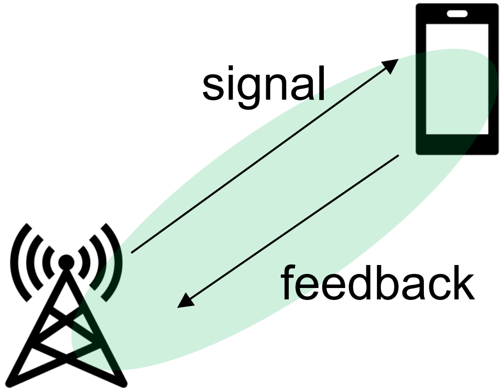
DISTRIBUTED BEAMFORMING BY MULTI-AGENT ACTIVE INFERENCE

Tatsuya Otoshi †, Masayuki Murata ‡

† Graduate School of Economics, Osaka University
‡ Graduate School of Information Science and Technology, Osaka University

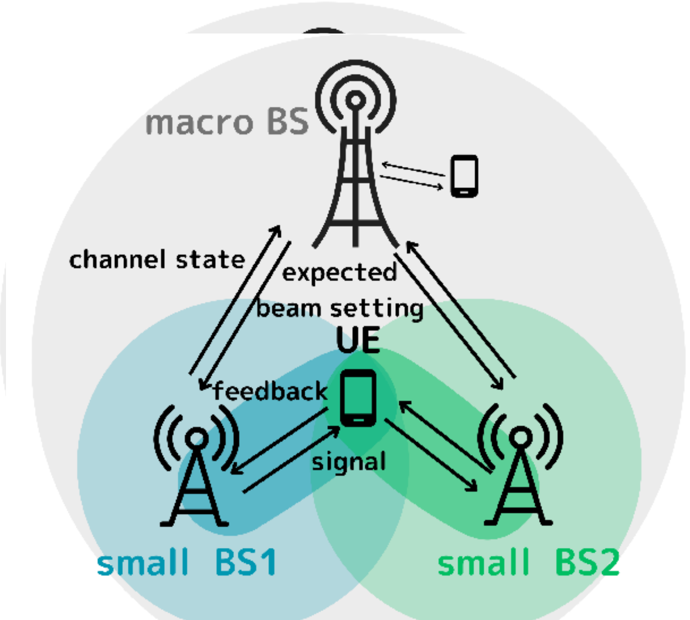
Beamforming

- 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



Beamforming with multiple base stations

- Efficient use of radio wave resources by cooperatively performing beamforming between base stations
- Appropriate collaboration methods vary depending on the accuracy of available information
  - Joint transmission:
    - Send the same signal from multiple base stations and amplify the signal
    - Need accurate channel information
  - Coordinated beamforming:
    - Beamforming is performed between base stations to avoid interference
    - Rough location is more important than precise channel information

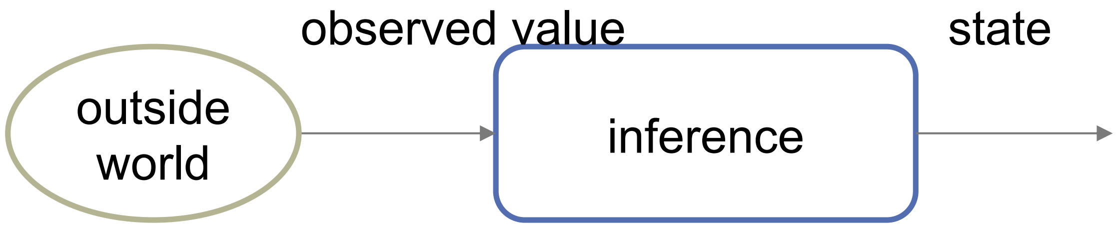


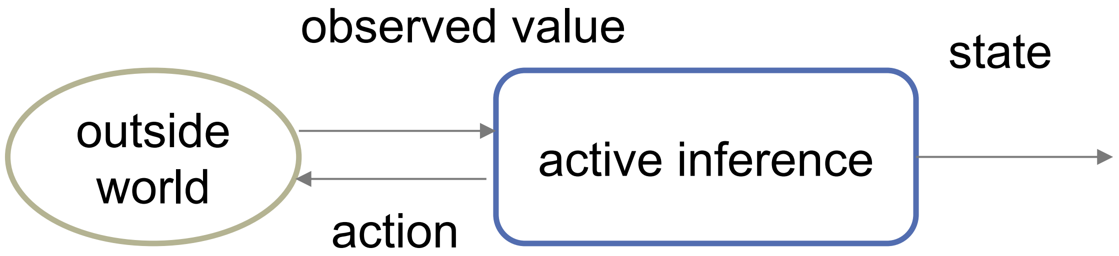
Trade-off between estimation accuracy and control performance

- In order to accurately grasp information, communication signal resources are sacrificed.
    - Allocate resources to measurement signals to improve estimation accuracy
    - A certain level of accuracy is necessary for optimizing communication signals, but anything more than that will result in a decline in communication performance.
- People constantly make trade-offs between accuracy and goal achievement under uncertainty.
    - Be proactive and obtain information to reduce uncertainty
    - Make decisions once a certain amount of information has been gathered, rather than aiming for zero uncertainty.
- Solving the trade-off between estimation accuracy and control performance by applying human active inference

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 reasoning 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 for
- 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
  - $\frac{\partial F_\tau}{\partial q^f} = 0$

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

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

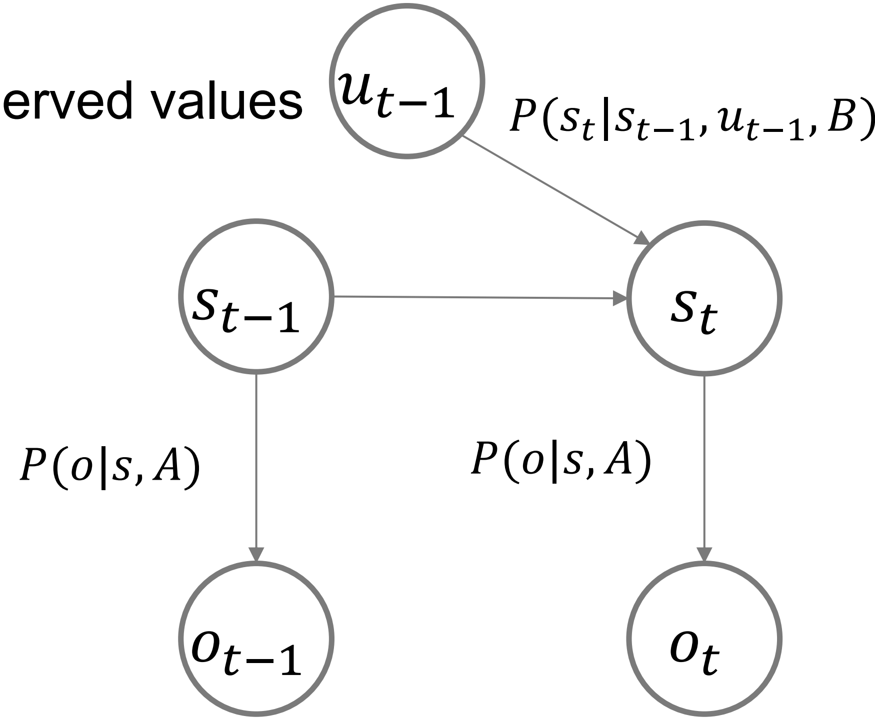
- A probability distribution that expresses the goodness of the observed value itself in determining behavior.
  - Corresponds to reward function and objective function
  - Predicted distribution of observations
    - Minimize surprise = Obtain observed values with high probability = Obtain observed 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

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

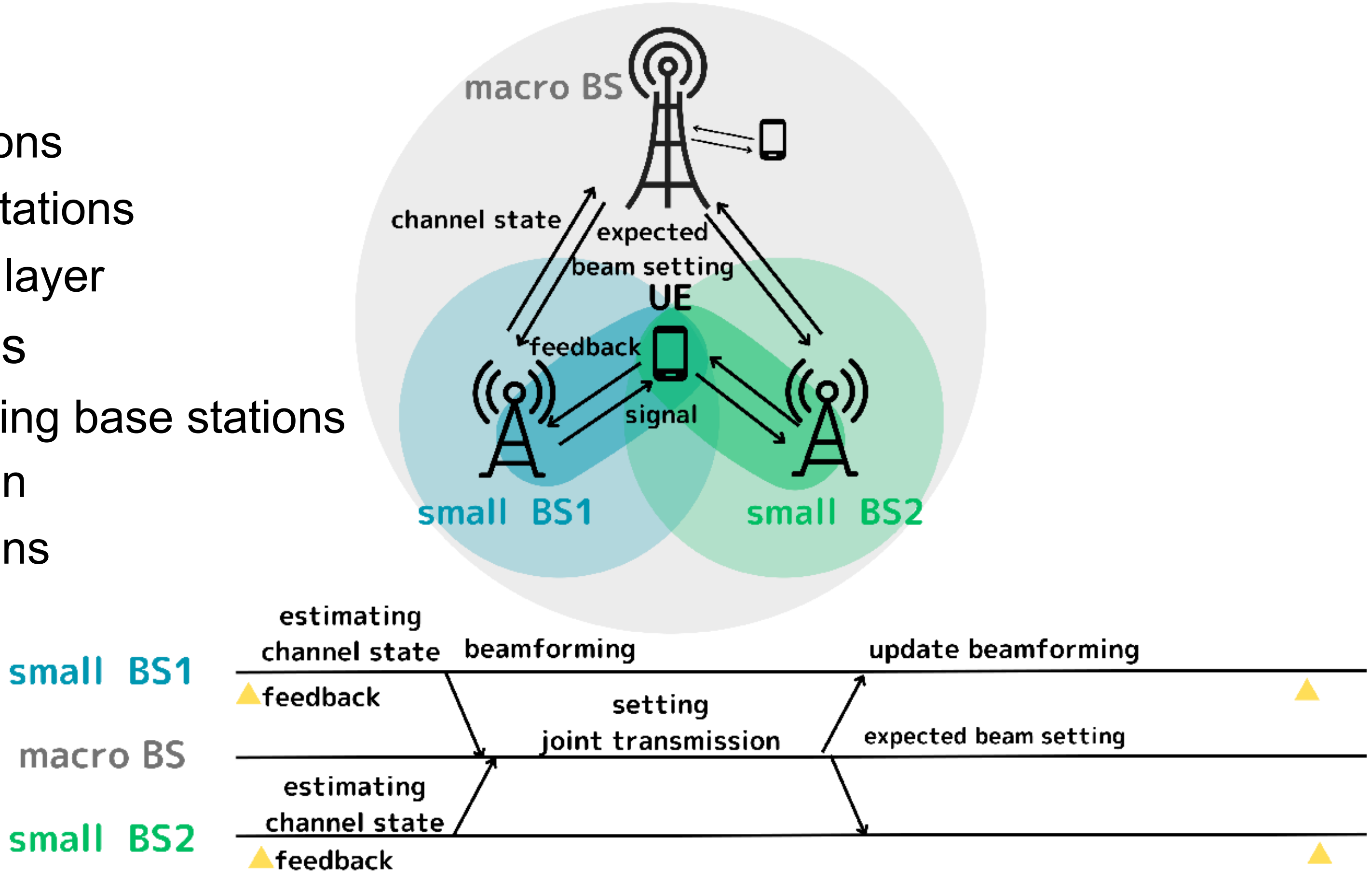


- State transition model : B
  - A probabilistic model that expresses the time change of state
  - Used to predict the state when deciding on actions

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

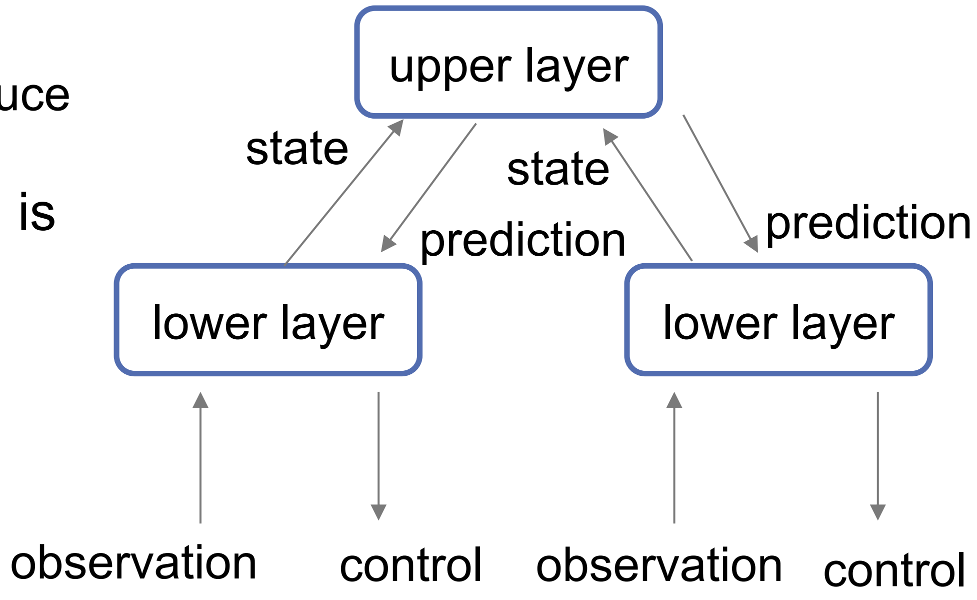
Coordination between base stations

- Perform cooperative operations by exchanging information between base stations
- How to exchange information
  - Sharing via upper base station
    - Integrate information at upper base stations
    - Information transmission to lower base stations
    - Load may be concentrated on the upper layer
  - Sharing between adjacent base stations
    - Exchange information between neighboring base stations
    - Each base station makes its own decision
    - Control may conflict between base stations



Cooperation between hierarchical FEP agents

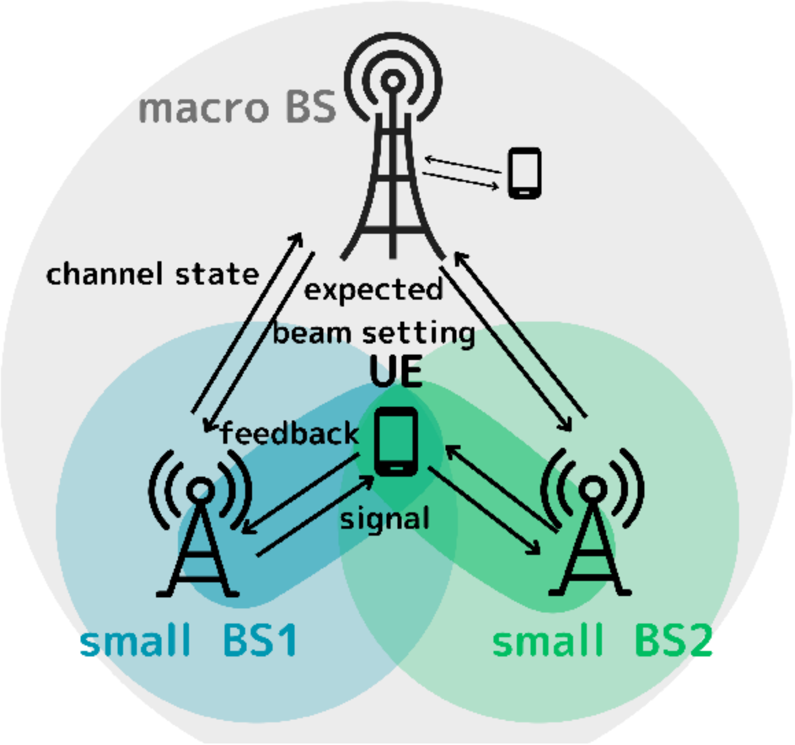
- Upper layer agent
  - Observes the state resulting from the inference of lower layer agents
    - By reducing the dimensionality of the state, it is possible to reduce the load concentrated on the upper layer.
  - Predicts the desired situation when lower layer cooperation is realized and feed it back 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
  - 2 BSs
  - 4 antennas/BS
- channel coefficient
  - multipath fading
    - 4 pathes
    - Each path is a complex Gaussian

- UE
  - 1 or 3 UEs
  - 1 antenna/UE
- beam
  - Direction: 4 types
  - Power: 5 levels



$$h_{ij} = \sqrt{\frac{\beta_{i,j}}{L}} \sum_{l}^{L} a^{\dagger}_{i,j}(\theta)$$

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

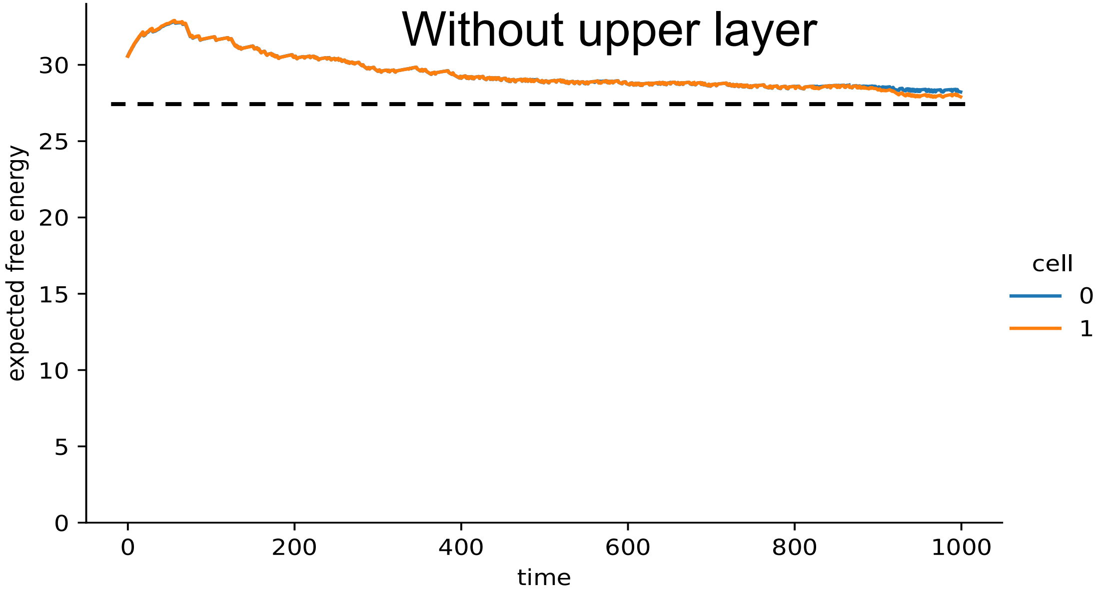
$$\exp(\pi i l \cdot (N-1)\cos\theta_l)))$$

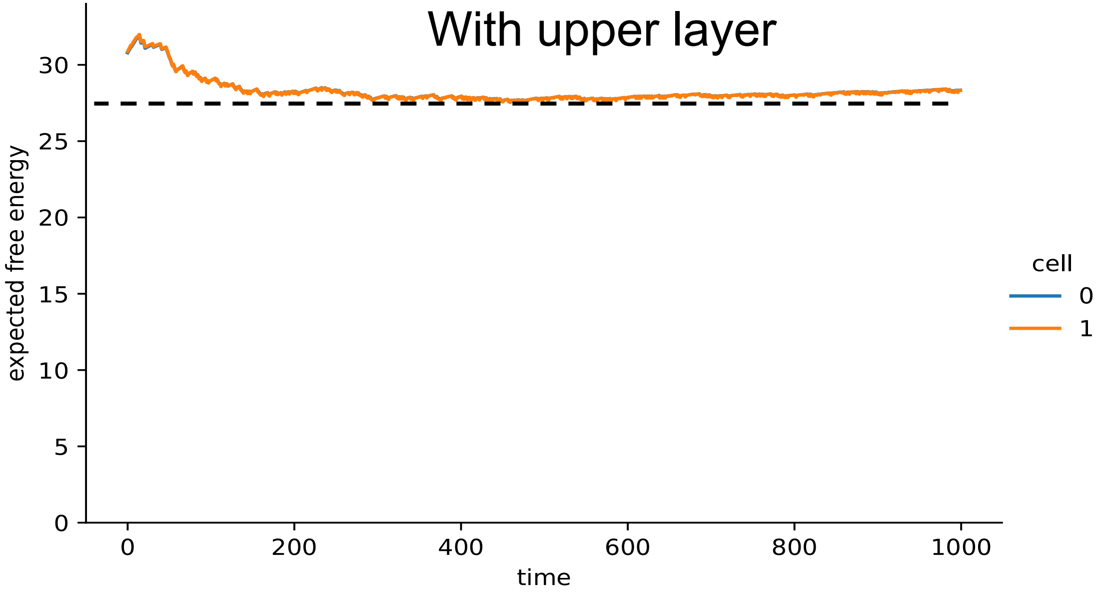
$$\vec{w}_n = (\frac{1}{\sqrt{N}} \exp(\frac{2\pi i(n-1)0}{N}), \cdots,$$

$$\frac{1}{\sqrt{N}} \exp(\frac{2\pi i(n-1)(N-1)}{N}))$$

Convergence of Free Energy

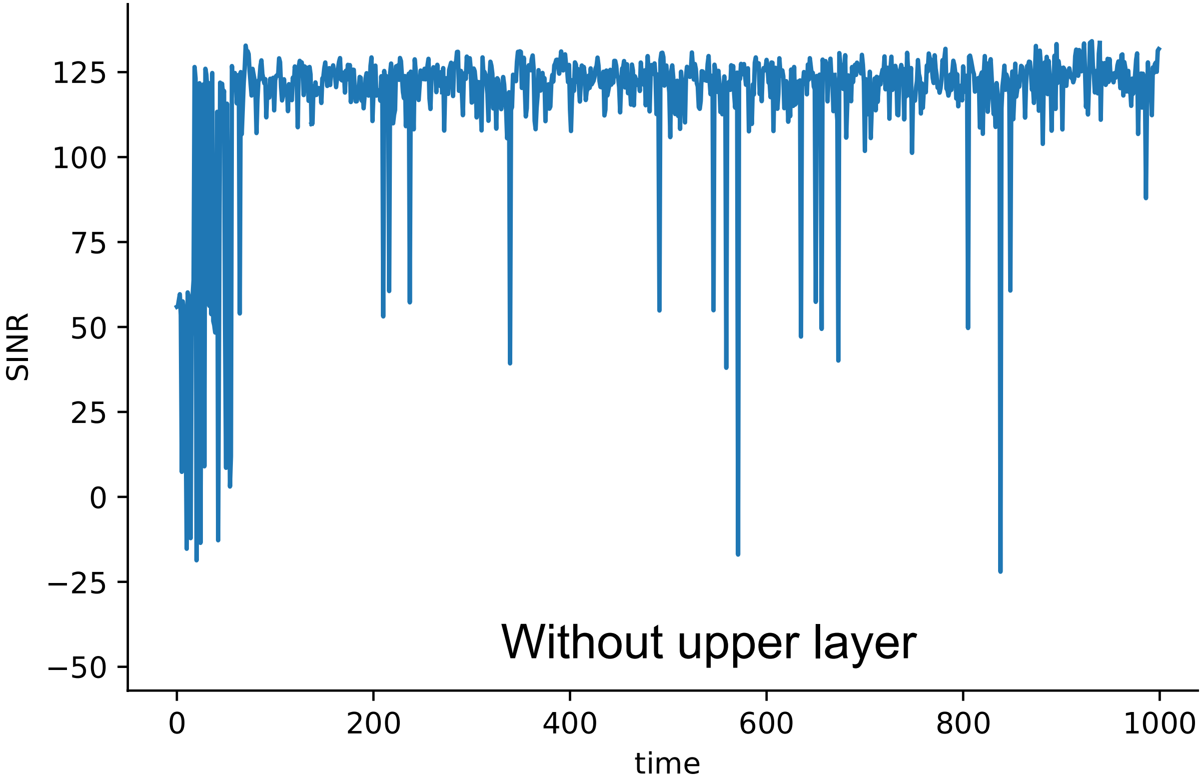
- Information exchange through upper layers speeds up control convergence.
  - When upper layer agents are deployed, convergence of expected free energy is achieved in about 200 steps.
  - Without upper layer agents, it takes about 1000 steps to converge.
- Convergence is possible with decentralized control, but convergence is faster with hierarchical control.

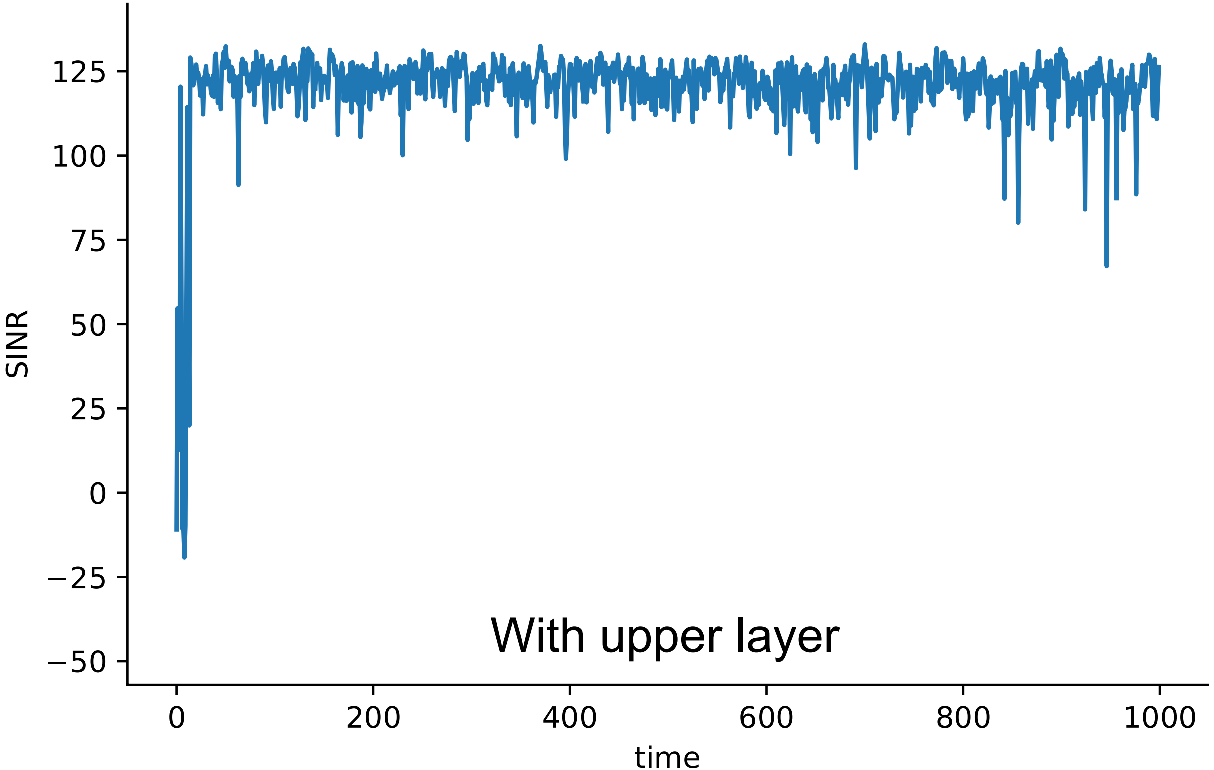




Communication performance

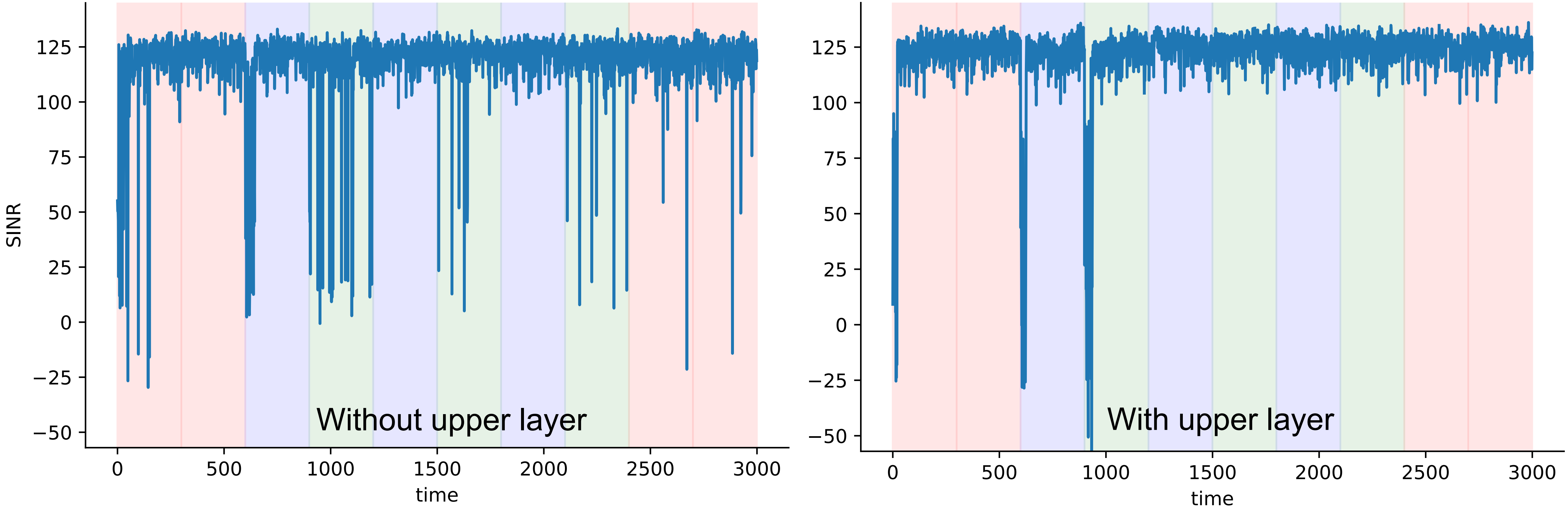
- Information exchange through upper layers quickly selects the appropriate beam
- Without upper layer agents in place, large SINR drops occur many times before convergence





Switching Multiple UEs

- Beamforming switched between three UEs.
  - Red, blue, and green in the graph show beamforming for different UEs.
- High SINR is maintained after beamforming for each UE.



Summary and future work

- summary
  - Proposed a solution based on the free energy principle framework for beamforming with coordination among multiple base stations
  - Achieved coordination as the aggregation of information by the upper layer and prediction for the lower layer, and the realization of prediction by the lower layer.
  - As a result, appropriate beam selection can be achieved in a short time with little feedback information exchange.

- Future work
  - Simulation considering the movement of UE
  - Multimodal information processing such as UE location information
  - Realization of shortcuts for collaboration without upper layers