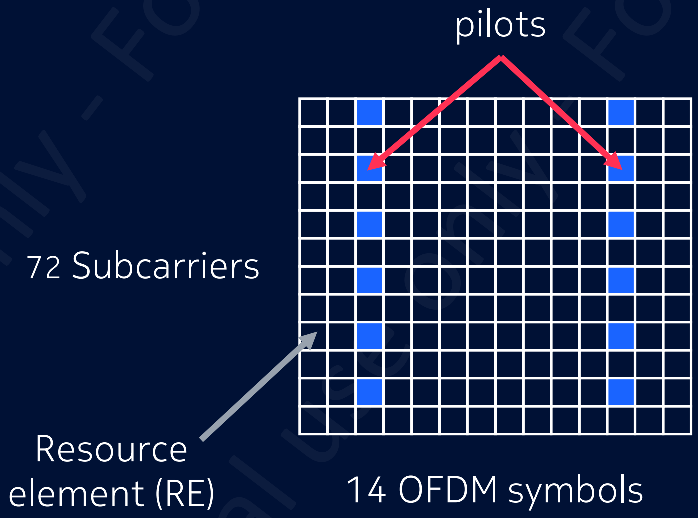
Curtesy of Sebastian Dörner, University of Stuttgart

S. Cammerer, et al. "Trainable communication systems: Concepts and prototype," IEEE Trans. Commun., no. 9, vol. 68, Sep. 2020.

Case study:

From Neural Receivers to Pilotless Transmissions

SISO doubly-selective channel



$$\boldsymbol{Y} = \boldsymbol{H} \circ \boldsymbol{X} + \boldsymbol{N}$$

$$\mathrm{vec}(\boldsymbol{H}) \sim \mathcal{CN}(0, \boldsymbol{R}_F \otimes \boldsymbol{R}_T)$$

Spectral correlation

- $[\boldsymbol{R}_F]_{i,k} = \sum_{l=1}^{L} S_l e^{j2\pi\tau_l D_s \Delta_F (i-k)}$
- Subcarrier spacing $\Delta_F = 30$ kHz
- Delay spread $D_s = 100$ ns
- TDL-A power delay profile

Temporal correlation

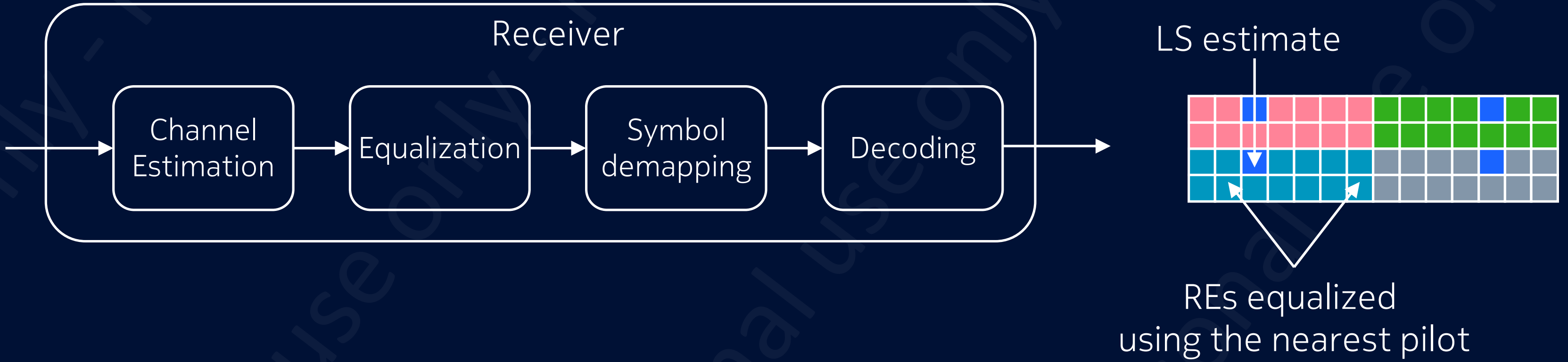
- $[\boldsymbol{R}_T]_{i,k} = J_0\left(2\pi \frac{v}{c} f_c \Delta_T (i-k)\right)$
- Carrier frequency $f_c = 3.5$ GHz
- Speed $v = 50$ km/h

Modulation & Coding

- 64 QAM
- 5G code n=1024, r=2/3

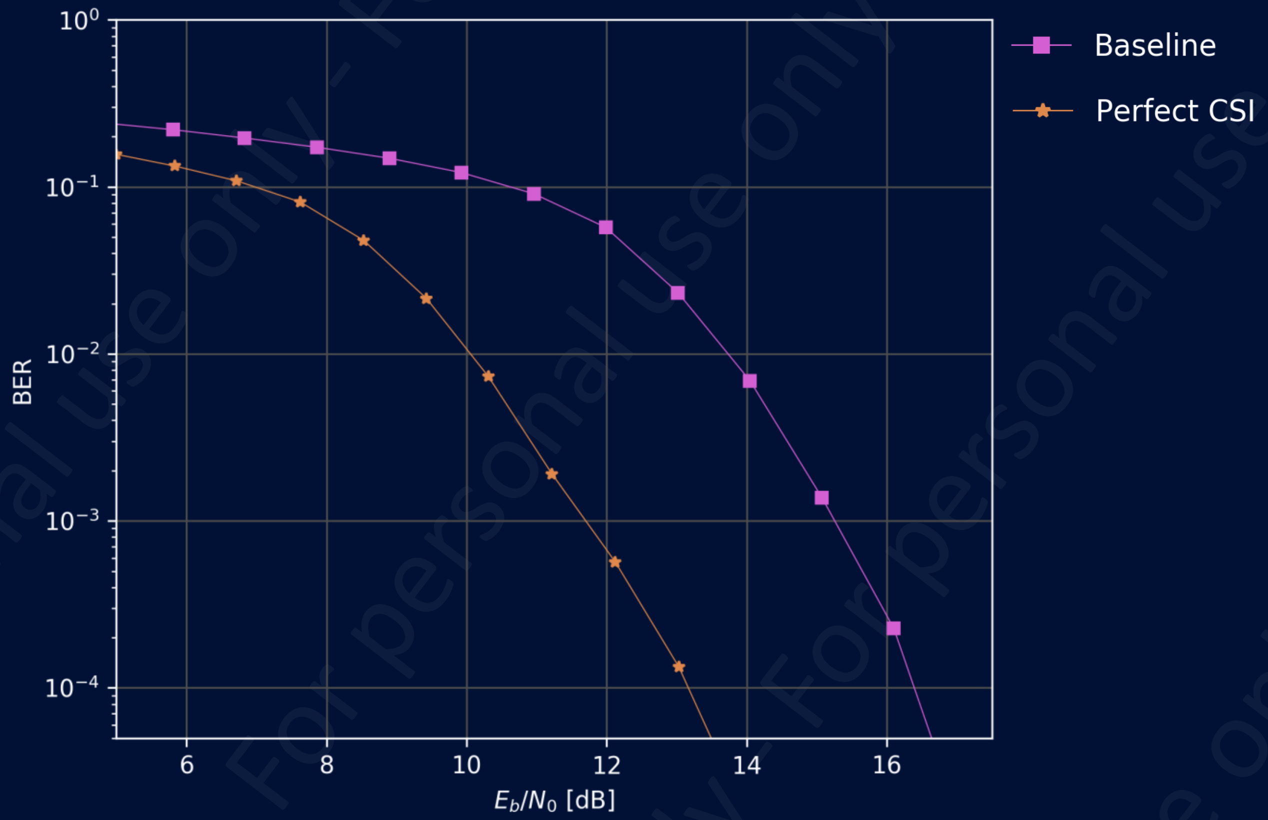
NOKIA Bell Labs

Baseline receiver

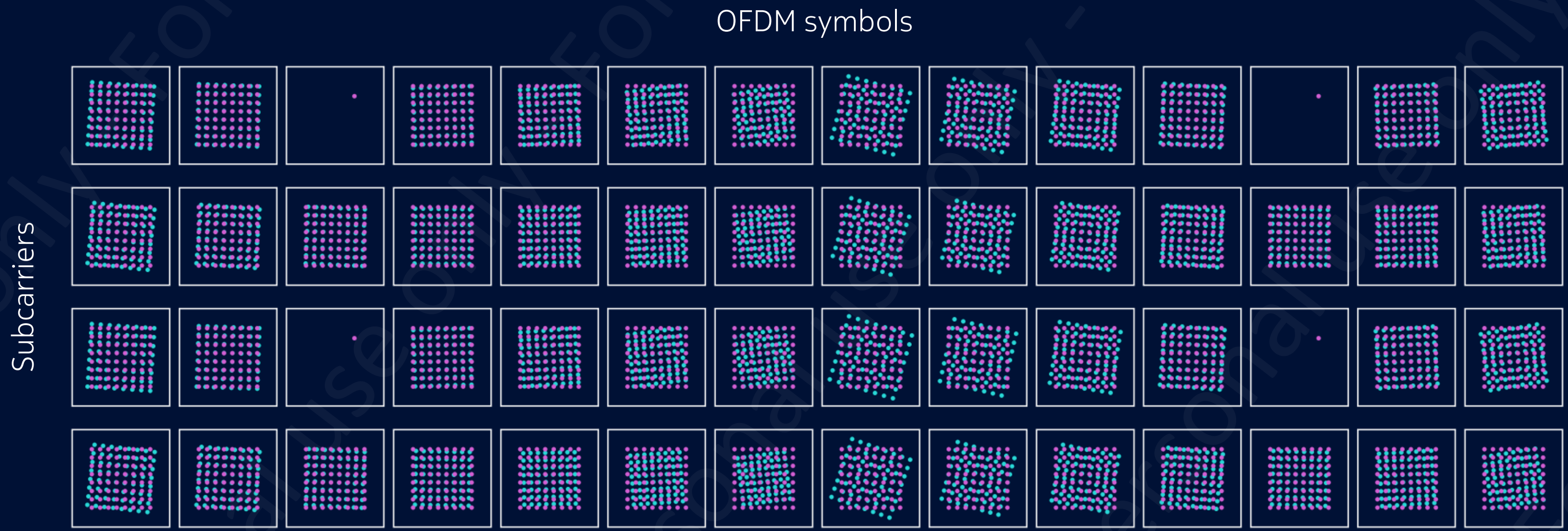


- Least-squares channel estimation at pilot positions
- Equalization using the nearest pilot
- Exact LLR computation assuming a Gaussian post-equalized channel
- Textbook sum-product BP decoder with 40 iterations

Potential performance enhancements



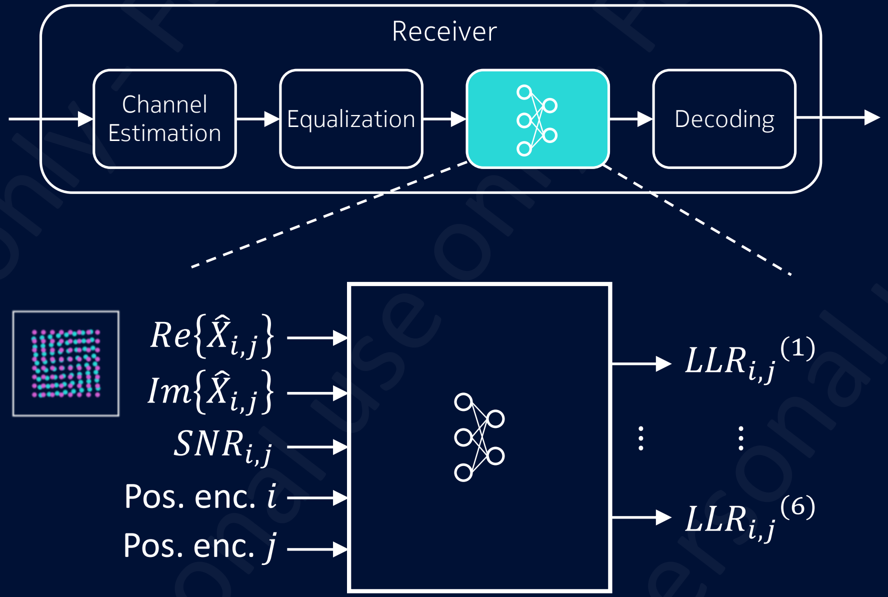
Deficits of the baseline

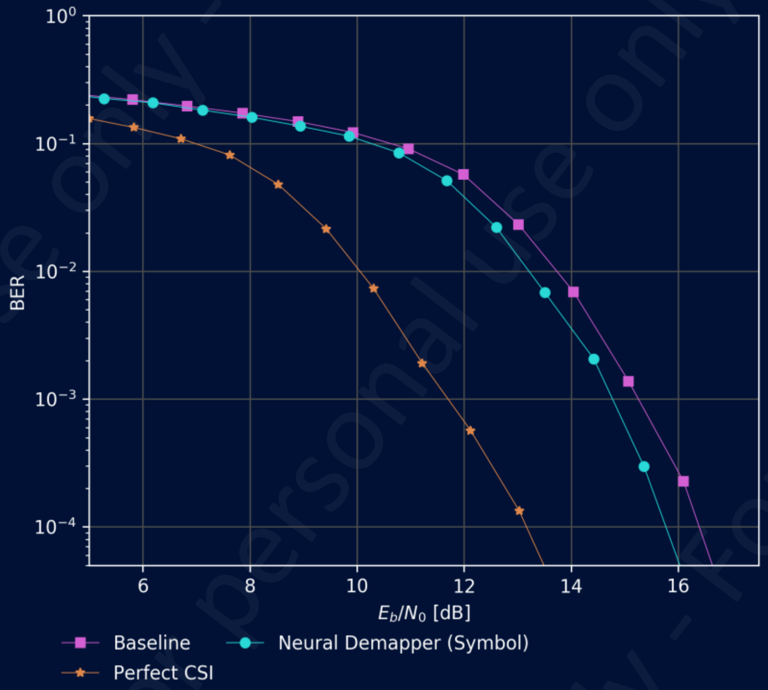


Imperfect channel estimation & channel aging lead to

- Mismatched LLR computation
- SNR degradation

Neural demapper (symbol-wise)

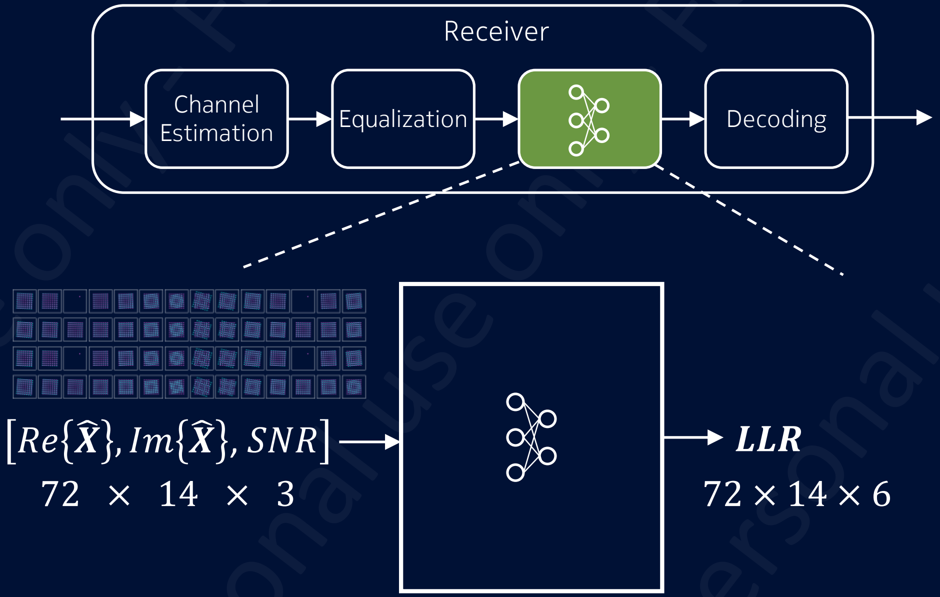


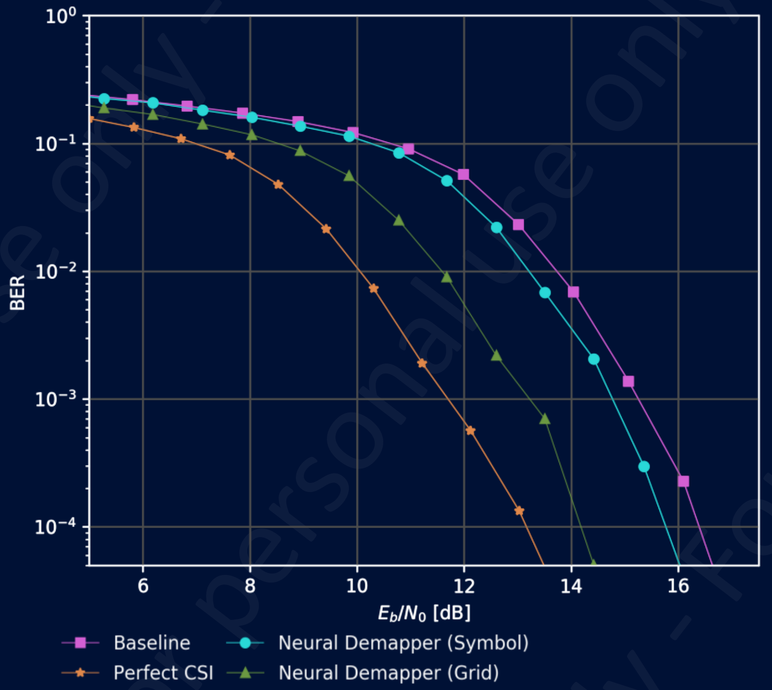


Learns grid position-dependent statistics for better LLR computation

NOKIA Bell Labs

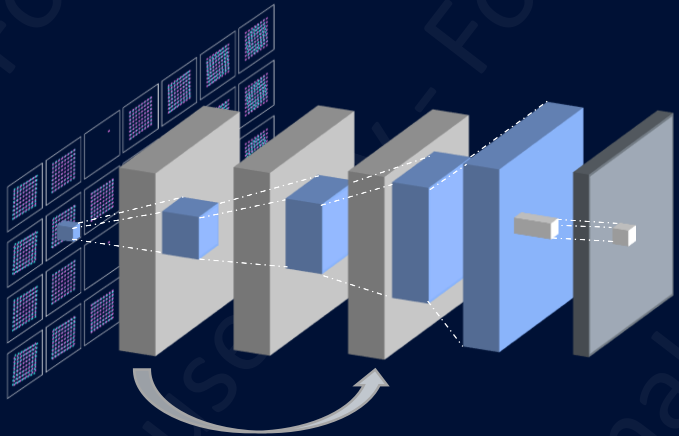
Neural demapper (grid-wise)



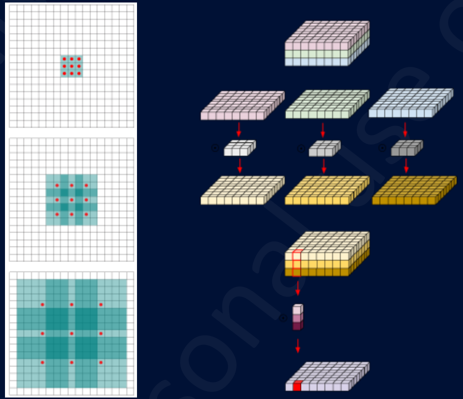


Leverages pilots and data to compensate for channel aging and mismatched LLR computation

Neural network architecture is key to success



Fully convolutional ResNet



Dilated separable convolutions

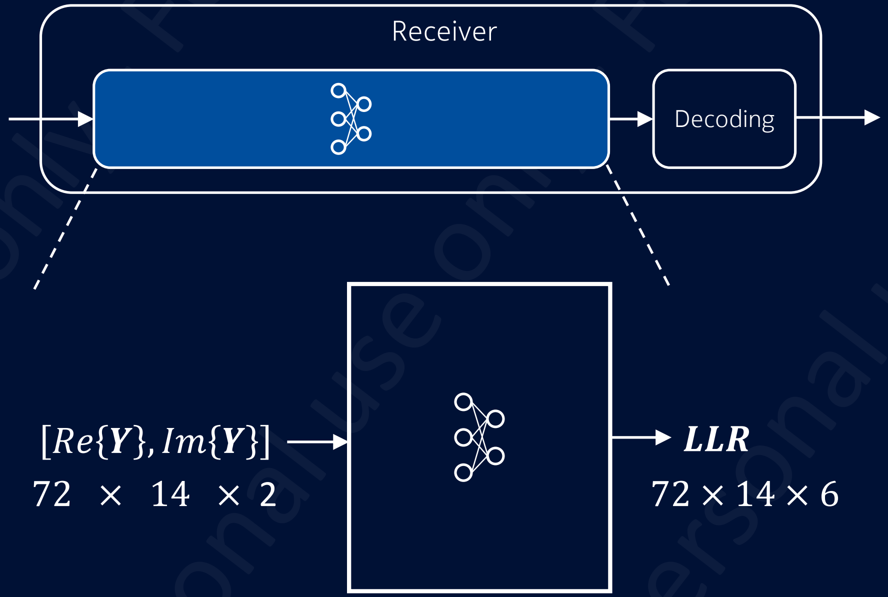
M. Honkala, et al., “DeepRX: Fully Convolutional Deep Learning Receiver”, arXiv2005.01494

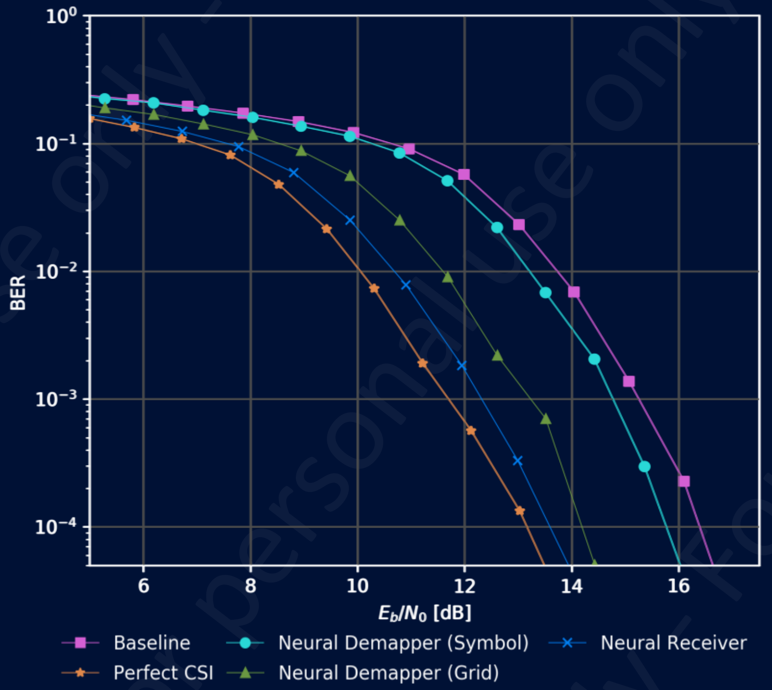
Each output value has a receptive field spanning the entire resource grid

H. Talebi et al., “Learned perceptual image enhancement”, arXiv:1712.02864
F. Yu et al., “Multi-scale context aggregation by dilated convolutions”, arXiv:1511.07122
https://medium.com/@zurister/depth-wise-convolution-and-depth-wise-separable-convolution-37346565d4ec

NOKIA Bell Labs

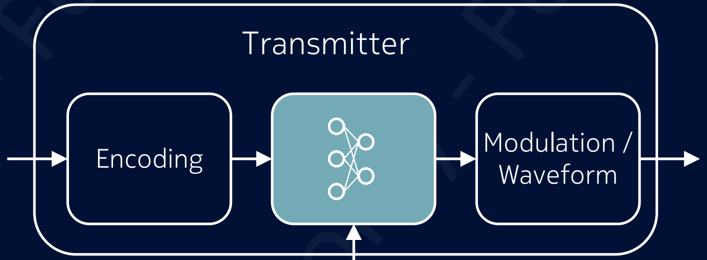
Neural receiver





Data-aided channel estimation, equalization, and demapping for unprecedented performance

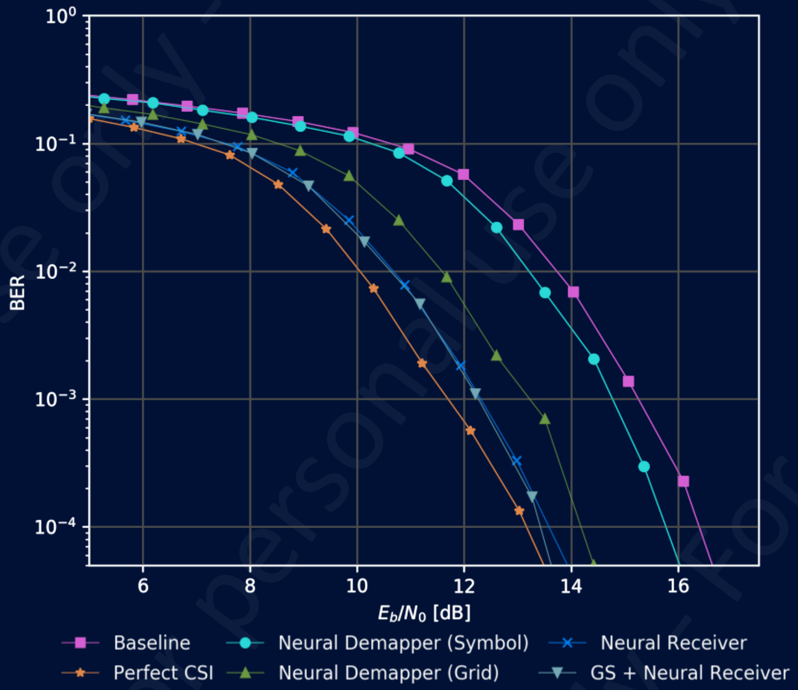
End-to-end learning with Neural receiver



64x2 trainable weights

$$\mathcal{C} = \frac{\tilde{\mathcal{C}} - \frac{1}{64}\sum_{c\in\tilde{\mathcal{C}}} c}{\sqrt{\frac{1}{64}\sum_{c\in\tilde{\mathcal{C}}}|c|^2 - \left|\frac{1}{64}\sum_{c\in\tilde{\mathcal{C}}} c\right|^2}}$$

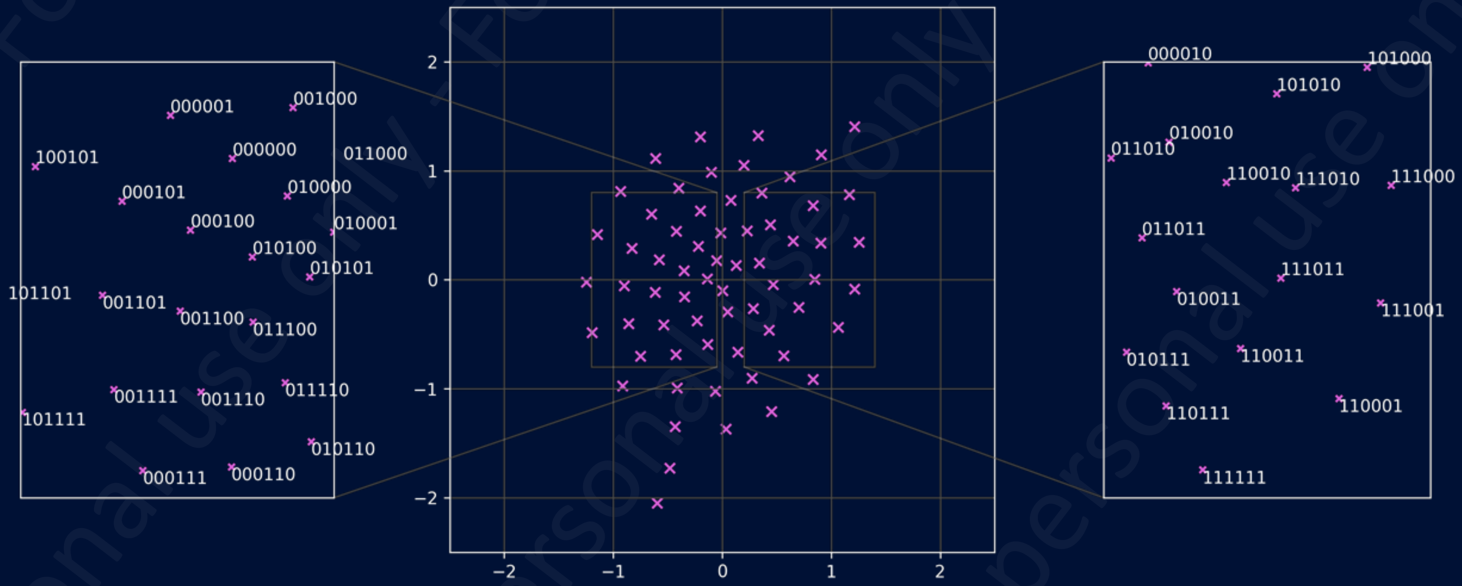
Zero-mean unit energy constellation



End-to-end learning enables pilotless transmissions without performance loss

NOKIA Bell Labs

Learned constellation for pilotless communication



F. Ait Aoudia, J. Hoydis, "End-to-end Learning for OFDM:
From Neural Receivers to Pilotless Communication", arXiv2009.05261

How could this be standardized?

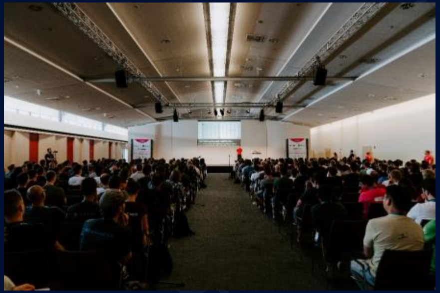
Important research topics for end-to-end learning

- New waveforms for new spectrum
- Learning for systems with (extreme) hardware constraints
- Joined communications + X
- Signals conveying a few bits of information
- Application-specific end-to-end learning
- Semantic communications
- Decentralized & federated learning
- Transfer & meta learning

The next frontier:
Protocol learning

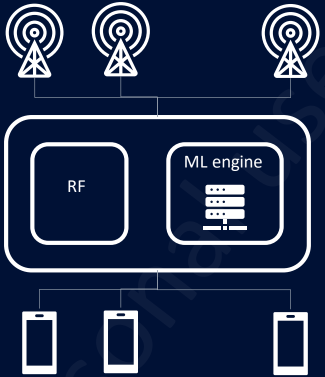
Emerging a RAN protocol

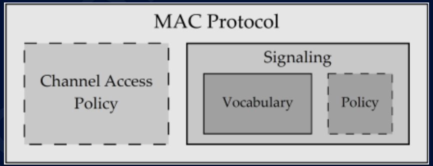
3GPP Way



VS

ML Way





A. Valcarce and J. Hoydis, “Towards joint learning of optimal MAC signaling and wireless access channel access”, *arXiv:2007.09948v2*

Can we learn a MAC protocol?

Thank you!

NOKIA Bell Labs