Aalto University
School of Electrical
Engineering

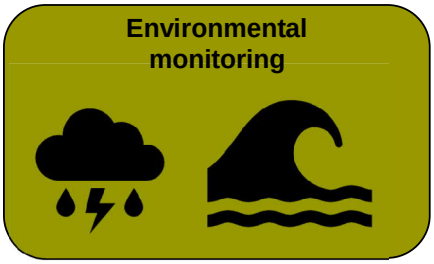
MIMO Multicarrier ISAC: model-based reinforcement learning for waveform optimization and resource allocation

Visa Koivunen, Department of Information and Communications Engineering
Aalto University

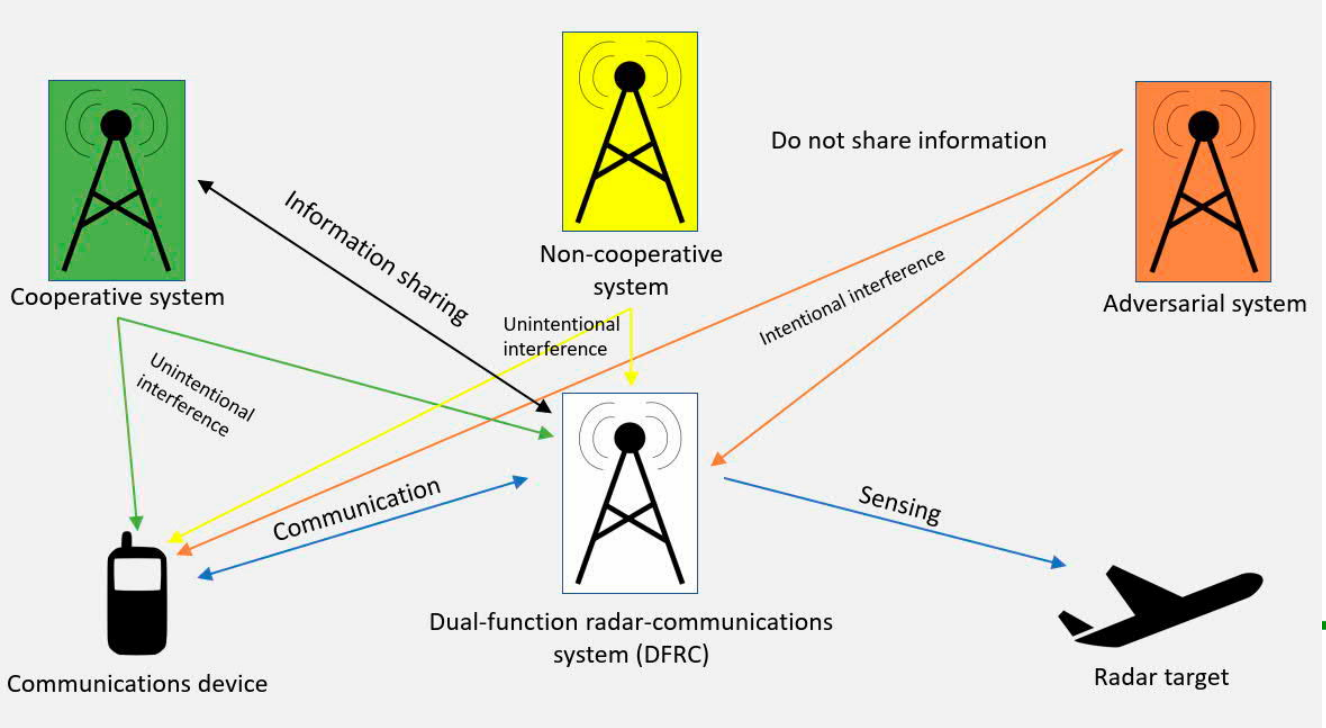
Joint work with Petteri Pulkkinen
Support by SAAB, Business Finland, MPKS, EU INSTINCT
ACK: ISAC collaboration with Mikko Valkama, Henk Wymeersch, and Furkan Keskin

Integrated Sensing & Communications (ISAC)

- **Paradigm shift:** integration of (radar) sensing and communications into unified ISAC systems
  - Both functions potentially operated on the same hardware, antennas, spectrum, and waveforms
  - Multifunction transceiver, circuit and large aperture antenna resources
- Reduce **the system cost** and **power consumption** while also **mitigating congestion** in the radio spectrum
- Oher sensing modalities: video, lidar…
- A key new technology envisioned for the emerging **6G wireless networks** and Wi-Fi



ISAC: co-designed, cooperating systems sharing awareness



Integrated Sensing and Communication (ISAC) drivers

Cooperation and Co-design of sensing and communication for mutual benefit

- **Technology Convergence:** Parallel RF convergence, Multifunction HW, large aperture antenna arrays, shared transceiver HW and antenna resources
- Understanding both wireless comms and sensing/radar tasks is crucial
  - More sensing and radar expertise needed
- ISAC has high potential to be one of the transformative technologies in 6G
- Potential 6G enablers studied include RIS, and THz (300GHz-3THz) comms will not be in 6G at this point
  - Multicarrier OFDM modulation and large aperture MIMO will be used,
  - Delay-Doppler domain OTFS is not included at this point
- FR3 changed the ball game (mmWave not used that much)
- Sensing as service, Sense to communicate, communicate to sense, multifunction operation

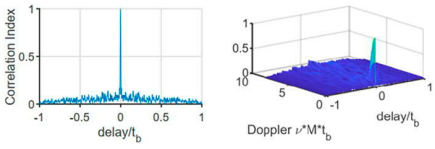
ISAC and 6G thoughts

- Standardization is needed to ensure significant investments, interest from devices manufacturers
- There must be business models for sensing and sensing related services. Financial incentives are needed for operators to invest on sensing
- In wireless standards, the transmitter is usually standardized
  - How to include AI in TX of heterogeneous devices, users, protocols, awareness
- If TX is standardized, defining a variety of sensing waveforms may be difficult
  - different sensing tasks have different KPIs requiring different waveforms, optimization, adaptation, resource allocation
- Means and protocols for sharing situational awareness, radio environment, channel state information, interference, pilot signaling, location information
- Similar Technologies find applications in multifunction Wi-Fi, EW, and radars

Typical Radar Sensing tasks

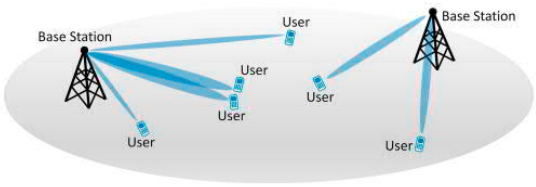
- Different optimality criteria and requirements for waveforms, transceiver algorithms and operational parameters for different sensing tasks
  - Target search and detection, $\boldsymbol{p}_{FA}$, $\boldsymbol{p}_{D}$, FDR control, track before detect, estimating noise and interference statistics (CFAR detectors)
  - Estimation, MI maximization, CRB, MMSE: Target parameter estimation
    - Determining target kinematic state: range, direction, velocity and bearing
    - Target tracking, dealing with multiple maneuvering targets
  - Target recognition, HRRPs, micro-doppler signatures, SAR/ISAR imaging
  - mitigating interference, waveform diversity and optimization, power allocation
  - Managing multifunction operation, scheduling, time-budget and resource management
  - Dealing with unwanted returns (clutter), multipath, unintentional and intentional (jamming) interferences
  - weather sensing (polarization), change-point or anomaly detection, spatio-temporal RF field modeling

$$SNR = \frac{P_S}{P_N} = \frac{P_T G_T G_R \lambda^2 \sigma}{(4\pi)^3 R^4 k T_0 B F_n L}$$



ISAC waveform design and resource management

Multibeam MIMO techniques – W. Hong *et al., 2017*



Giraffe 4A radar antenna – Saab.com

Objectives:

- Data rate/Sum Rate/throughput
- Capacity, Mutual Information
- Reliability (low outage, Quality of Service)
- Low Latency (for control, robotics)
- Energy efficiency (lower power consumption)

Objectives:

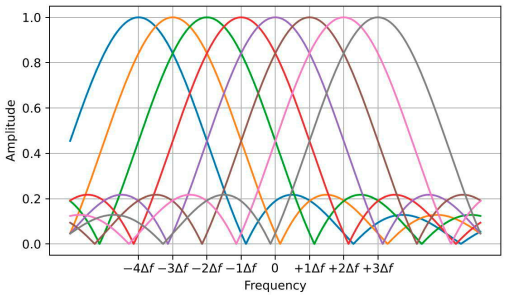
- Detection probability (detect targets with false alarm control)
- Estimation Performance (MMSE, CRLB, localization, track quality)
- Information theory: Rate-Distortion, Mutual Information, bounds
- Coverage (radar equation R^4 attenuation), Energy efficiency
- Resolution (distinguish closely spaced targets, BW and aperture)

ISAC design considers both **radar and communications objectives!**

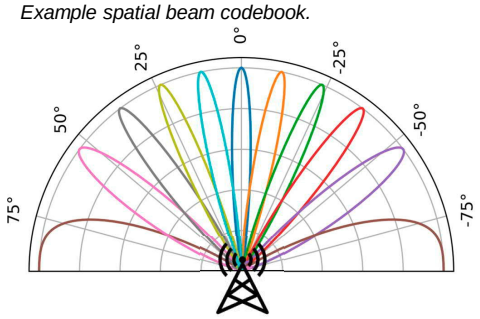
- Objectives may be conflicting → balancing and trade-offs needed
- E.g., via multi-objective or constrained optimization

Multicarrier and multiantenna ISAC Systems

Example set of orthogonal subcarriers.



Example spatial beam codebook.



Multicarrier waveforms (e.g., OFDM)

- Basis of most wireless comms standards (4G, 5G, emerging 6G, WiFi), 6G standardization has selected OFDM
- Many different variants, including OTFS
- Used also in passive and active radars

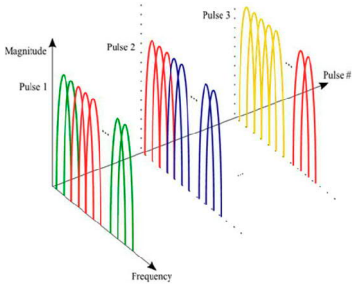
Multi-antenna configurations

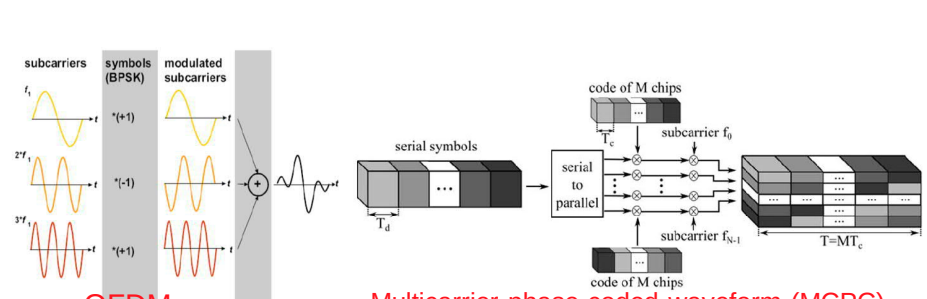
- Enables spatial beam steering and multiplexing
- Separation of users/targets
- Interference management
- Beamformers, precoders, wireless localization, mapping

→ Combination offers significant flexibility for ISAC

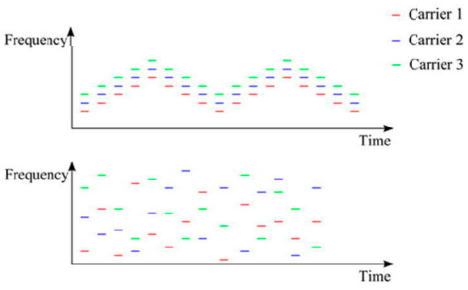
Multicarrier waveforms for ISAC

- Multicarrier Sensing Benefits: Frequency diversity, interference avoidance, resource allocation, estimation of Doppler from single pulse, range resolution
  - Existing comms HW design knowledge can be used
- Main problem: high peak to average power ratio
  - Inefficient use of amplifiers, may reduce the coverage
- Often combined with multiantenna systems (MIMO-OFDM)





OFDM



Multicarrier phase-coded waveform (MCPC)

Multiobjective ISAC waveform optimization problem

$$\min_{R} \quad (f_{\text{comm}}(R), f_{\text{sense}}(R))$$

$$\text{s.t.} \quad \mathbf{h}(R) \leq \mathbf{0},$$

- $f_{comm}\,(R)$ is a communication metric such as achievable rate, QoS, latency, BER, …
- $f_{sense}\,(R)$ is a radar sensing metric such as MI, P_{D} with false alarm constraint P_{FA}
- $\mathbf{h}(R)$ represents constraints sun as total power, desired SINR, AF similarity, PAPR
- Problem can be scalarized with subobjective weights, objectives moved as constraints, or finding Pareto solutions where improvement in an objective cannot be done without worsening others

ISAC waveform optimization problem formulations

- **Radar centric** approach: maximize radar sensing performance (e.g. MI) under constraints on minimum tolerable performance for wireless transmission
- **Communication centric** approach: maximize communications performance under constraints on minimum tolerable radar performance (detection, estimation)
- **Multiobjective** optimization (e.g Liu et al, TR-SP 18)
- Example radar centric objectives:
  - Frequency and power allocation for optimal detection
  - Target parameter CRB minimization
  - Beam allocation for target track quality
- Other constrains can be imposed
  - Constant modulus, per antenna power, total power
  - Similarity constraint on AF or waveform
  - Low cross correlation for MIMO waveforms
  - Antennas resources

maximize	*Probability of Detection*
subject to	*Probability of False Alarm Constraint*
	Desired Data Rate for Communications Users
	Total Radar Power Budget

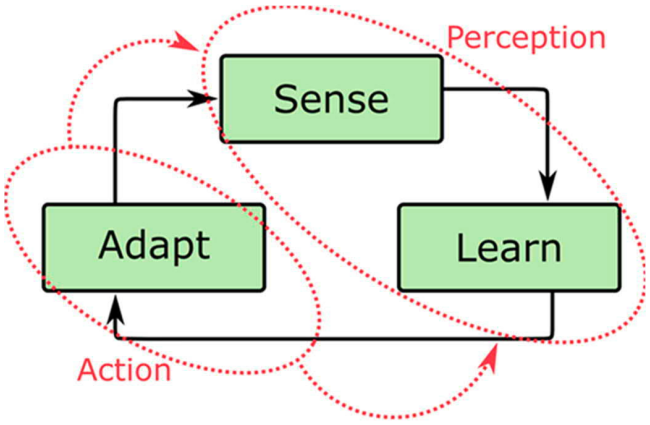
minimize	*CRB*
subject to	*Desired Data Rate for Communications Users*
	Total Radar Power Budget
	Subcarrier Power Ratio

Structured optimization drawbacks in ISAC

- Structured optimization methods assume that all necessary model parameters needed in optimization are known or estimated reliably without uncertainty.
- In ISAC, they rely on ideal models and assumptions about the state of the radio environment, acquired data, hardware, and user and target scenarios.
- Modeling a large-scale distributed radio system with *highly dynamic propagation environments and spectrum usage patterns* using realistic and rigorous mathematical modeling may not be feasible in practice.
- Typically structured optimization methods employ high complexity batch algorithms and may not be suitable for dynamic scenarios.
- Moreover, no learning from the past experiences takes place

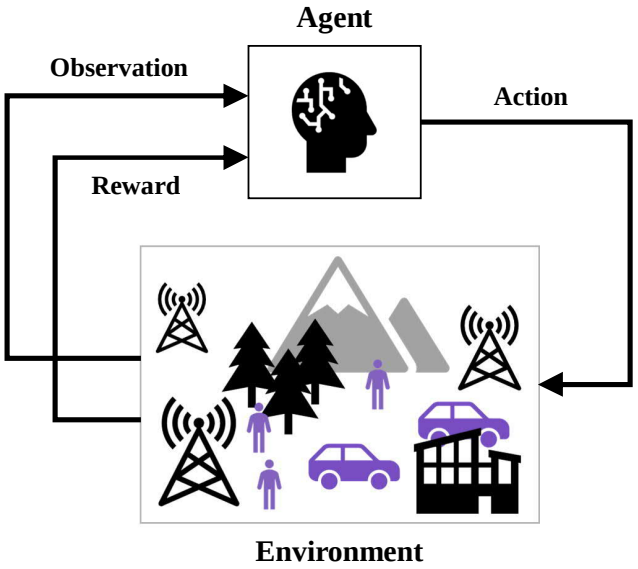
Cognitive processing in learning

- System displays intelligence, adapting its operation and its processing in response to a changing operational environment
- Cognitive system learns to adapt operating parameters over extended time periods. It has memory and capability learn from past experiences and autonomously improve its performance
- Online learning based on situational awareness
- Reinforcement Learning and Learning Model-Based Control provide a framework for ISAC with cognitive capabilities
  - Learning dynamic models for state transitions and rewards and model prediction



Reinforcement Learning (RL)

- Reinforcement learning (RL) algorithms learn optimal policies through interactions with the operational environment
- An agent interacts with an environment by taking different actions. It observes the environment state and receives a scalar reward and takes the next action.
- Addresses the inherent complexity, modeling deficits and highly dynamic nature of ISAC optimization and adaptation problems
- **Key challenges:**
  - High-dimensional decision spaces
  - Data inefficiency
  - Partial observability
  - Lack of interpretability

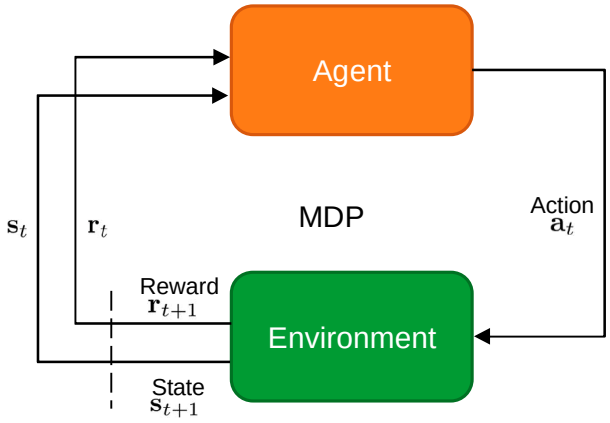


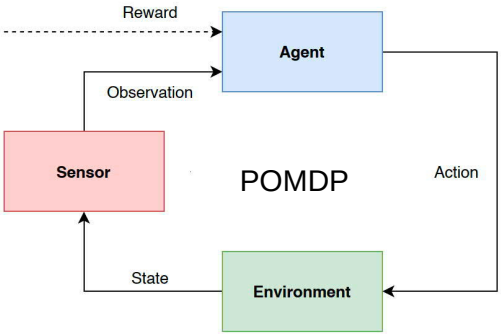
Underlying model: Markov Decision Process (MDP)

- State space: $\mathbf{s} \in \mathcal{S}$
- Action space: $\mathbf{a} \in \mathcal{A}$
- Dynamics model: $\mathrm{T}(\mathbf{s}_{t+1}, r_{t+1}|\mathbf{s}_t, \mathbf{a}_t) = \Pr(\mathbf{s}_{t+1}, r_{t+1}|\mathbf{s}_t, \mathbf{a}_t)$
- Expected reward: $r(\mathbf{s}_t, \mathbf{a}_t, \mathbf{s}_{t+1}) = \mathbb{E}[r_{t+1}|\mathbf{s}_t, \mathbf{a}_t, \mathbf{s}_{t+1}]$
- Policy
  - Deterministic: $\mathbf{a}_t = \pi(\mathbf{s}_t)$
  - Stochastic: $\pi(\mathbf{a}_t|\mathbf{s}_t) = \Pr(\mathbf{a}_t|\mathbf{s}_t)$
- Action-value function (Q-value): Value of action

  $Q^{\pi}_{\gamma}(\mathbf{s}, \mathbf{a}) = \mathbb{E}_{\pi}\left[\sum_{i=t}^{\infty} \gamma^{i-t} r_{i+1} | S_t = \mathbf{s}, A_t = \mathbf{a}\right]$
- Value function: Discount factor [0, 1)

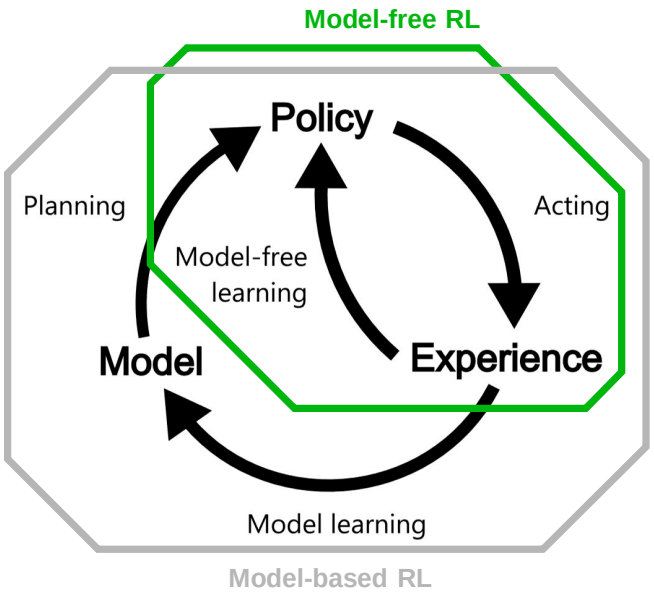
  $V^{\pi}_{\gamma}(\mathbf{s}) = \mathbb{E}_{\pi}\left[\sum_{i=t}^{\infty} \gamma^{i-t} r_{i+1} | S_t = \mathbf{s}\right]$ Value of state: expected cumulative rewards
- Optimal policy: $\pi^* = \arg\max_{\pi} V^{\pi}_{\gamma}(\mathbf{s}) \; \forall \; \mathbf{s} \in \mathcal{S}$





Aalto University
School of Electrical Engineering

Model-free and model-based RL



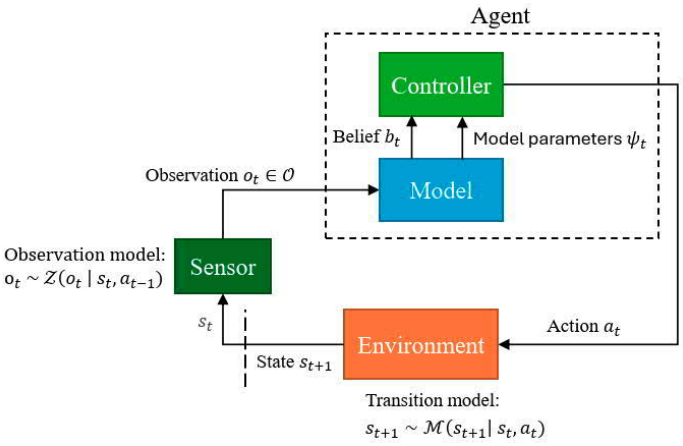
Reasons to use model-based RL

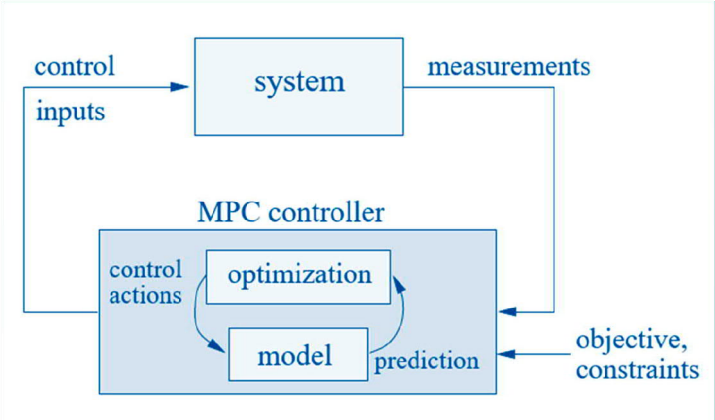
- ❑ Exploit rich structural information about communications and radar systems
- ❑ Predict rewards and states without tedious real-world trial and error interactions
- ❑ **Use learning only where it gives added value!**

Method	Model Used	Model Learned	Policy Learned
Model-free RL	✗	✗	✓
Model-based RL	✓	✗	✓
	✓	✓	✓
	✓	✓	✗
Fully model-based	✓	✗	✗

Could be considered as Learning
Model-Based Control system: learned dynamic model and model predictive optimal controller

Model-Based RL (POMDP) and Learning Model-Predictive Control





Learning dynamic model of state transitions and rewards

Constrained POMDP Objective

Performance evaluation via (discounted) cumulative rewards:

$$V_\pi(b) = \mathop{\mathbb{E}}_{\pi}\left[\sum_{i=0}^{\infty} \gamma^i r_t | B_t = b\right]$$
$$= \mathbb{E}\left[r_t + \gamma r_{t+1} + \gamma^2 r_{t+2} + \gamma^3 r_{t+3} \ldots | B_t = b\right]$$

$\gamma \in [0, 1)$; far future rewards emphasized less than near future rewards

E.g., radar task objective

$$\arg\max_{\pi} \; V_\pi^{(0)}(b)$$
$$\text{s.t.} \quad V_\pi^{(i)}(b) \geqslant C_i \quad \forall\, i = 1, \ldots, N_C$$

E.g., communications constraints

Radar rewards: classical *statistical* and *information-theoretic* measures or ones that support system or task level objectives: Mutual Information, low CRB, CFAR detector, track quality

Communications rewards: Sum rate**,** Data throughput, QoS, capacity, fairness, BER, SER. latency

RL for joint spatial and frequency resource allocation

- Closed-loop MBRL-based frequency and spatial resource allocation method that optimizes the radar task performance while meeting the communications rate or QoS constraints
- Allocation of beamspace-domain orthogonal beams and frequency-domain resource blocks to sensing (GLRT target detection and tracking) and comms tasks
- Bayesian Thompson sampling method to trade off exploration and exploitation
  - allocation of the power resources to beamspace beams and frequency domain resource blocks by solving a regret minimization problem

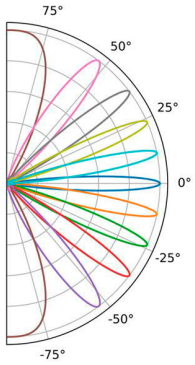
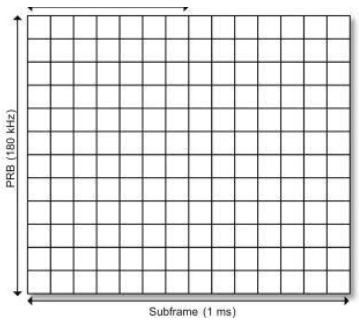


Fig. Example of DFT beam codebook with 10 beams.



Transmit signal model

- Transmit signal on *n*:th subcarrier and *m*:th OFDM symbol

Diagonal matrix

$$\mathbf{x}_{n,m} = \sum_{k=0}^{K} \mathbf{U}\boldsymbol{\Lambda}_n^{(k)}\mathbf{c}_{n,m}^{(k)}$$

Spatially multiplexed data

Beamspace transform (e.g., DFT beams)

- Beamspace transform (BT) maps antenna element domain processing to orthogonal beam domain processing
- Comprises one signal component dedicated for sensing and *K* signal components for communications
  - Index $k = 0$ correspond to the sensing waveform
  - Index $k = 1, \ldots, K$ correspond to the signal dedicated to user *k*
- Matrices $\Lambda_n^{(k)}$ for all n and k is used to realize different power allocations to different beams and sub-carriers/resource blocks
  - For $k > 0$ we consider constraint that allows only single spatial data stream per user per sub-channel
- Allocated power to one resource block element: $p_{n,l}^{(k)} = [\boldsymbol{\Lambda}_n^{(k)}(\boldsymbol{\Lambda}_n^{(k)})^{\mathrm{H}}]_{l,l}$

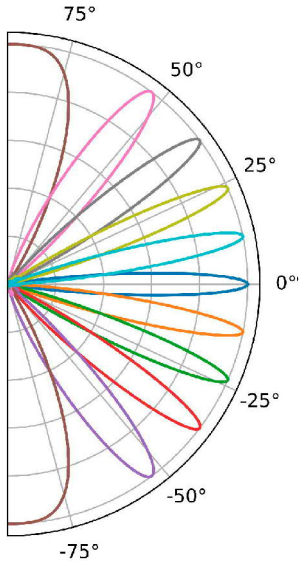


Fig. Example of DFT beam codebook with 10 beams.

Sensing signal model

Target scattering coefficient

OFDM symbol duration (including CP)

Two-way propagation delay

RX/TX steering vectors

White complex gaussian noise with variance σ^2

$$\mathbf{y}_{n,m} = \sum_{q=1}^{Q} \alpha_q e^{j2\pi(m\rho_q T_\mathrm{o} - n\Delta f \tau_q)} \mathbf{u}_\mathrm{R}(\theta_q)\mathbf{u}_\mathrm{T}^\mathrm{H}(\theta_q)\mathbf{x}_{n,m} + \mathbf{v}_{n,m}$$

Doppler shift

Target azimuth angle

- Generalized likelihood ratio test (GLRT) detector used to detect targets in 3D range-Doppler-azimuth data cube
  - Test statistic: $\eta(\tau,\rho,\theta) = \frac{1}{\sigma^2 C(\theta;\mathbf{X})}|S(\tau,\rho,\theta;\mathbf{X})|^2$

Matched filter

Beamshape

$$\eta(\tau,\rho,\theta)\,|\,\mathcal{H}_0 \sim \mathcal{E}(\eta;1)$$
$$\eta(\tau,\rho,\theta)\,|\,\mathcal{H}_1 \sim \mathcal{E}\left(\eta; C(\theta;\mathbf{X})\frac{\sigma_\alpha^2}{\sigma^2}+1\right)$$

where $\mathcal{E}(\eta;\mu) = \mu^{-1}e^{-\frac{\eta}{\mu}}$

  - α,τ,ρ are nuisance parameters
  - Constant false alarm rate (CFAR) property
  - Probability of detection (PoD) of Swerling I target: $P_\mathrm{d}(\theta,\sigma_\alpha^2;\mathbf{X}) = P_\mathrm{fa}^{\left(C(\theta;\mathbf{X})\frac{\sigma_\alpha^2}{\sigma^2}+1\right)^{-1}}$

Communications signal model

Downlink communications channel to user $k = 1,\ldots,K$ on sub-carrier $n = 1,\ldots,N$

$$r_{n,m}^{(k)} = \underbrace{(\mathbf{h}_n^{(k)})^\mathrm{H}\mathbf{x}_{n,m}^{(k)}}_{\text{signal of interest}} + \overbrace{\sum_{k'\neq k}(\mathbf{h}_n^{(k)})^\mathrm{H}\mathbf{x}_{n,m}^{(k')}}^{\text{interference}} + \underbrace{v_{n,m}^{(k)}}_{\text{noise}}$$

- The communications rate for user k is bounded from above by:

Beamspace channel gain: $g_{n,l}^{(k)} := [|\mathbf{U}\mathbf{h}_n^{(k)}|^2]_l$

$$\mathcal{R}_k(\mathbf{G};\mathbf{P}) = \frac{1}{T_\mathrm{o}}\sum_{n=1}^{N}\log_2\left(1+\frac{\sum_{l=1}^{L} g_{n,l}^{(k)} p_{n,l}^{(k)}}{\sum_{k'\neq k}\sum_{l=1}^{L} g_{n,l}^{(k)} p_{n,l}^{(k')} + \sigma_k^2}\right)$$

Considered utilities and constraints

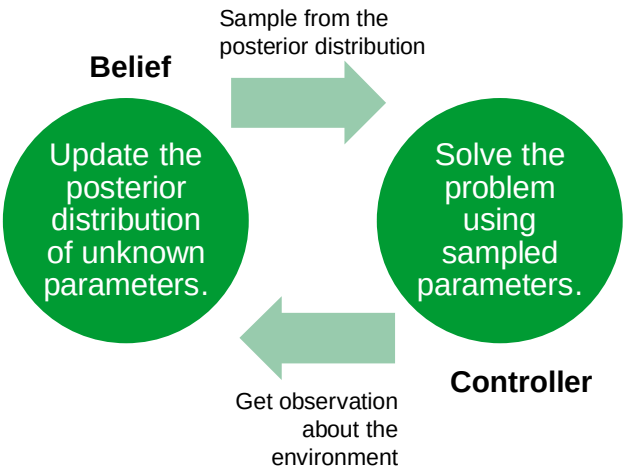
- Tracking and target search utilities

$$\tilde{u}_0(\mathbf{b}_t, \mathbf{a}_t) \coloneqq \mathcal{E}_{\mathrm{T}}(\mathbf{b}_t, \mathbf{a}_t) + \tilde{u}_{\mathrm{I}}(\mathbf{b}_t)\mathcal{E}_{\mathrm{S}}(\mathbf{b}_t, \mathbf{a}_t),$$

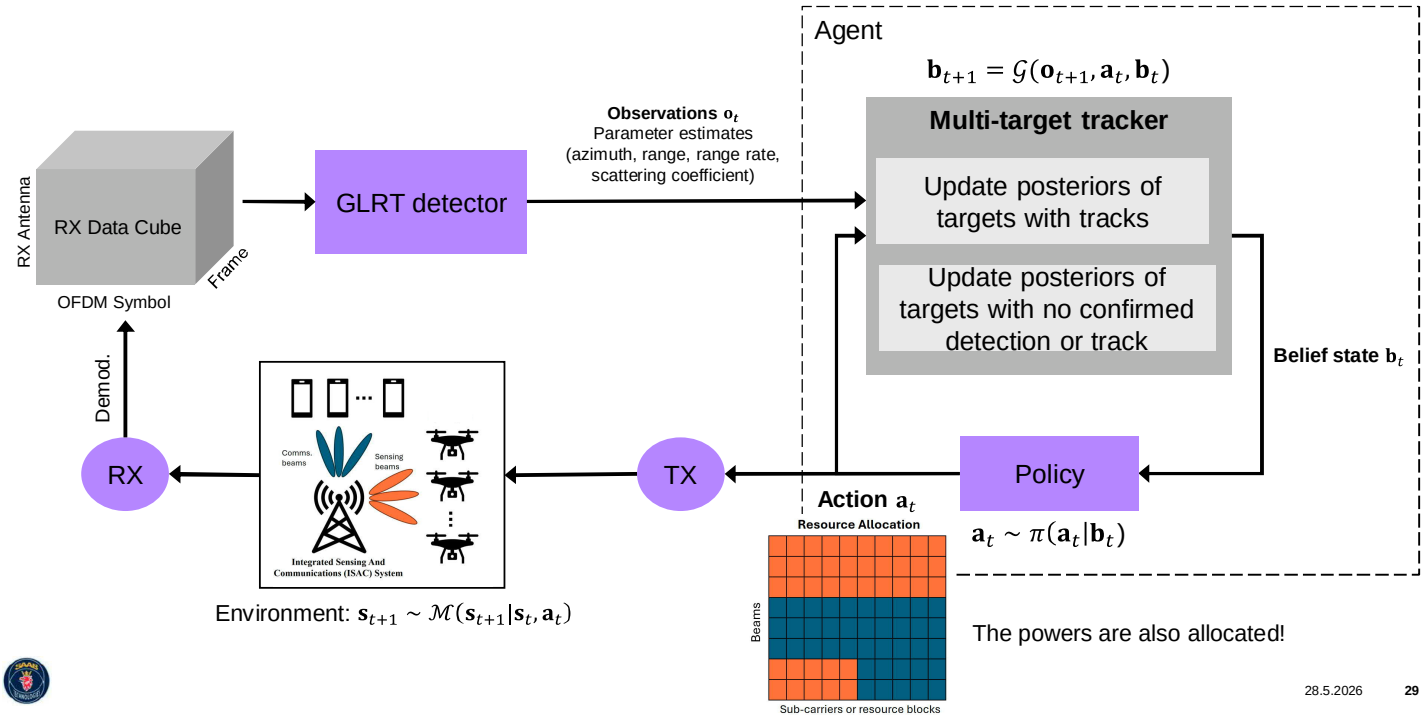
- Sum of P_{d} under strict P_{fa} constraints
- Track sharpness, max eigenvalue of error ellipse in Multitarget Tracking
- **Constraints**: the total power constraint, non-negativity of the allocated powers, SPR constraint (RMS BW for range resolution), desired data rate for each user (MI, Shannon-Hartley).
- Minimal CRB or Bayesian bound constraint

Proposed MBRL algorithm

- Model-based: Bayes' rule to update the posterior distribution of the unknown parameters for sensing and comms (dynamics of transitions and rewards)
- sampling these parameters from the posterior distribution and solving the problem as if they were known.
- Proposed method based on Thompson Sampling (TS) for choosing actions
  - Bayesian online learning method that balances exploration and exploitation tradeoff
- Application to our problem:
  - Unknown sensing parameters include target kinematics, gains, and uncertainty about the number of targets
  - Controller maximizes the sensing utilities using sampled parameters while meeting the communication rate constraints



Joint frequency–beamspace resource allocation



Joint frequency–beamspace resource allocation

Pseudo-myopic control strategy (learning-based policy)

Optimism under uncertainty principle: Overestimated long-term tracking utilities

Weight between search-and-track operations (heuristic)

Thompson sampling algorithm: Sample targets for new tracks from the posterior

$$\arg\max_{\mathbf{a}_t \in \mathcal{A}} \mathcal{E}_{\mathrm{T}}(\mathbf{b}_t, \mathbf{a}_t) + \tilde{u}_{\mathrm{I}}(\mathbf{b}_t)\mathcal{E}_{\mathrm{S}}(\mathbf{b}_t, \mathbf{a}_t)$$

$$\text{s.t. } u_k(\mathbf{b}_t, \mathbf{a}_t) \geq 0 \quad \forall\, k \in [K],$$

Greedy strategy for the communications constraints (no exploration needed due to decoupled channel estimation scheme)

Simulation numerology

- Methods evaluated over a sequence of time steps to investigate dynamic behavior of the algorithm

Table 1: Simulation parameters.

Description	Symbol	Value
# of subcarriers	N	1200
# of symbols	M	128
Carrier frequency	f_{c}	6.0 Ghz
Sub-carrier spacing	Δf	15.0 kHz
Cyclic prefix	T_{cp}	16.7 μs
Total power	P_{tot}	30.0 dBm
Noise power	$\sigma^2, \sigma_k^2 \; \forall \; k \in [K]$	-132.2 dBm
# of TX antennas	L_{T}	32
# of RX antennas	L_{R}	32
RB size		12
Rate requirement	$C_k \; \forall \; k \in [K]$	20 Mbps
# of targets	Q	8
Mean RCS	σ_{rcs}	1.0
# of users	K	4
# of Monte Carlo iterations		80
Prob. of False Alarm	P_{fa}	1.0e-06

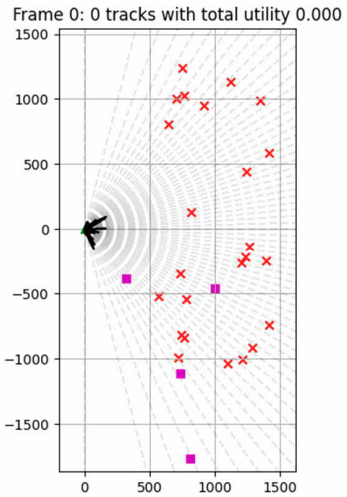
Table 2: Sensing properties of the considered ISAC system.

Unambiguous range	9.99 km
Max range (CP limited)	2.50 km
Range resolution	8.33 m
Unambiguous velocity	539.41 km/h
Max velocity (Doppler limited)	269.81 km/h
Velocity resolution	4.21 km/h
Integration time	10.67 ms

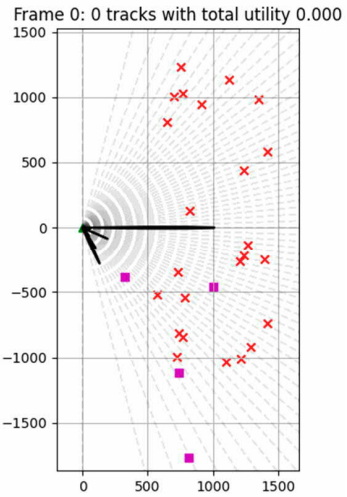
Aalto University
School of Electrical Engineering

Joint allocation of beam and frequency resources

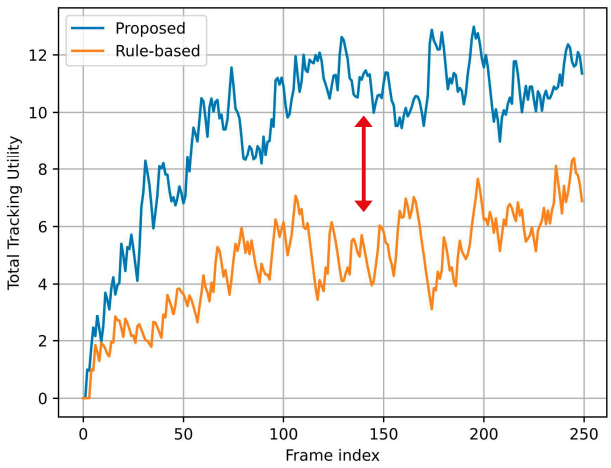
Proposed model-based RL method



Rule-based method



The proposed method can find targets faster and track them with better track qualities!



- Average utilities under different SNRs and rate constraints

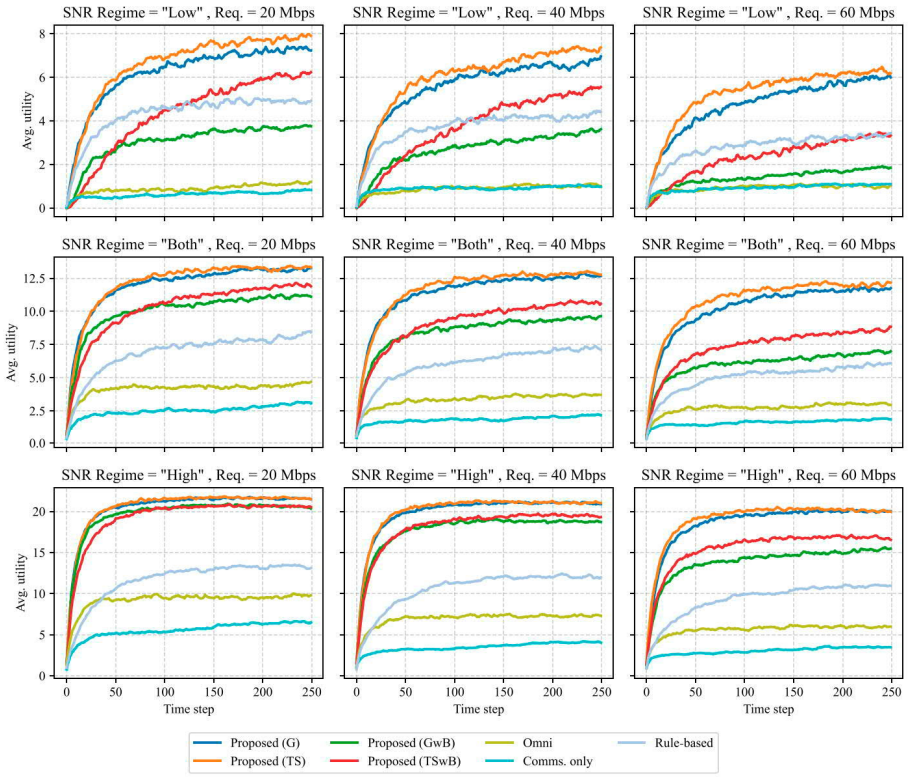


Fig. 4. The average utility as a function of time in different SNR regimes and with different communications rate constraints. The proposed TS-l algorithm achieves the highest average utility in all scenarios, but the difference is not large compared to the *greedy* variant. Furthermore, it can be seer the proposed adaptive search bonus improves the sensing performance significantly.

Conclusions and take-home messages

- Parallel technology convergence, shared HW, antenna, circuit and spectrum resources
- Standards need to be defined for ISAC systems to ensure investments
- Business models for sensing and sensing related services have to be found
- ISAC systems with intelligent waveform design and resource management for different tasks
- Waveform design and optimization can be formulated as structure optimization or machine learning problems. ML allows for dealing with modelling and algorithmic deficits
- Model-based RL / Learning Model-Based Control enable:
  - Real-time adaption and learning in a congested and dynamic radio spectrum
  - Exploiting rich knowledge on propagation, interference, communication and sensing systems behavior
  - predicting the dynamics and rewards performance while avoiding tedious trial and error steps of MFRL,
  - Learning from experiences, continuous autonomous improvement of performance
  - Explainability, stability, using analytical tools from control and estimation

Some of our references

- V. Koivunen, M. F. Keskin, H. Wymeersch, M. Valkama and N. González-Prelcic, "Multicarrier ISAC: Advances in waveform design, signal processing, and learning under nonidealities," in IEEE Signal Processing Magazine, vol. 41, no. 5, pp. 17-30, Sept. 2024
- MM. Bica, K. -W. Huang, V. Koivunen and U. Mitra, "Mutual information based radar waveform design for joint radar and cellular communication systems," 2016 IEEE ICASSP, Shanghai, China, 2016, pp. 3671-3675
- Bică and V. Koivunen, "Radar Waveform Optimization for Target Parameter Estimation in Cooperative Radar-Communications Systems," in IEEE Transactions on Aerospace and Electronic Systems, vol. 55, no. 5, pp. 2314-2326, Oct. 2019
- M. F. Keskin, V. Koivunen and H. Wymeersch, "Limited Feedforward Waveform Design for OFDM Dual-Functional Radar-Communications," in IEEE Transactions on Signal Processing, vol. 69, pp. 2955-2970, 2021
- T. E. Abrudan, A. Haghparast and V. Koivunen, "Time Synchronization and Ranging in OFDM Systems Using Time-Reversal," in IEEE Transactions on Instrumentation and Measurement, vol. 62, no. 12, pp. 3276-3290, Dec. 2013
- K. V. Mishra, M. R. Bhavani Shankar, V. Koivunen, B. Ottersten and S. A. Vorobyov, "Toward Millimeter-Wave Joint Radar Communications: A Signal Processing Perspective," in IEEE Signal Processing Magazine, vol. 36, no. 5, pp. 100-114, Sept. 2019
- V. Koivunen, M. F. Keskin, H. Wymeersch, M. Valkama and N. González-Prelcic, "Multicarrier ISAC: Advances in waveform design, signal processing, and learning under nonidealities," in IEEE Signal Processing Magazine, vol. 41, no. 5, pp. 17-30, Sept. 2024
- P. Pulkkinen and V. Koivunen. Model-Based Online Learning for Resource Sharing in Joint Radar-Communication Systems. In Proceedings of the IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP), Singapore, pp. 4103-4107, May 2022.
- P. Pulkkinen and V. Koivunen. Model-Based Online Learning for Joint Radar-Communication Systems Operating in Dynamic Interference. In Proceedings of the 30th European Signal Processing Conference (EUSIPCO), Belgrade, Serbia, pp.992-996, Aug 2022.
- P. Pulkkinen and V. Koivunen. Model-Free Online Learning for Waveform Optimization In Integrated Sensing And Communications. In Proceedings of the IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP), Rhodes Island, Greece, pp.1-5, Jun 2023.
- P. Pulkkinen and V. Koivunen. Model-Based Online Learning for Active ISAC Waveform Optimization. Journal of Selected Topics in Signal Processing, vol. 18, no. 5, pp. 737-751, July 2024.
- P. Pulkkinen and V. Koivunen. Partially Observable Model-Based Learning For ISAC Resource Allocation. In Proceedings of the IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP), Seoul, Korea, pp.12996-13000, Apr 2024.
- P. Pulkkinen, M. Esfandiari and V. Koivunen. Cognitive Beamspace Algorithm for Integrated Sensing and Communications. In Proceedings of the IEEE Radar Conference (RadarConf), Denver, C0, USA, pp.1-6, May 2024.
- P. Pulkkinen, M. Esfandiari and V. Koivunen. Beamspace and Frequency Domain ISAC Resource Allocation using Reinforcement Learning. In Proceedings of the 58th Annual Asilomar Conference on Signals, Systems, and Computers, Pacific Grove, CA, USA, pp. 443-449, Oct 2024.
- P. Pulkkinen, M. Esfandiari, H.V Poor and V. Koivunen. Multicarrier MIMO ISAC Resource Allocation Using Model-Based Reinforcement Learning. Submitted to IEEE Transactions on Signal Processing, May 2025.

Thank you for your interest!

Questions?

Aalto University
School of Electrical
Engineering