BEAMFORMING WITH FREE ENERGY PRINCIPLE UNDER HIERARCHICAL CODEBOOK

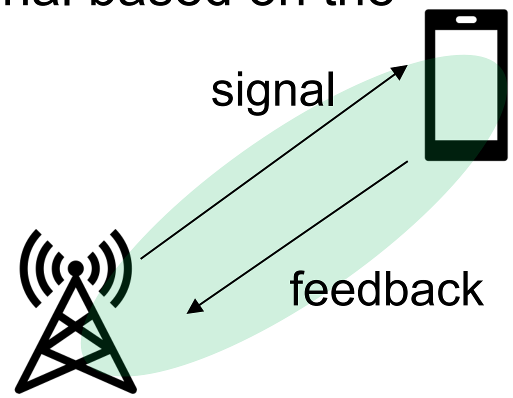
Tatsuya Otoshi †, Masayuki Murata ‡

† Graduate School of Economics, Osaka University

‡ Graduate School of Information Science and Technology, Osaka University

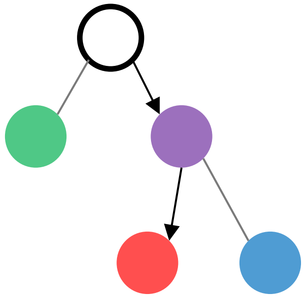
Beamforming

- Effective use of a large number of antennas at the base station
  - Generally, base stations have many antennas and terminals have few antennas
  - Since the terminal side has only a few antennas, it is difficult to receive the throughput gain of MIMO.
  - Signal propagation can be made more directional by using a large number of antennas at the base station
- Signal control based on channel propagation conditions is necessary
  - Base station estimates the propagation state based on feedback from the terminal
  - Controls the signal to be transmitted to amplify the signal received by the terminal based on the propagation state
- Handling of channel state fluctuations is an issue



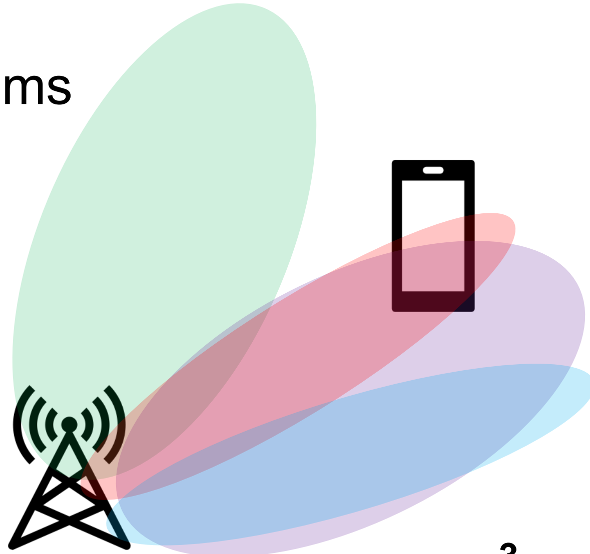
Hierarchical Codebook

- Propagation state estimation becomes more difficult as the number of antennas increases
  - Number of feedbacks required increases with the number of antennas
  - Increased overhead due to signal transmission for measurement → lower throughput



- Hierarchical beam search by Hierarchical codebook
  - Candidate beams are prepared in advance as a codebook
  - Hierarchical codebook is constructed from spatially rough beams to finer beams
  - Reduction of feedback signal by limiting the search range

- Redo search when propagation conditions change



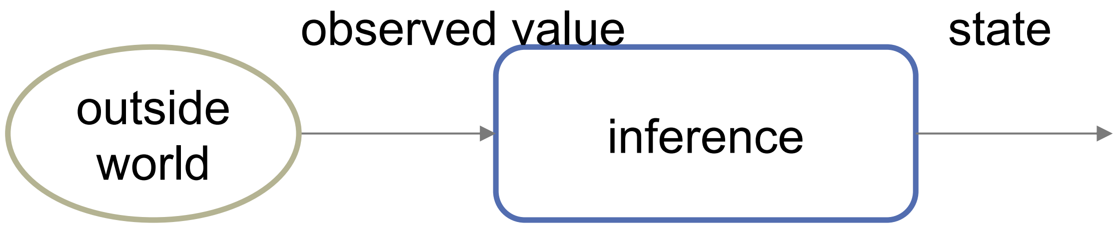
Approach

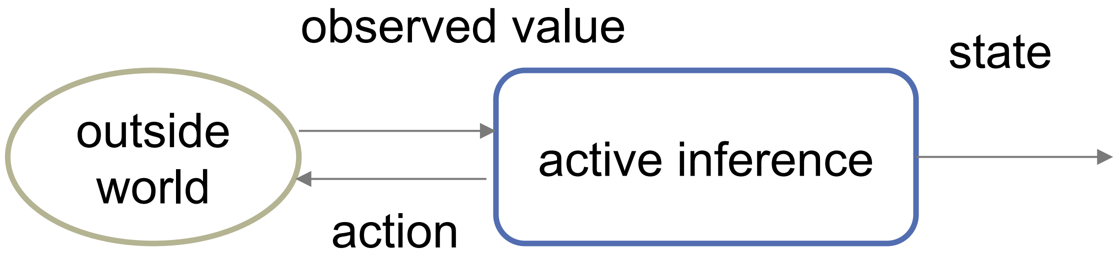
- Beamforming has the overhead of estimating the propagation state.
- Hierarchical codebooks enable somewhat efficient search, but require re-search as conditions change

- Lightweight continuous search, rather than re-exploration, is the best way to deal with changes
- The way people perceive the world is active inference, which always combines search and execution
- Application of active inference to keep up with changing conditions with small overhead

Active inference

- Ordinary inference estimates a "good" state given observed values.
- Active inference estimates a "good" state, including changing observed values through actions.
  - Example: Peek under the table to see what is hidden under the table.
  - Example: Switching between various beams to estimate channel conditions
- Using free energy as a measure of goodness





Free energy principle

- A theory that comprehensively describes the functioning of the brain
  - Describe inference and actions as minimization of "free energy"
- Free energy
  - $F = D_{KL}[Q(s)|P(s|x)] - \log P(x)$
    - First term: Posterior distribution of state s $\mathrm{P}(s|x)$ and approximate distribution $\mathrm{Q}(s)$ Kullback-Leibler information amount of
    - Second term : Shannon surprise
- Inference
  - Estimate the posterior distribution
- Action
  - Select the action that will yield the observed value x that reduces the Shannon surprise in addition to the accuracy of the inference

Inference

- observed value x
  - Feedback of signal strength from the device
- condition s
  - Propagation channel information
- State estimation
  - $\frac{\partial F_\tau}{\partial q^f} = 0$

$$\Longrightarrow \; Q^*(s_\tau^f) = \sigma\left(\mathbb{E}_{q^{i\setminus f}}\left[\ln P(\mathbf{o}_\tau|\mathbf{s}_\tau)\right] + \ln\left(\mathbb{E}_{P(s_{\tau-1}^f, u_{\tau-1}^f)}[P(s_\tau^f|s_{\tau-1}^f, u_{\tau-1}^f)]\right)\right)$$

Action

- action u
  - beam vector w
  - Transmission power p
- policy $P(u_t|\pi)$
  - Determine behavior by estimating the policy using the behavior distribution as a policy (control as inference)
  - The actual action shall be the one with the highest probability.
- Policy estimation

$$Q^*(\pi) = \underset{Q(\pi)}{\operatorname{argmin}} \ \mathcal{F}$$

$$\Longrightarrow \quad Q^*(\pi) = \sigma(-\mathbf{G}(\pi) + \ln P(\pi_0) - F(\pi))$$

Direct gain
(preference for observed values)

$$\mathbf{G}(\pi) = \sum_{\tau} \mathbf{G}_\tau(\pi) \qquad \mathbf{G}_\tau(\pi) \geq -\underbrace{\mathbb{E}_{Q(o_\tau|\pi)}\left[\mathrm{D}_{KL}[Q(s_\tau|o_\tau,\pi) \parallel Q(s_\tau|\pi)]\right]}_{\text{Epistemic Value}} - \underbrace{\mathbb{E}_{Q(o_\tau|\pi)}[\ln \tilde{P}(o_\tau)]}_{\text{Utility}}$$

Estimated value of obtaining the observed value o

Preference prior

- A probability distribution that expresses the goodness of the observed value itself in determining behavior.
  - Corresponds to reward function and objective function
  - Predicted distribution of observations
    - Minimize surprise = Obtain observed values with high probability = Obtain observed values with strong preferences
- Objective function
  - Transmission rate
- Preference distribution reflecting objective function
  - Boltzmann distribution with negative transmission rate as energy

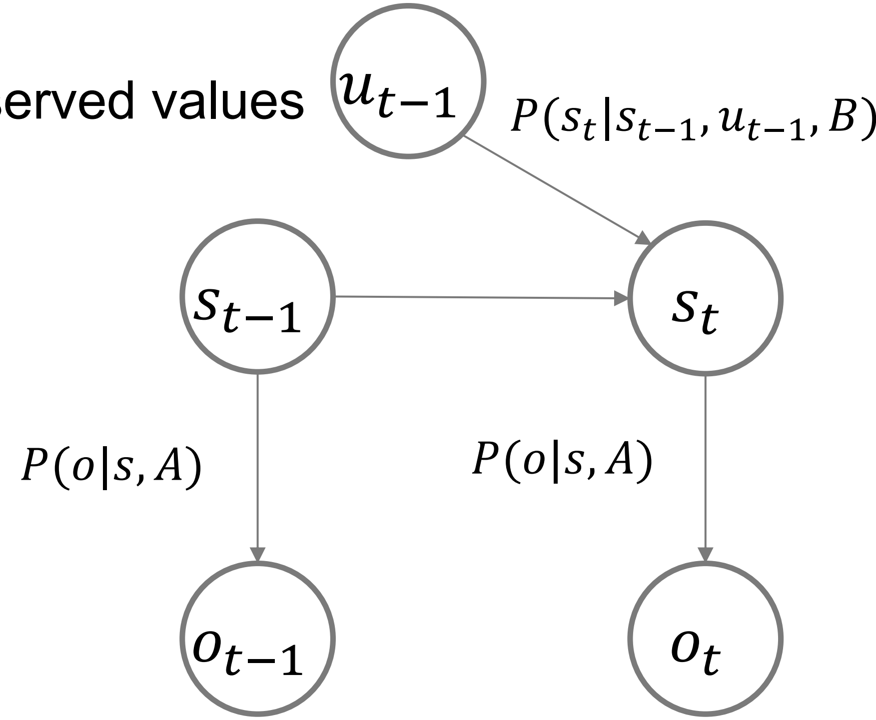
$$\tilde{P}(o_t) \propto \exp(-\beta\epsilon) = \exp\big(\beta R(o_t)\big)$$

Learning

- Observation model: A
  - A probabilistic model that expresses the relationship between
  - Used to estimate the state from observed values and predict observed values

$$\frac{\partial F}{\partial \ln A} = 0 \implies \mathbf{a}^* = a + \sum_{\tau=1}^{T} o_\tau \otimes \mathbf{s}_\tau$$

- State transition model : B
  - A probabilistic model that expresses the time change of state
  - Used to predict the state when deciding on actions

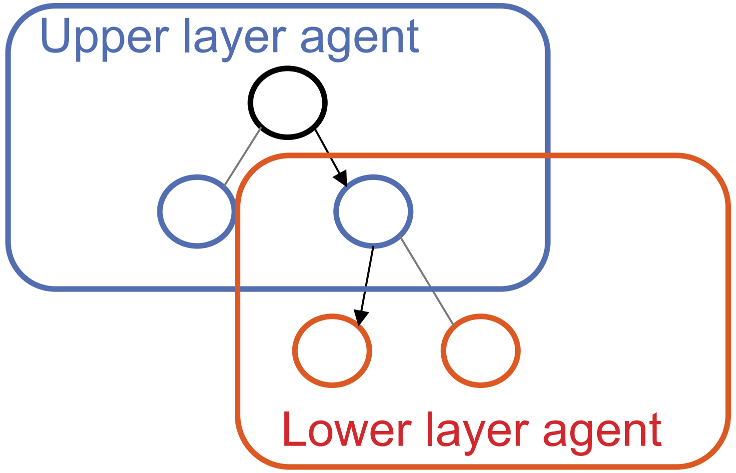


$$\frac{\partial F}{\partial \ln B} = 0 \implies \mathbf{b}_u* = b_u + \sum_{\tau=2}^{T} \sum_{\pi} Q(u_\tau|\pi)Q(\pi)\left(Q(s_\tau|\pi) \otimes Q(s_{\tau-1}|\pi)\right)$$

Searching on hierarchical codebook with FEP

- FEP agents at each level of the Hierarchical codebook
- Active inference is used to make beam decisions at each level
- Information is shared between agents at the upper and lower levels to achieve coordinated operation.
  - Subtle changes are handled by the lower layers.
  - Switching to the upper layer when the lower layer cannot handle the change

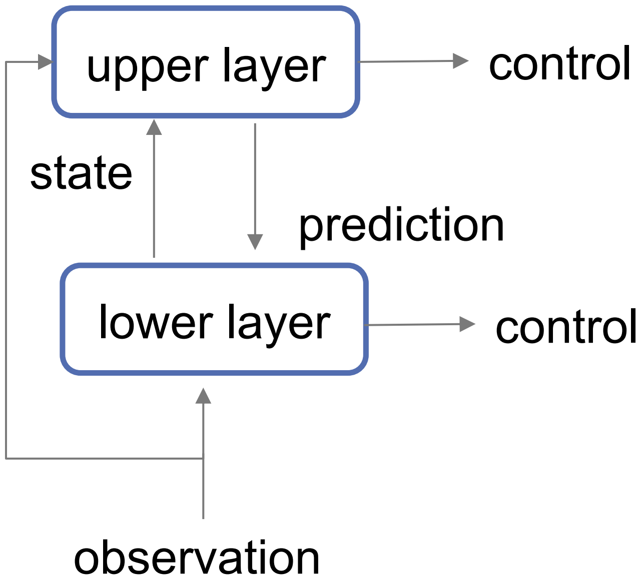
Hierarchical codebook



Cooperation between hierarchical FEP agents

- Upper layer agent
  - Observes the state resulting from the inference of lower layer agents
    - By reducing the dimensionality of the state, it is possible to reduce the load concentrated on the upper layer.
  - Predicts the desired situation when lower layer cooperation is realized and feed it back to the lower layer.

- Lower layer agent
  - Make control decisions by inferring the state from actual observed values
  - Achieving cooperation by using the predictions of upper-layer agents as the prior distribution for inference
    - Achieving cooperation through inference and control to minimize prediction errors



Simulation environment

- base station
  - 4 antennas
- channel coefficient
  - multipath fading
    - 4 pathes
    - With time variation

$$h_{ij}(t) = \sqrt{\frac{\beta_{i,j}}{L}} \sum_{l=1}^{L} g_{i,j}(t,l) a_{i,j}^{\dagger}(\theta)$$

$$g_{i,j}(t,l) = \rho g_{i,j}(t-1,l) + \sqrt{1-\rho^2} e_{i,j}(t)$$

$$a_{i,j}(\theta_l) = \frac{1}{\sqrt{N}}(1, \exp(\pi i l \cdot 1 \cos\theta_l), \cdots, \exp(\pi i l \cdot (N-1)\cos\theta_l)))$$

- UE
  - 1 antenna
  - Concentric circles moving at a constant speed
- Hierarchical codebook
  - 2 layers
  - 2^l beams/layer

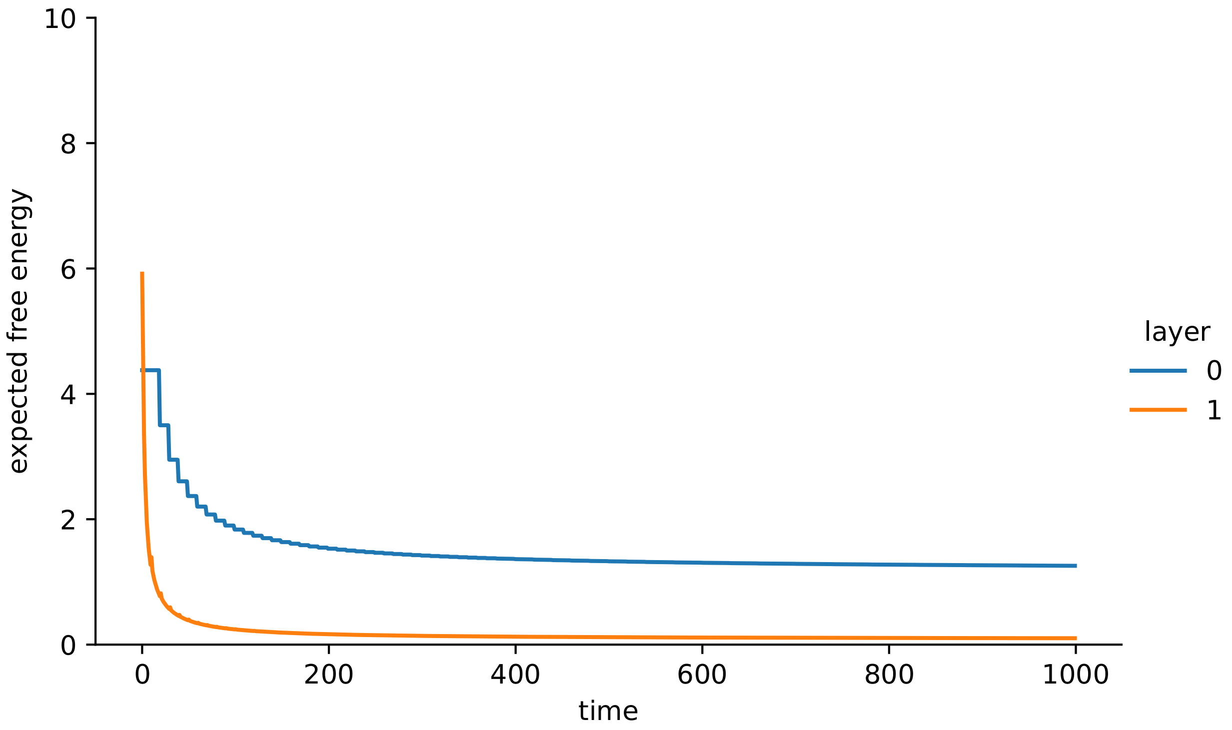
$$\boldsymbol{w}_n^{(l)} = (\boldsymbol{a}(2^l, -1 + \tfrac{2n-1}{2^l}), \boldsymbol{0}_{N-2^l})$$

$$\boldsymbol{a}(N, \Omega) = \tfrac{1}{\sqrt{N}}(\exp(i\pi 0 \Omega), \cdots, \exp(i\pi(N-1)\Omega))$$

Convergence of Free Energy

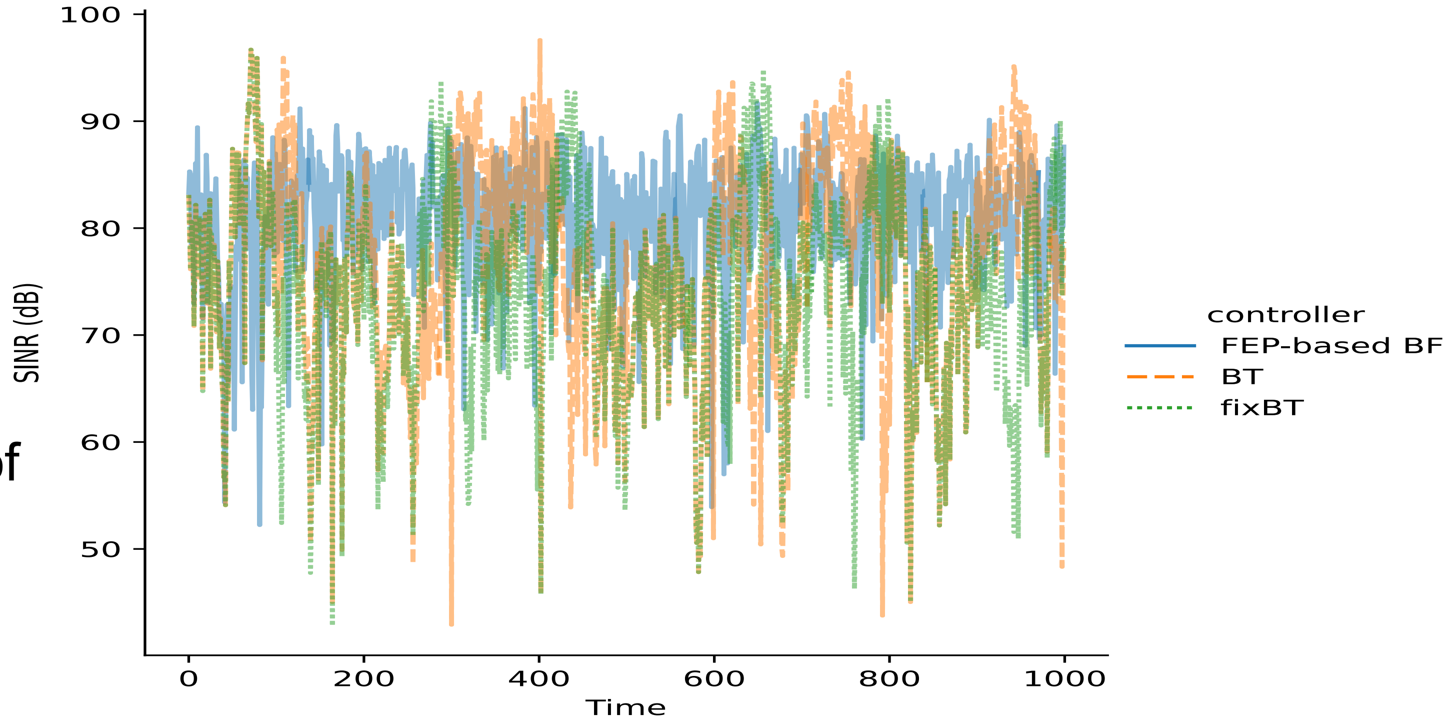
- The upper and lower agents operate independently and inferece and control to minimize free energy
  - Therefore, they do not necessarily converge.

- As shown in the figure on the right, free energy is converging.
  - Appropriate coordination is achieved through information exchange between upper and lower agents.



Communication performance

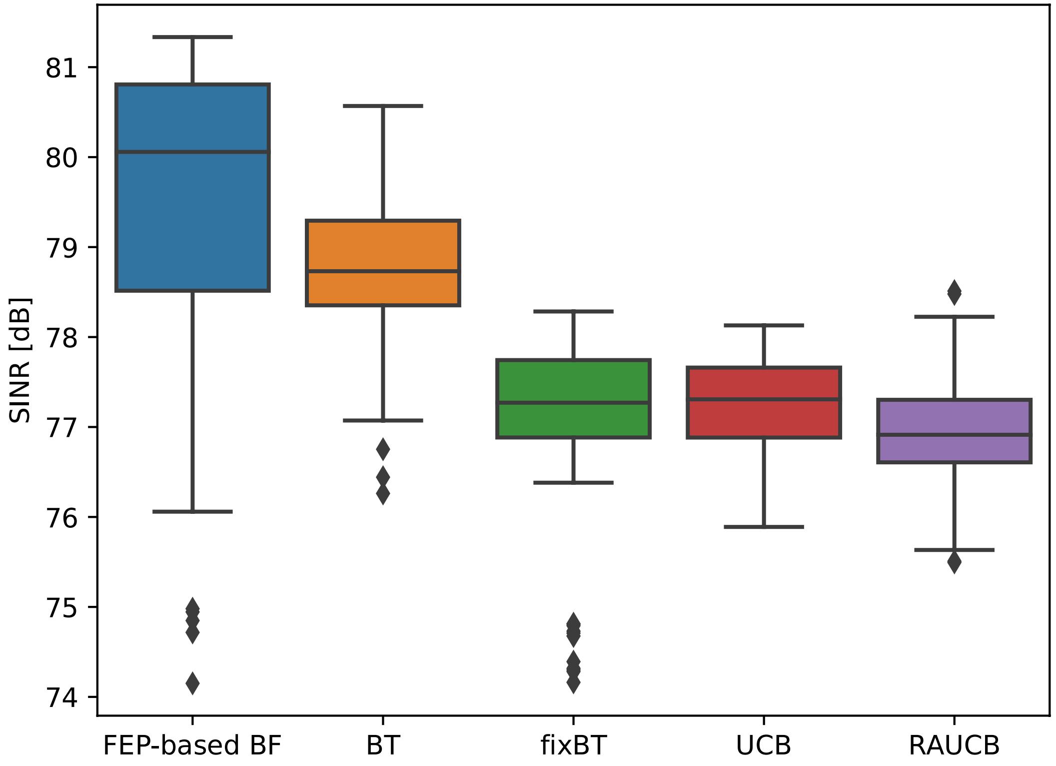
- Methods
  - FEP-based BF
  - BT: Beam training is performed periodically
  - fixBT: Beam training
- Communication performance
  - FEP-based BF maintains stable and high SNR
    - Adaptive control over time variations in the environment
  - BT and fixBT have temporary periods of low SNR
    - This is due to the lack of beam selection suitable for environmental fluctuations.



Comparison

- FEP-based BF has the highest average and quartile SNR
  - In some cases, the minimum value is smaller than BT, but this is due to exploratory behavior
- BT has higher SNR than fixBT
  - Compared to fixBT, which fixes the beam, BT periodically selects the beam and is better able to follow environmental changes.
- UCB and RAUCB have similar or lower SNR than BT
  - Both UCB and RAUCB are basically dealing with uncertainty in a static environment.

- FEP-based BF
- BT: Beam training is performed periodically
- fixBT: Beam training first
- UCB: multi-armed bandit with UCB
- RAUCB: randomly reject UCB



Summary and future work

- Summary
  - Proposed a solution based on the free energy principle framework for beamforming with hierarchical codebook
  - FEP agents are deployed at each level of hierarchy, and state predictions are shared among agents for coordination between upper and lower levels
  - A higher SNR can be maintained on average than with periodic beam training

- Future work
  - Simulation considering the realistic movement of UE
  - Multimodal information processing such as UE location information

Thank you for your attention.