RADIO PROTOCOL EMERGENCE

Dr. Alvaro Valcarce
Head of department on Wireless AI/ML
Nokia Bell-Labs, France

ROHDE & SCHWARZ

Make ideas real

AI-AI: Radio Stack Learning

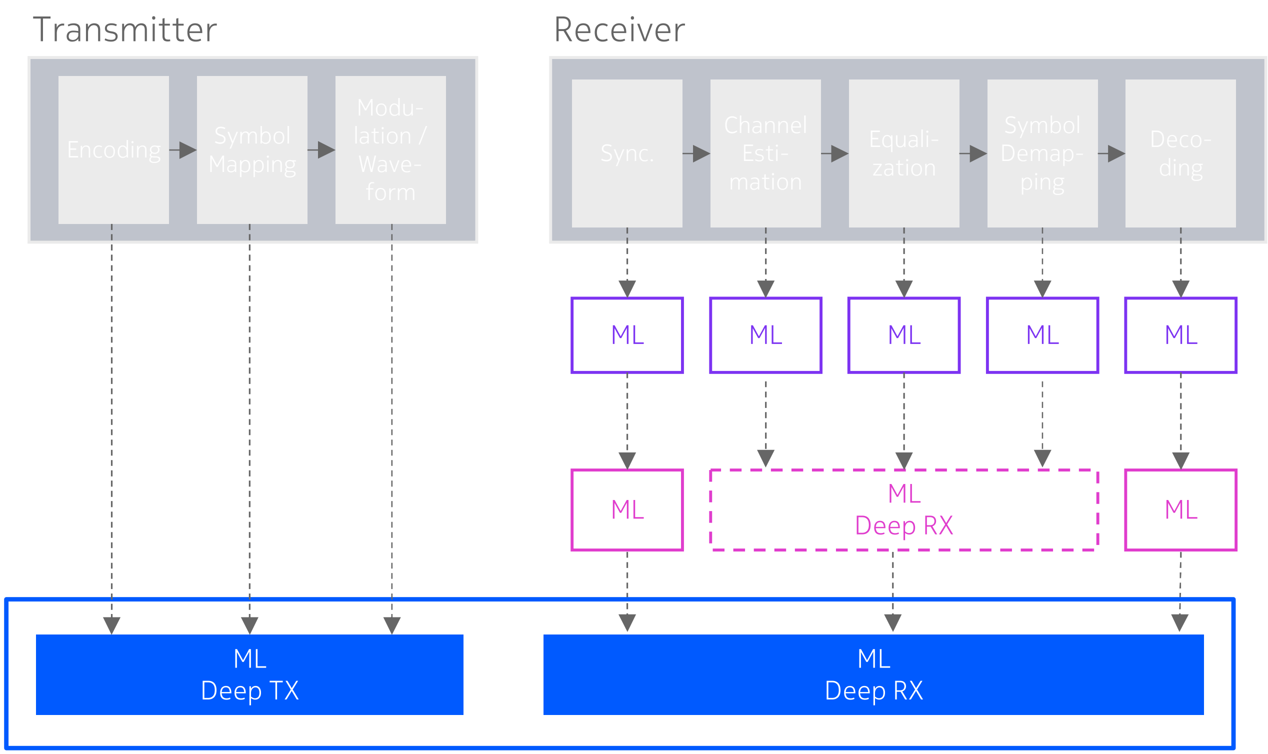
Alvaro Valcarce Rial

November 20th 2024

Venue: Mobile Test Summit (MTS) R&S

6G AI native Air Interface

The roadmap to learning radios



5G

The classical architecture

5G-Advanced phase 1

ML replaces/enhances individual processing blocks

5G-Advanced phase 2

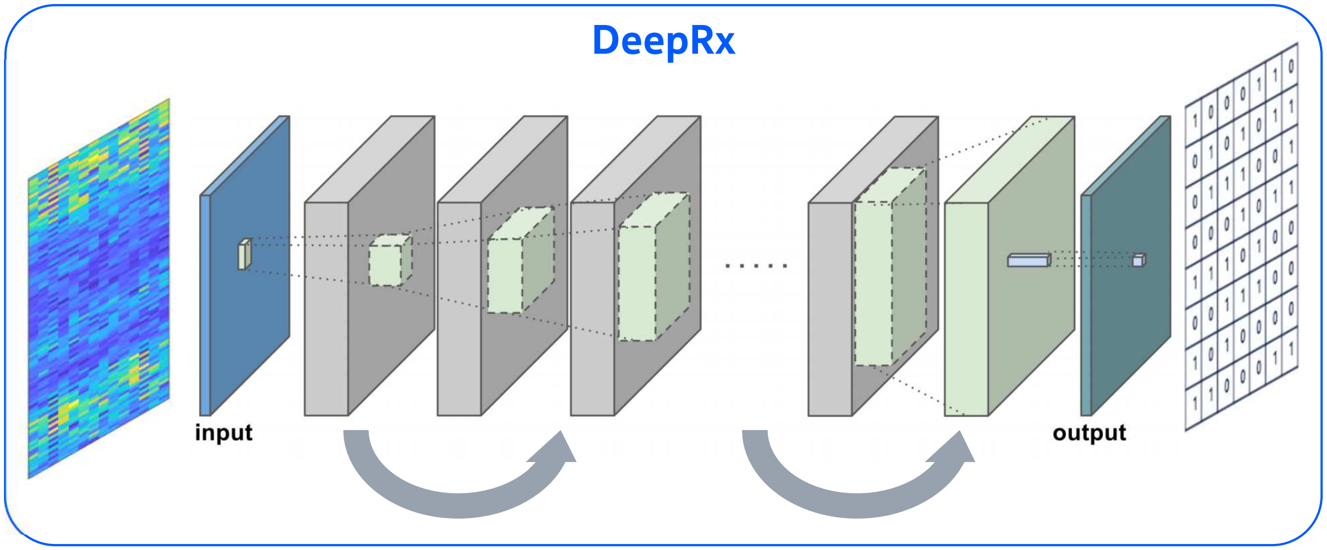
ML replaces multiple processing blocks

6G

ML designs part of the PHY itself

The L1 AI-AI: Learning a receiver

ML in L1: Learning a Receiver



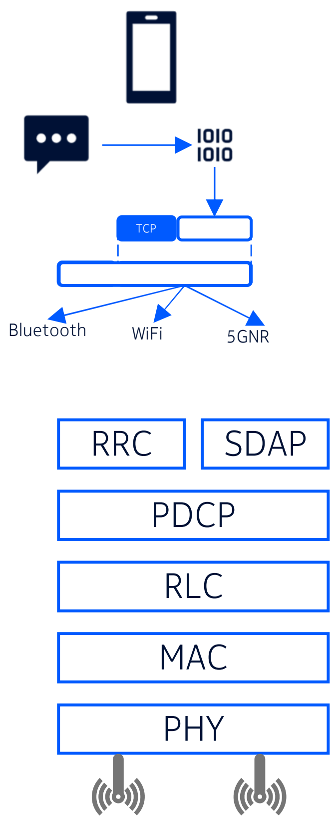
- It is possible to enhance detection accuracy in the receiver by implementing it as a neural network
- OFDM reception is quite similar to *image segmentation*
- **Main benefits:**
  - A learned receiver, such as DeepRx, can operate with fewer reference signals
  - Other variants can operate under severe distortion, which can help in enhancing coverage in 6G

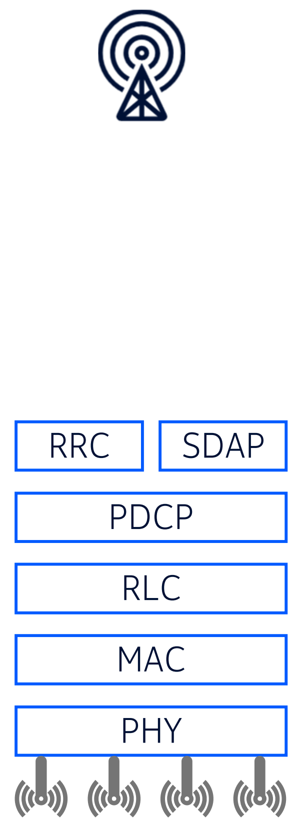
Picture modified from: H. Talebi, and P. Milanfar. "Learned perceptual image enhancement." in Proc. IEEE ICCP, 2018.

The L2 AI-AI: Protocol learning

Radio protocols in cellular networks

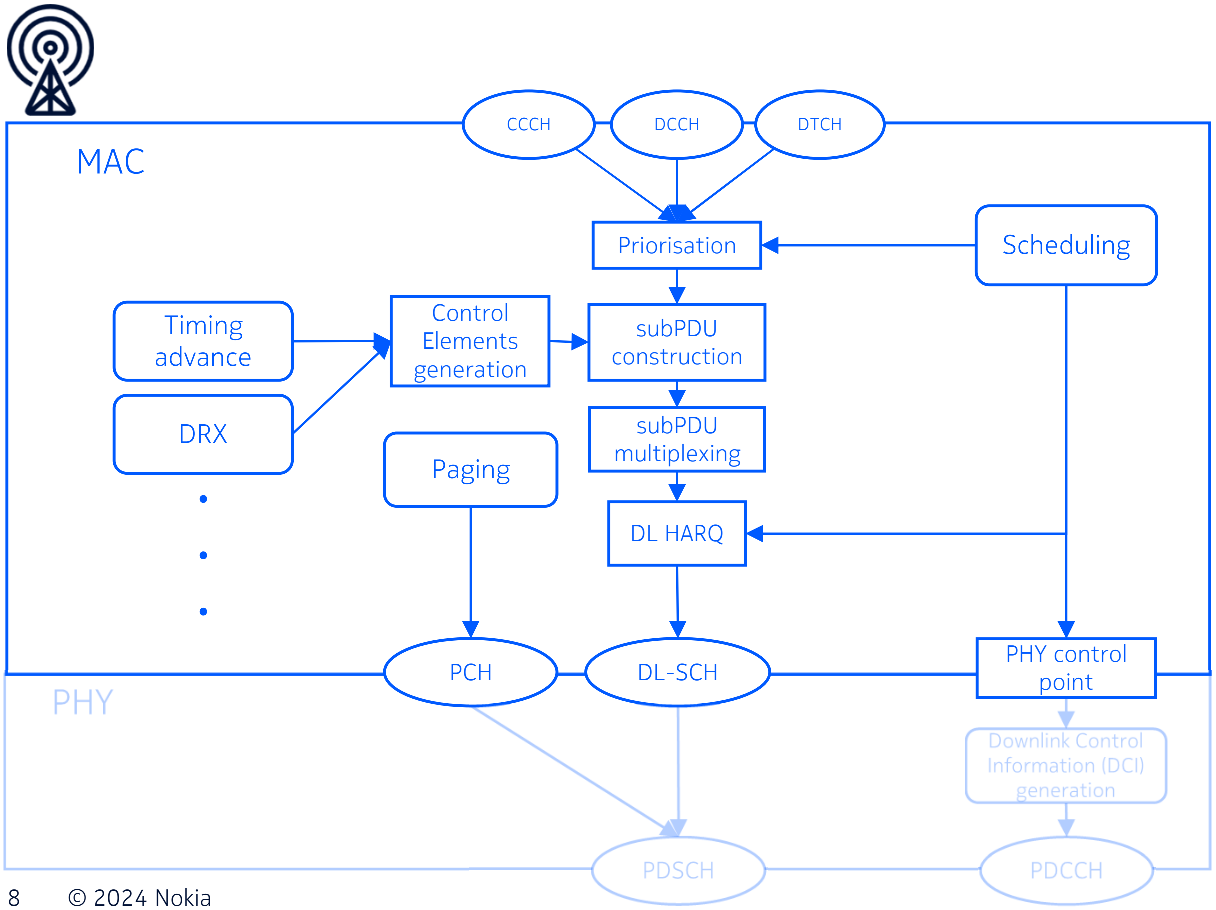
A recap





Radio protocols in cellular networks

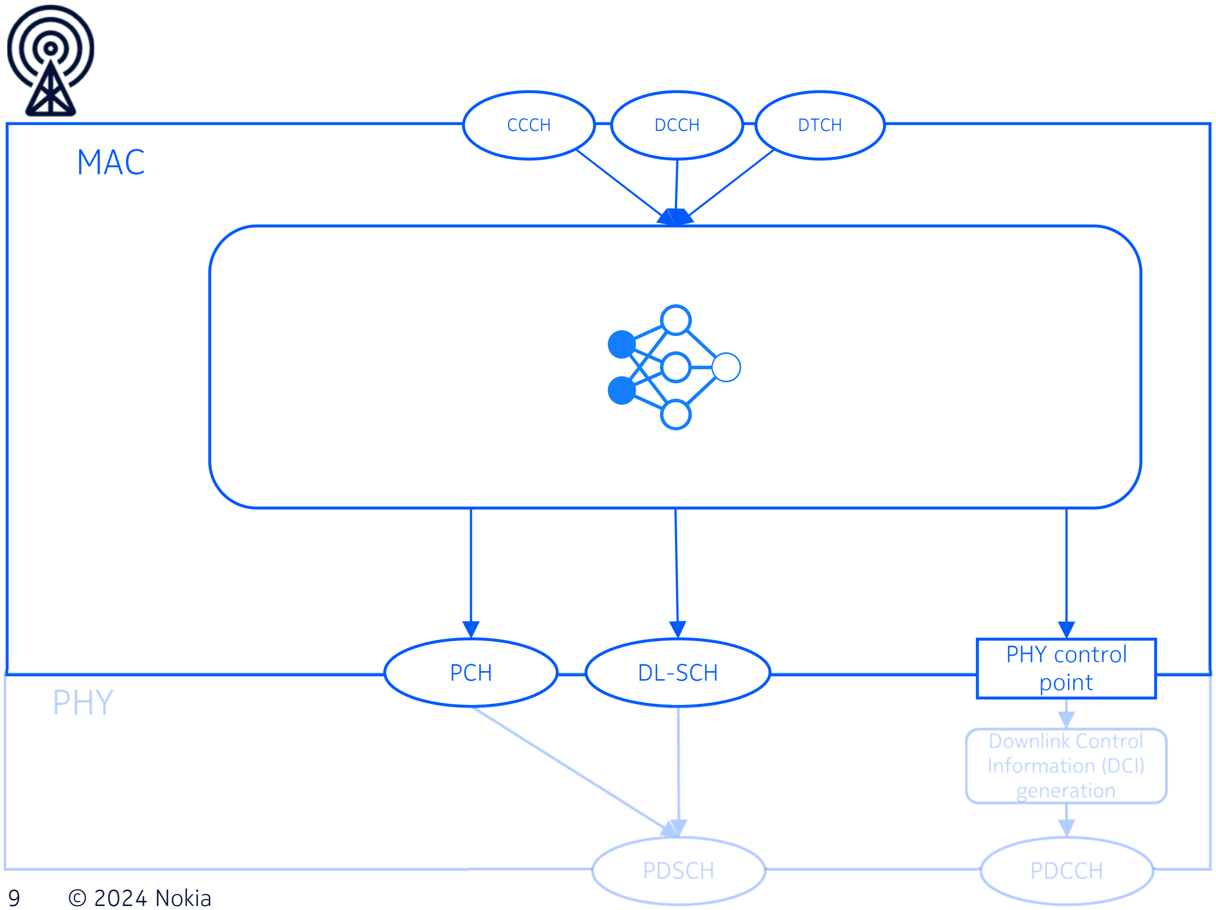
The 5GNR MAC & AI opportunities



- Growing number of functionalities
- Heuristic features (non provably optimal)
- Data abundance
- Relaxed real-time requirements

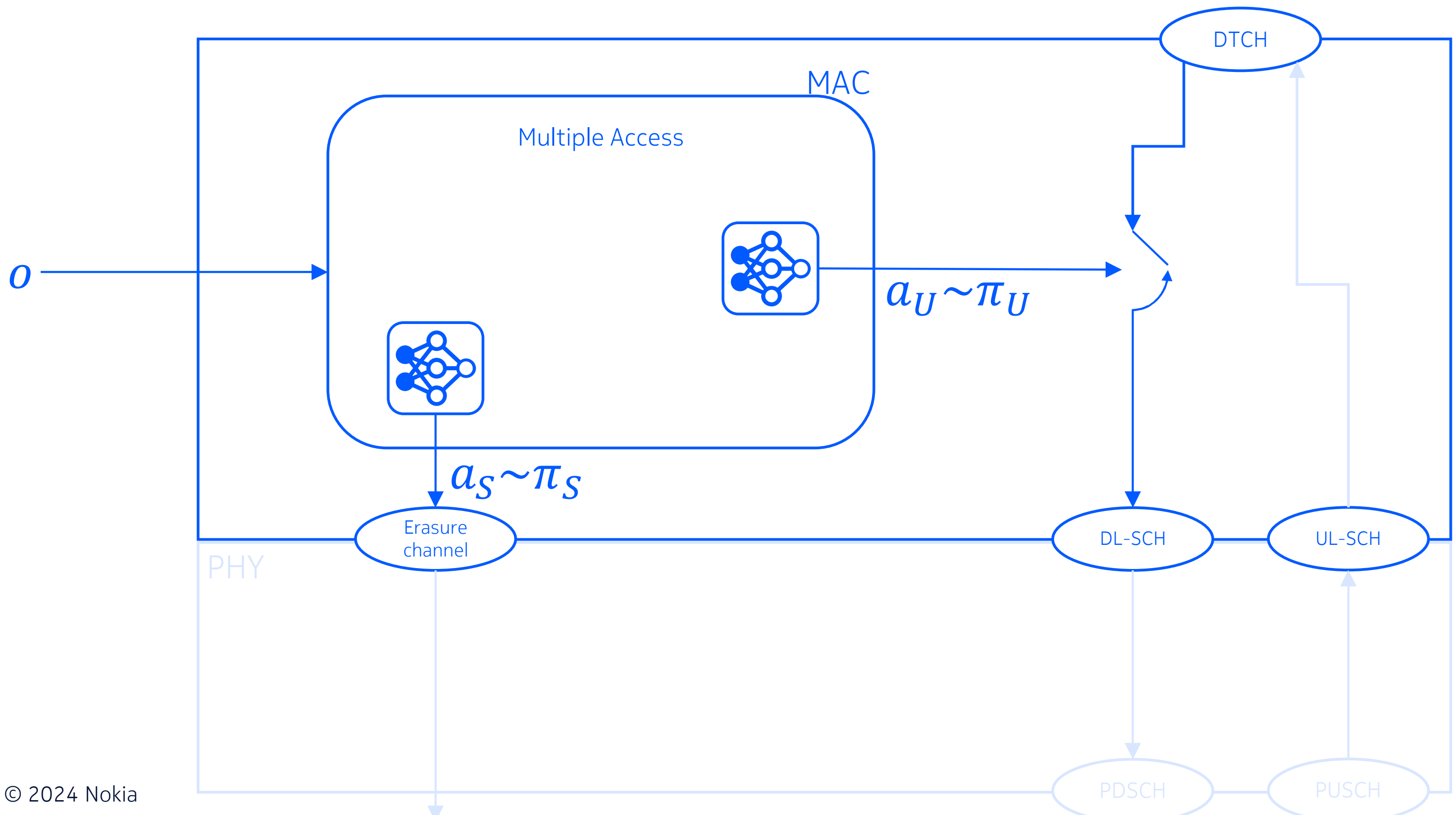
Radio protocols in cellular networks

The learned MAC



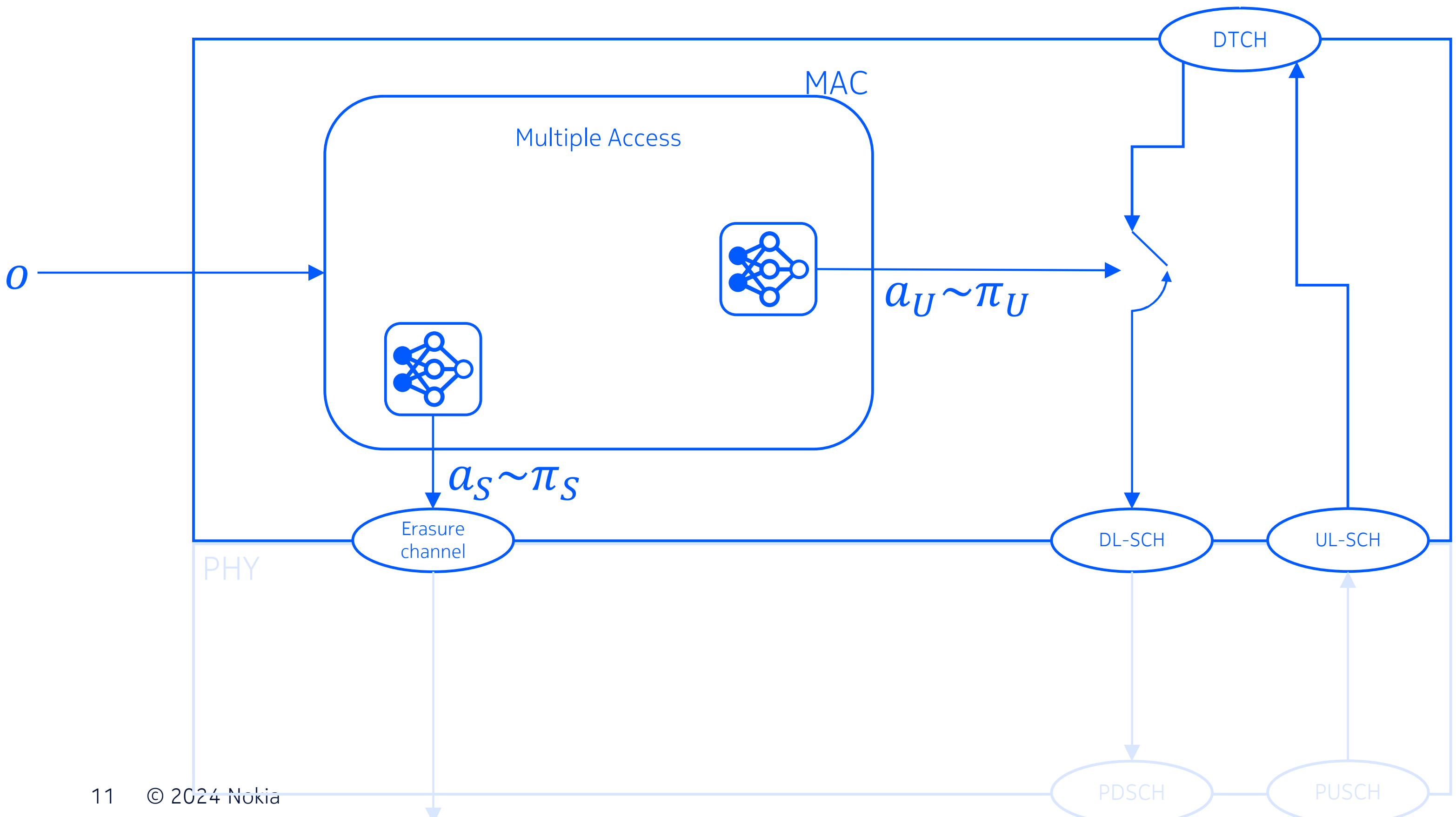
Learning a radio MAC layer

Feature-focused protocol learning



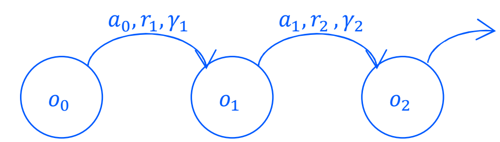
Learning a radio MAC layer

Feature-focused protocol learning



$$a^{BS} = \begin{pmatrix} a_S \\ a_U \end{pmatrix}$$

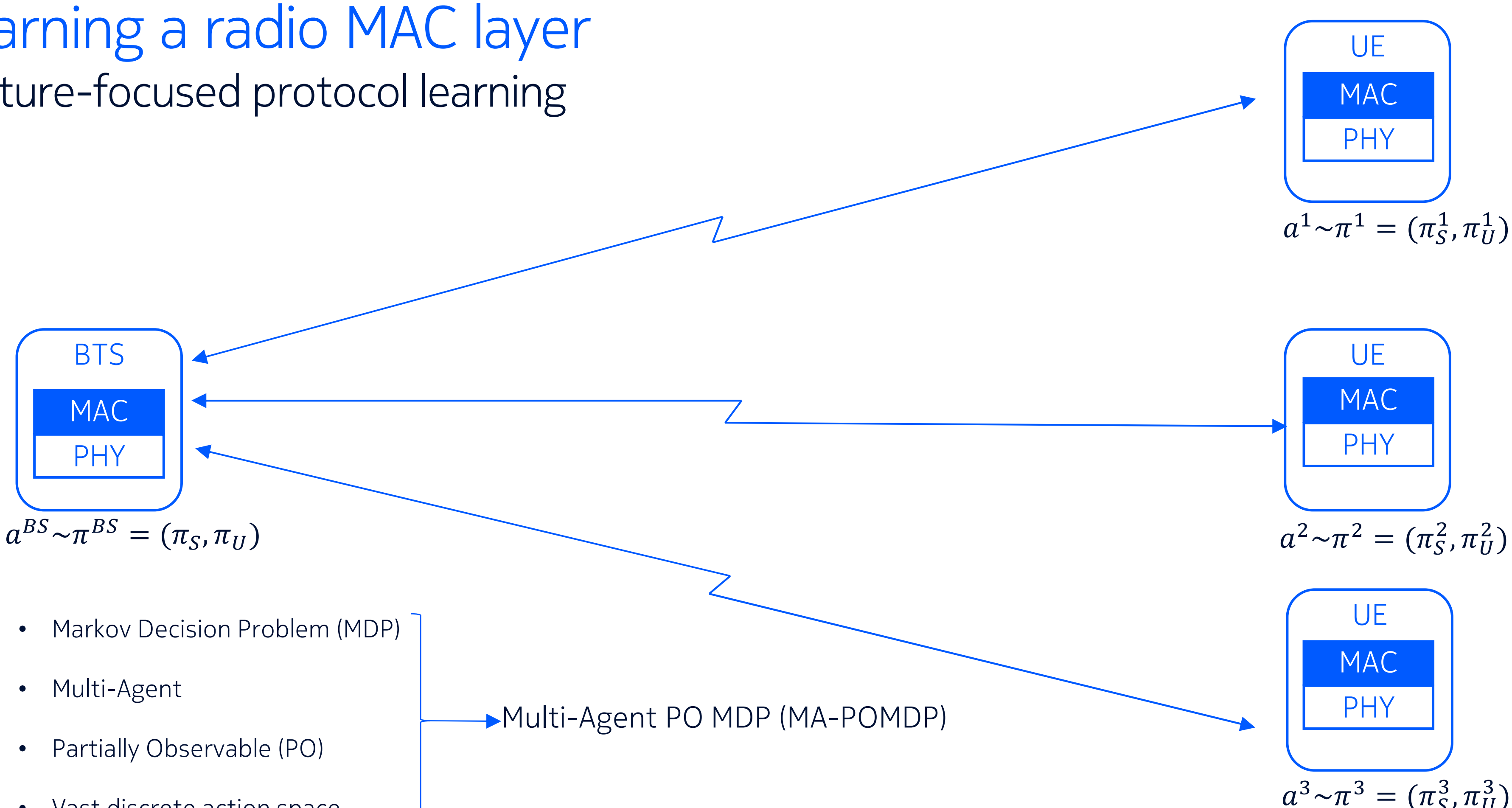
Large signaling space



Markov Decision Process (MDP)

Learning a radio MAC layer

Feature-focused protocol learning



- Markov Decision Problem (MDP)
- Multi-Agent
- Partially Observable (PO)
- Vast discrete action space

→ Multi-Agent PO MDP (MA-POMDP)

Emerging a radio MAC feature
With MADDPG

Centralized training with decentralized execution

Training gradient for the policy π_i (parametrized by θ_i):

$$\nabla_{\theta_i} J(\pi_i) = \mathbb{E}_{x,a\sim D}[\nabla_{\theta_i}\pi^i(a^i|o_i)\nabla_{a^i} Q_\pi^i(x,a^0,\dots,a^N)|_{a^i=\pi^i(o_i)}]$$

Training loss of centralized action-value functions:

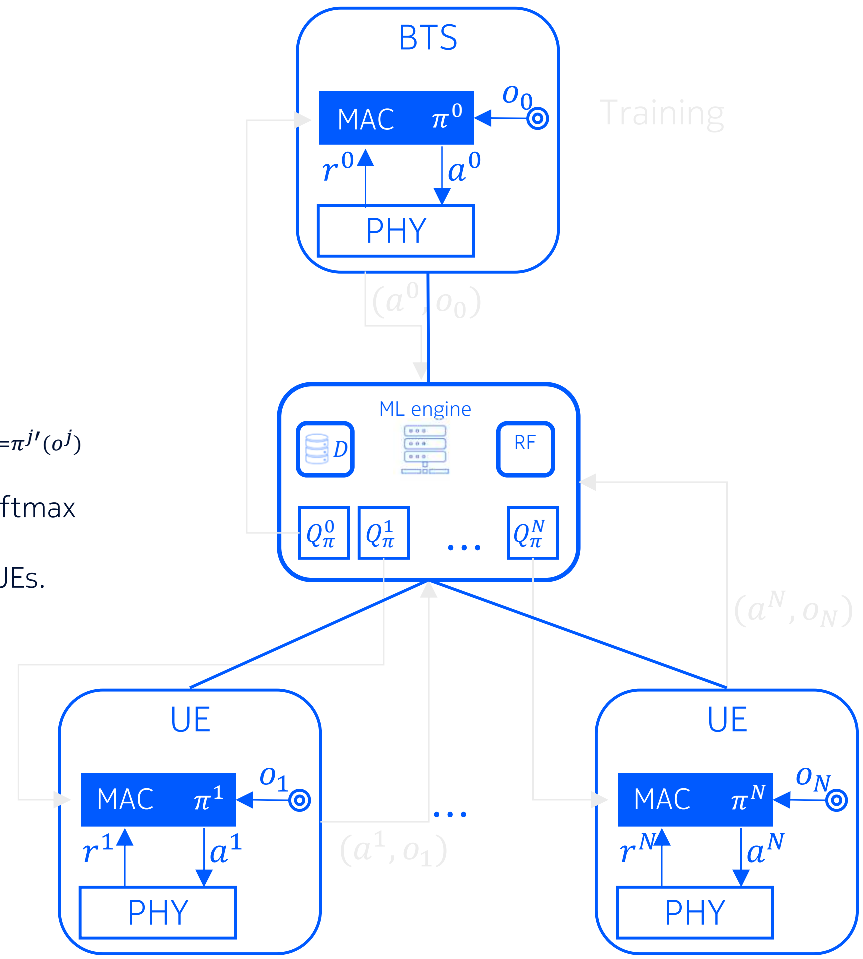
$$\mathcal{L}(\theta_i) = \mathbb{E}_{x,a,r,x'}[(Q_\pi^i(x,a^0,\dots,a^N) - y)^2] \text{ with } y = r_i + \gamma Q_{\pi'}^i(x', a^{0\prime},\dots,a^{N\prime})|_{a^{j\prime}=\pi^{j\prime}(o^j)}$$

Signaling messages are soft approximations to discrete messages using Gumbel-Softmax

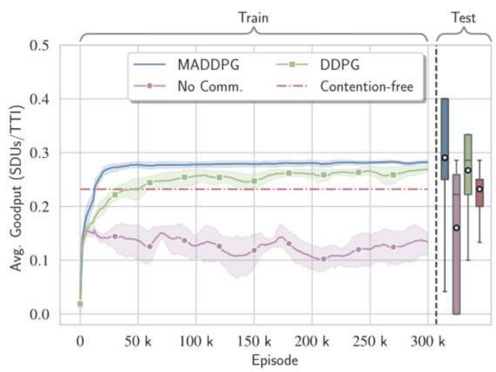
General-purpose protocol: A single UE policy is trained, and regularly copied to the UEs.

Goal-specific protocol: UE-specific policies can be trained.

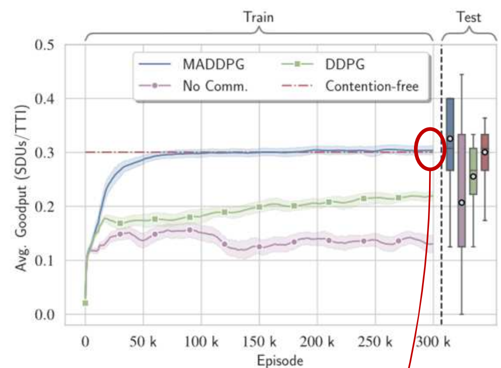
Population-based training: Survival of the fittest protocol



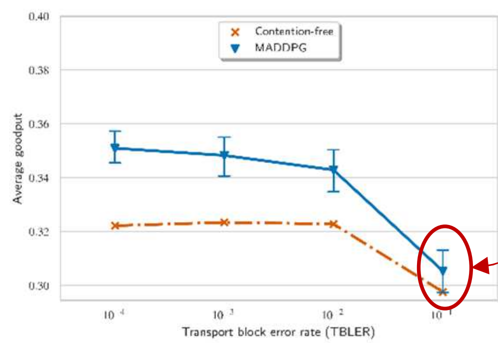
Emerging a radio MAC feature
With MADDPG



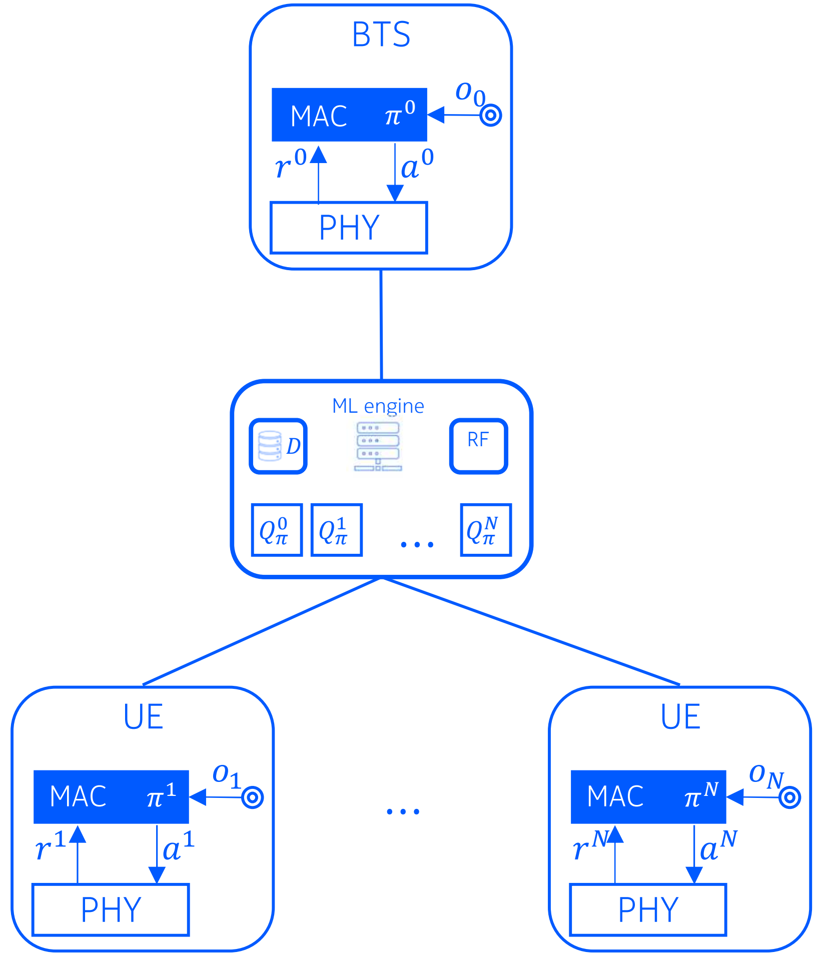
1 SDUs per UE



2 SDUs per UE



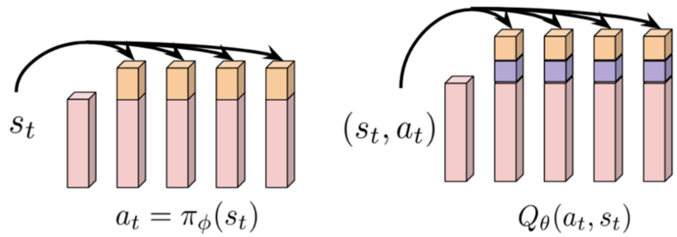
Emerged protocols outperform baseline in different scenarios



Emerging a radio MAC feature

Scalability tricks

D2RL (Deep Dense Architectures in RL)



"D2RL: Deep Dense Architectures in Reinforcement Learning", S. Sinha et al, 2020.

JSRL (Jump Start Reinforcement Learning)

- Pre-training
- Knowledge transfer through a guide policy

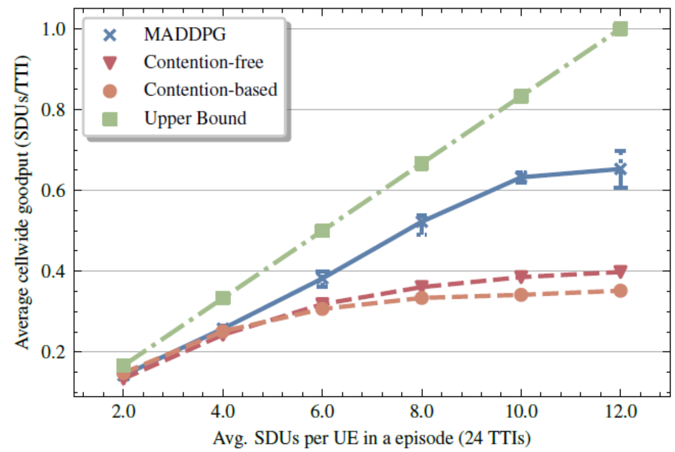
I. Uchendu, T. Xiao, Y. Lu, Banghua Zhu, M. Yan, J. Simon, M. Bennice, C. Fu, Cong Ma, J. Jiao, S. Levine, and K. Hausman. 2023. "*Jump-start reinforcement learning*". In Proceedings of the 40th International Conference on Machine Learning (ICML'23), Vol. 202.

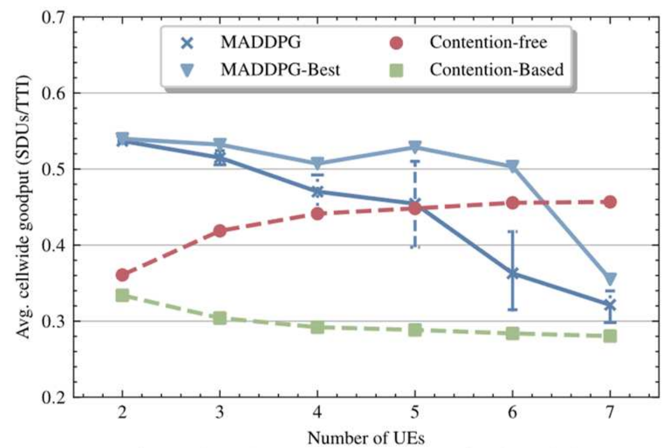
Reward standardization

```
Initialize:
    rewards_sum = 0
    rewards_squared_sum = 0
    num_rewards = 0

For each reward r:
    rewards_sum += r
    rewards_squared_sum += r * r
    num_rewards += 1

    mean = rewards_sum / num_rewards
    variance = rewards_squared_sum / num_rewards - mean * mean
    std_dev = sqrt(variance + epsilon)
    standardized_reward = (r - mean) / std_dev
```





M. P. Mota, A. Valcarce and J. -M. Gorce, "*Scalable Joint Learning of Wireless Multiple-Access Policies and their Signaling,*" 2022 *IEEE 95th Vehicular Technology Conference: (VTC2022-Spring)*, Helsinki, Finland, 2022, pp. 1-5, doi: 10.1109/VTC2022-Spring

The AI-based Air Interface (AI-AI)

Key takeaways

Customization

Customer/application/site-specific radio protocol stacks

Performance

KPI-led system design

Automation

Create features rapidly, reproducibly and without human intervention

Science

Explore exotic solutions (protocols and policies)

The AI-based Air Interface (AI-AI)

Open challenges

Standardization

Data collection, training metrics, enablers, model management

Testing

Pass/fail criteria, debugging methods

Computational complexity & scalability

Training/inference time, real-time requirements

Interpretability

black-box models, interpreting large models, privacy

NOKIA BELL LABS

MOBILE TEST SUMMIT 2024

Test. Measure. Innovate

THANK YOU VERY MUCH

ROHDE & SCHWARZ

Make ideas real