Distributed Beamforming by Multi-Agent Active Inference

Tatsuya Otoshi
Graduate School of Economics
Osaka University
Osaka, Japan
t-otoshi@econ.osaka-u.ac.jp

Masayuki Murata
Graduate School of Information Science and Technology
Osaka University
Osaka, Japan
murata@ist.osaka-u.ac.jp

***Abstract*—Massive MIMO systems have emerged as a promising technology for next-generation wireless communication networks. Beamforming, a key technique in Massive MIMO, significantly enhances the system's performance by exploiting the spatial domain. This paper presents a collaborative beamforming approach that leverages vertical collaboration between base stations to improve adaptation and mitigate fading variations, while also applying the principle of free energy to optimize beamforming. The approach involves sharing estimated channel states and learned models among base stations, enabling them to collectively optimize beamforming based on the principle of free energy. Evaluations in single-user operation scenarios demonstrate that both with and without vertical collaboration, the expected free energy, representing beamforming performance, decreases over time. However, vertical collaboration reduces temporary decreases in signal-to-interference-plus-noise ratio (SINR) caused by independent adaptation to fading variations. Furthermore, in a multi-user switching scenario, the proposed approach ensures stable control by utilizing learned models and state estimation results, leading to improved beamforming performance during switching.**

***Index Terms*—Beamforming, Active Inference, Free Energy Principle, 5G NR, CoMP**

I. Introduction

In recent years, beamforming has gained significant importance in the context of the fifth-generation (5G) communication technology [1]. Beamforming is a technique that enhances communication quality and capacity by controlling the directionality of the transmitted radio waves using a transmit antenna array.

In particular, the combination of beamforming with Massive MIMO (Massive Multiple Input Multiple Output) has further improved the directional capabilities [2]. Massive MIMO utilizes a large number of transmit and receive antennas, enabling simultaneous communication with multiple User Equipment (UE) devices.

However, in the border regions of the coverage area, providing sufficient communication quality solely through a single base station can be challenging. This necessitates the need for Coordinated Multi-Point (CoMP) operations, where multiple base stations cooperate to perform beamforming and improve communication quality for UE devices in the border regions [3], [4].

In CoMP, there is a trade-off between the accuracy of feedback information for beamforming and system performance. Utilizing accurate feedback information enables optimal beamforming, but there can be delays or errors associated with obtaining such information. When accurate information is available, high throughput can be achieved through Joint Transmission (JT), where multiple base stations transmit the same signal. On the other hand, when accurate information is not available, Coordinated Scheduling/Beamforming (CS/CB) is employed, where multiple base stations transmit different signals while avoiding interference. In this case, the throughput is lower compared to JT.

In addition, there is a trade-off between channel state estimation and data transmission in beamforming. To obtain accurate information, it is necessary to acquire a substantial amount of feedback information. Typically, dedicated pilot signals are transmitted for feedback purposes. The transmission of pilot signals for channel state estimation limits the resources available for data transmission. Even if the ACK (Acknowledgment) for data transmission is used as feedback, the feedback information obtained varies depending on the beam configuration. Therefore, it is necessary to switch the beams for measurement purposes. This implies that there are occasions where the opportunity to sacrifice beams optimized for data transmission arises in order to obtain feedback.

To address these challenges, we use the Free Energy Principle (FEP) [5], drawing inspiration from the workings of the human brain. FEP aims to estimate and utilize the information necessary for the system to make optimal decisions. In FEP, the goal is to minimize the free energy, which is a combination of the exploration and exploitation value. By minimizing free energy, FEP allows for the selection of actions that strike a balance between exploration and exploitation. This means that FEP enables the system to choose actions that are optimal in terms of balancing the trade-off between exploration and exploitation. In the context of beamforming or similar scenarios where a trade-off exists between observation and control, FEP can be leveraged to make optimal decisions. By using FEP, the cooperation among multiple base stations in beamforming can be achieved, leading to an improvement in communication quality.

In this paper, we propose a cooperative beamforming technique using the Free Energy Principle (FEP). To achieve this, we deploy FEP agents at each base station and place a higher-

979-8-3503-7099-7/24/$31.00 ©2024 IEEE

level FEP agent to coordinate their actions. The lower-level agents perform beamforming using information from the higher-level agent and feedback information from the User Equipment (UE). The higher-level agent observes the state of the lower-level agents and provides feedback in the form of predicted states. By employing this hierarchical structure, coordinated beamforming is realized, ensuring coherence across the system.

The evaluation compares the performance of hierarchical coordination with non-coordination scenarios, elucidating the effectiveness of hierarchical cooperation.

This paper is organized as follows. Section II discusses related work. Section III describes the system model. Section IV presents the proposed method. Section V evaluates the proposed method. Finally, Section VI concludes the paper.

II. Related Work

In previous studies targeting hierarchical beamforming with macrocells and small cells, most of them assumed sparse placement of small cells without overlapping with each other [6]–[8]. In this case, user terminals connect to the small cell at their location, and if they do not belong to any small cell, they connect to the macrocell, ensuring a unique correspondence between cells and terminals. However, in practice, considerations must be given to terminal allocation between cells, such as the overlap between small cells or allocating terminals near the periphery of small cells to the macrocell.

In the work by [9], a method for performing independent beamforming in situations with overlapping cells and interference between them is investigated. However, this approach does not consider the hierarchical structure of cells.

Each cell determines beamforming and transmission power using local information within the cell, such as the transmission rate to the terminal, channel gains, and interference power including noise. Beamforming and transmission power selection are performed using Deep Q-network (DQN) from a discrete set of multiple candidates.

DQN learns the relationship between observed information and reward (power efficiency as the transmission rate per unit transmission power) using a neural network. The observed information uses only local information, while the reward incorporates global information by considering the overall power efficiency of all cells. The learning of the Q-network involves accumulating observed information and rewards, followed by offline iterative retraining.

The evaluation demonstrates that the overall power efficiency improves compared to randomly selected beamforming and transmission power choices or when using a greedy approach.

In the study by [7], beamforming for small cells within the coverage of macrocells is investigated. In this case, the interference between cells is only considered between the macrocell and small cells, without considering interference among other small cells. The macrocell has the ability to determine the transmitted radio waves with higher priority, while small cells perform beamforming in the presence of interference from the macrocell.

In the small cells, beamforming decisions are made based on channel coefficients, interference power from the macrocell, and noise power. Additionally, for each user terminal, the decision of whether to block it or not is made. Blocked terminals cannot transmit radio waves, reducing interference to other terminals. The beamforming and blocked user terminals are computed by solving an optimization problem to maximize the sum of transmission rates for all terminals.

In the context of active inference based on the principle of free energy, the policy $\pi = (a_1, \cdots, a_T)$ representing a sequence of selected actions is determined to minimize the following expected free energy:

$$G(o_{1:T}, s_{1:T}, \pi) = E_Q[\log Q(s_{1:T}, \pi) - \log \tilde{P}(o_{1:T}, s_{1:T}, \pi)] \quad (1)$$

Here, $o_{1:T}$ represents the observations, $s_{1:T}$ represents the states, Q represents the approximate posterior distribution, and $\tilde{P}$ represents the target distribution.

Given π, the expected free energy at each time step τ can be independently calculated as follows:

$$G_\tau(\pi) = E_{Q(o_\tau, s_\tau|\pi)}[\log Q(s_\tau|\pi) - \log \tilde{P}(o_\tau, s_\tau|\pi)] \quad (2)$$

$$\geq -E_{Q(o_\tau|\pi)}[D_{KL}(Q(s_\tau|o_\tau, \pi)||Q(s_\tau|\pi))] - E_{Q(o_\tau|\pi)}[\log \tilde{P}(o_\tau)] \quad (3)$$

The first term in the above equation provides the information gain of posterior distribution update by obtaining the new observation o_τ. The second term represents the expected utility based on the observation o_τ. The utility needs to be specified in advance by the prior distribution $\tilde{P}(o_\tau)$. Both the first and second terms are negative, and as the information gain and utility increase, the expected free energy decreases.

Hierarchical active inference models primarily focus on hierarchical organization within a single individual, where the upper layers provide goals to lower layers to determine specific actions [10]–[13].

In the context of small cell-macrocell systems, the macrocell not only modifies the goals of the small cells but also controls its own behavior. Additionally, the macrocell needs to interact with multiple small cells. While the coordination of multiple agents has been addressed in [13], it does not involve hierarchical coordination but rather coordination between equal agents.

In [12], hierarchical active inference is applied to robot motion control. The hierarchical structure in this case refers to the temporal hierarchy within a single robot, where the upper layers determine long-term trajectories while the lower layers decide short-term motor controls.

The upper layers aim to reach suitable positions based on preference distributions, while the lower layers aim to minimize deviations from the assumptions made by the upper layers using the upper layers' estimated distributions as prior beliefs.

In the upper layers, the path with the minimum expected free energy is determined by weighing the expected free energy at each point along the trajectory. On the other hand, in the lower

layers, individual movements are determined through sampling based on the magnitude of the expected free energy.

Using an actual robot and only RGB camera data, the ability to correctly move to the target position has been confirmed. However, this work does not consider comparisons with other methods or environmental changes.

In [13], a hierarchical active inference approach is applied to a cooperative cooking game called "overcooked," where agents (cooks) cooperate with each other. To facilitate agent cooperation, each agent estimates the goals of other agents and shares those goals.

In this approach, hierarchical organization occurs within each agent, with an upper layer that sets goals and a lower layer that determines specific actions. The estimated goals of others and top-down goals from the upper layer are reflected as the Kalman gain in the Kalman filter, and the lower layer makes decisions for integrated actions.

It has been demonstrated that this approach achieves comparable performance to Bayesian Delegation, which combines Bayesian estimation and Q-learning, in a shorter computation time.

III. System Model

A. Components

The system consists of M macrocells and S small cells that communicate with D User Equipment (UE). We consider a scenario where downlink communication involves interference among the beams transmitted from the base stations and interfering with each UE.

We define the set of base stations for macrocells as $B^{(m)}$, the set of base stations for small cells as $B^{(s)}$, and the set of UEs as U:

$$B^{(m)} = \{b_i^{(m)} | i = 1, \cdots, M\} \tag{4}$$

$$B^{(s)} = \{b_i^{(s)} | i = 1, \cdots, S\} \tag{5}$$

$$U = \{u_i | i = 1, \cdots, D\} \tag{6}$$

The set of all base stations is denoted as $B = B^{(m)} \cup B^{(s)}$.

B. Channel Coefficients and Beamforming

Each base station b transmits signals using M_b antennas, and the signals are transformed by the spatial characteristics and reach the M_u antennas of the receiving UE u. The spatial characteristics are represented by the matrix $H^{u,b}(t)$, where the element $h_{ij}^{u,b}(t)$ represents the channel coefficient when transmitting from antenna i to antenna j.

At each time, each UE u is connected to one of the base stations b, and this correspondence is represented by $A(t)$. The element $a_{u,b}(t)$ of $A(t)$ is 1 if u is connected to b, and 0 otherwise.

Each base station adjusts the phase and amplitude of the transmitted signals for each antenna and performs beamforming. The phase is represented by the beamforming vector $\boldsymbol{w}_b(t)$, and the amplitude is represented by $P_b(t)$. The relationship between the transmitted signal $x_b(t)$ and the received signal $y_{u,b}(t)$ is given by:

$$\boldsymbol{y}_{u,b}(t) = \sqrt{P_b(t)} H^{u,b}(t) \boldsymbol{w}_b(t) \circ \boldsymbol{x}_b(t) \tag{7}$$

where $\circ$ denotes element-wise multiplication.

1) Electromagnetic Interference: UE u receives signals that include noise and signals transmitted from other base stations, causing interference. The received signal at UE u is given by:

$$\boldsymbol{y}_u(t) = \boldsymbol{y}_{u,b}(t) + \sum_{b' \in B \setminus \{b\}} \boldsymbol{y}_{b'}(t) + \sigma_u(t) \tag{8}$$

The Signal-to-Interference-plus-Noise Ratio (SINR) is defined as follows:

$$\gamma_u(t) = \frac{P_b(t) \| H^{u,b}(t) \boldsymbol{w}^b(t) \|^2}{I_u(t) + \sigma_u^2(t)} \tag{9}$$

$$I_u(t) = \left\| \sum_{b' \in B \setminus \{b\}} \boldsymbol{y}_{b'}(t) \right\|^2 \tag{10}$$

where $I(t)$ represents the strength of interference signals, and $\sigma^2(t)$ represents the power of noise.

C. Objective Function

The data transmission rate to UE u is estimated using SINR as follows:

$$C_u(t) = \log(1 + \gamma_u(t)) \tag{11}$$

In the paper by [9], the reward is defined as the total transmission rate of all cells per unit power consumed by all cells, aiming to maximize the overall power efficiency. In this case, the power efficiency at time t is given by:

$$\mathrm{EE}(t) = \sum_u \mathrm{EE}_u(t) \tag{12}$$

$$\mathrm{EE}_u(t) = \frac{C_u(t)}{\sum_b P_b(t)} \tag{13}$$

However, when using this objective function, fairness among UEs is not considered, which may lead to unfairness among UEs, such as prioritizing nearby UEs that can achieve higher transmission rates while making communication difficult for distant UEs. In the paper by [14], for the case of beamforming in a single base station, they propose using objective functions that consider fairness, such as the logarithmic sum of power efficiency for each UE or the maximization of the minimum power efficiency. Similar objective functions can be applied in the case of multiple base stations to achieve beamforming considering fairness.

IV. Distributed Beamforming by FEP

A. Active Inference by Multiple Agents

To minimize the expected free energy, it is necessary to calculate the expected free energy for all possible actions based on the observations and states. In the case of beamforming with multiple base stations, the number of observations, states, and actions increases with the number of base stations. In

particular, for states and actions, considering their combinations is necessary for distribution calculations, making it impractical to minimize the expected free energy of the entire system. Therefore, it is desirable for the active inference agents deployed at each base station to decide local actions based on local information.

In this regard, each base station $b \in B$ determines a local action sequence $\pi^{(b)} = (a_1^{(b)}, \cdots, a_T^{(b)})$ to minimize the following local free energy under the local observation information $o_\tau^{(b)}$ and the local state $s_\tau^{(b)}$:

$$G_\tau(\pi)^{(b)} = E_{Q(o_\tau^{(b)}, s_\tau^{(b)}|\pi^{(b)})}[\log Q(s_\tau^{(b)}|\pi^{(b)}) - \log \tilde{P}(o_\tau^{(b)}, s_\tau^{(b)}|\pi^{(b)})] \quad (14)$$

In this case, collaboration with other base stations is achieved through the exchange of local observation information $o_\tau^{(b)}$ or prediction and prediction errors. In FEP, collaboration at the same hierarchical level is achieved by other agents observing the results of an agent's actions through observations [13]. On the other hand, collaboration between upper and lower levels is achieved through the exchange of predictions and prediction errors, where the upper level predicts the state of the lower level from a higher level of abstraction, and the lower level conveys the prediction error to the upper level [12].

In the case of beamforming, collaboration at the same hierarchical level involves collaboration where the settings of each beamforming are observed as equal partners, while collaboration between upper and lower levels involves the upper level predicting the desired states of each small cell as a result of mediating the interference of each lower level, and the lower level providing feedback on prediction errors to the upper level. In this case, collaboration between the upper and lower levels requires more overhead for exchanging information since it is necessary to transmit information about the distribution rather than the determined value of beamforming results in the case of collaboration at the same hierarchical level. Therefore, collaboration between the upper and lower levels is expected to be performed on a longer timescale than collaboration within the same hierarchical level.

B. Active Inference in Vertical Collaboration

We provide a detailed description of the operation of active inference agents in the vertical collaboration between base stations. Specifically, we define observations, states, actions, and preference distributions in the framework of FEP.

1) Lower-level Agents: Lower-level agents are deployed in each small cell and determine the shape of the beam based on the feedback from the user equipment (UE). Although lower-level agents are independent of each other, they achieve cooperation among agents by receiving the predicted distribution from the upper level and implementing control along the preference distribution.

a) Observations: We assume that base station b can observe the transmission rate to each UE based on the feedback from the UE, denoted as $\boldsymbol{C}_b(t)$. Therefore, it corresponds to $\boldsymbol{o}_b(t) = \boldsymbol{C}_b(t)$.

As feedback, the receiving side SINR is measured through CSI or SS, and the transmission rate is estimated from the SINR.

b) States: Based on the observed transmission rate $\boldsymbol{C}_b(t)$ and the information of its own beamforming, base station b estimates the channel coefficients $H^{u,b}(t)$. Therefore, the system state is represented as $s_b(t) = (H^{u,b}(t), \boldsymbol{C}_b(t))$.

The prior distribution of states is given as the predicted distribution of lower-level states by the upper level, and the lower level achieves cooperation by controlling to align with the assumptions of the upper level. Therefore, the prior distribution of states in the lower level is given as follows:

$$\tilde{P}_C(s_b(t)) = P(s_b(t)|o_{b^h}(t)) \quad (15)$$

where b^h represents the base station of the macrocell.

c) Actions: The base station specifies the shape of the beam by the beam vector $\boldsymbol{w}_b(t)$ and determines the transmit power $P_b(t)$. Therefore, $a_b(t) = (\boldsymbol{w}_b(t), P_b(t))$ corresponds to the actions.

d) Preference Distribution: The preference distribution sets the desired states for control as the prior distribution with respect to the observations. When a specific objective function $R_b(o_\tau)$ is set, such as power efficiency or maximization of transmission rate, the preference distribution can reflect the objective function $R(o_\tau)$ by using the Boltzmann distribution as the preference, where utility is measured in terms of the logarithm of the preference distribution. It can be represented as follows:

$$\tilde{P}_R(o_\tau) \propto \exp(\beta R(o_\tau)) \quad (16)$$

where β is a parameter that determines how much the preference biases depending on the magnitude of the objective function.

2) Upper-level Agents: Upper-level agents are deployed in the macrocell and determine the Coordinated Multi-Point (CoMP) configuration (i.e., the group of lower-level base stations performing joint transmission) based on the observed states estimated by the lower-level agents. The predicted distribution of upper-level states is used as the prior distribution for the lower-level states.

a) Observations: The estimated results of the states $P(s_{b^l}(t)|o_{b^l}(t))$ in the lower-level base stations b^l are used as observations. Therefore, $\boldsymbol{o}_b(t) = P(s_{b^l}(t)|o_{b^l}(t))$ corresponds to the observations.

b) States: In the upper level, control is performed based on the channel coefficients $H^{u,b^l}(t)$ for each lower-level base station. Therefore, $s_b(t) = (H^{u,b_1^l}(t), H^{u,b_2^l}(t), \cdots)$ corresponds to the states.

c) Actions: In the upper level, the group of lower-level base stations performing joint transmission (JT) is determined. In other words, a subset of small cells is determined as $G \subseteq B^{(s)}$, and $a_b(t) = G$ is defined.

d) Preference Distribution: Similar to the lower level, the preference distribution in the upper level is a Boltzmann distribution based on power efficiency or transmission rate, but

it uses aggregated power efficiency and transmission rate for the entire system instead of individual values for each base station.

C. Beamforming for Multiple Terminals

In the case of multiple terminals, each UE is assigned a time slot, and only one UE is served simultaneously. Since each UE has physically different antennas, the channel coefficients $H^{u,b}$ between the UE and the base station are independent for each terminal. Therefore, for each UE u, an agent state $s_b^{(u)}$ is maintained, and the state is switched depending on the UE to which the transmission is directed, thereby achieving beamforming for multiple terminals.

Here, we assume that the assignment of time slots to UEs is determined in advance as an external control. Also, regarding the feedback from UEs to base stations, it is assumed to be the observation of the estimated time slots for each UE.

V. Evaluation

To confirm the effectiveness of beamforming coordination by vertical coordination using FEP, simulation will be used to verify the operation.

A. Setting

To verify the effectiveness of coordinated beamforming through vertical collaboration using FEP, we conducted simulations to confirm its operation.

1) System Configuration: We consider a system with 1 UE equipped with a single antenna and 2 base stations equipped with $N = 4$ antennas each. The goal is to transmit signals from the two base stations to the UE using beamforming techniques. Initially, we investigate the performance in a static environment where the channel coefficients remain fixed without temporal variations. The noise intensity is set to -114 dBm.

Following the approach in [9], we assume that the channel coefficients undergo time variations modeled as L-path Rayleigh fading. The channel coefficient $h_{ij}(t)$ between the i-th base station and the j-th antenna of the UE at time t can be expressed as:

$$h_{ij}(t) = \sqrt{\frac{\beta_{i,j}}{L}} \sum_{l=1}^{L} g_{i,j}(t,l) a_{i,j}^{\dagger}(\theta) \quad (17)$$

$$g_{i,j}(t,l) = \rho g_{i,j}(t-1,l) + \sqrt{1-\rho^2} e_{i,j}(t) \quad (18)$$

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

Here, $e_{i,j}(t)$ represents complex Gaussian noise, and ρ is a parameter that determines the temporal continuity of Rayleigh fading. θ_l denotes the transmission angle of path l, which follows a uniform distribution $(\bar{\theta}_l - \vartheta/2, \bar{\theta}_l + \vartheta/2)$. Figure 1 illustrates an example of the time variation of the channel gain $||H(t)||^2$ under fading conditions.

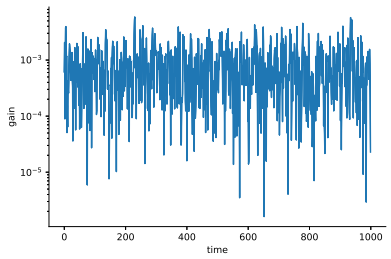


Fig. 1. Example of the time variation of the channel gain under fading conditions

2) Base Station Operation: Each base station observes the received SINR from the UE as feedback. Upon receiving the feedback, the base station updates its internal models and states and determines the transmit power and beamforming vector for the next transmission.

The beamforming vector is selected from a pre-defined set of N codebooks [1]. The i-th beamforming vector $\boldsymbol{w}_n$ is given by:

$$\boldsymbol{w}_n = (\frac{1}{\sqrt{N}} \exp(\frac{2\pi i(n-1)0}{N}), \cdots, \frac{1}{\sqrt{N}} \exp(\frac{2\pi i(n-1)(N-1)}{N})) \quad (19)$$

We use solver pymdp [15] for inference, action, and learning, in which, the SINR and channel coefficients are discritized into 10 bins and 50 bins.

3) Internal Model Initialization: The internal models $P(o_\tau|s_\tau,\pi)$ and $P(s_\tau|s_{\tau-1},\pi)$ can also be learned as FEP for expected free energy minimization. For this simulation, we set the initial models with random values as the initial models for FEP.

B. Results

In this section, we present the results of the evaluation for both single-user operation and multi-user switching scenarios.

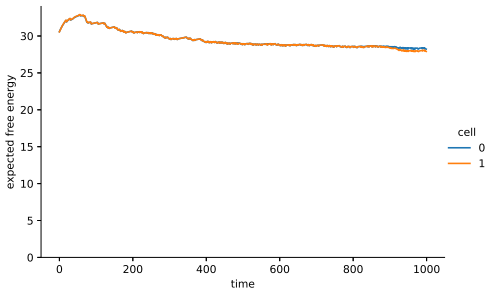
C. Single-User Operation

Figures 2 and 3 show the results of the expected free energy and SINR over time for single-user operation without and with vertical collaboration, respectively.

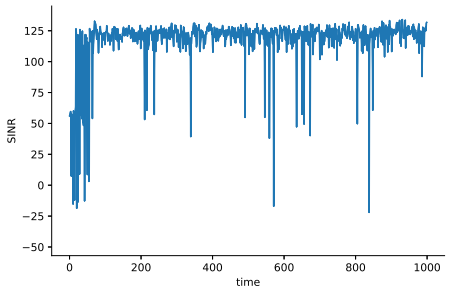
In Fig. 2, the results without vertical collaboration show the temporal variations of expected free energy and SINR when each base station independently performs beamforming. Fig. 3 illustrates the results with vertical collaboration, where the base stations share the estimated states through higher layers.

From both figures, it can be observed that the expected free energy decreases over time, indicating the improvement of the beamforming performance. However, due to the stochastic variations of the channel coefficients, the expected free energy shows small fluctuations and does not converge to zero, unlike in static environments.

The SINR, on the other hand, reaches a high level after a certain number of iterations in both cases, indicating that

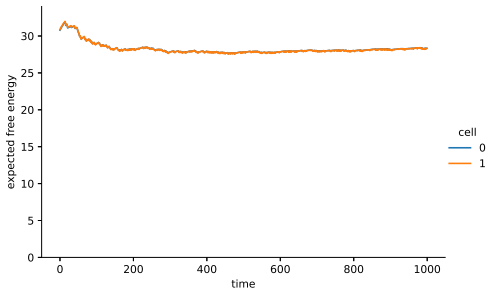


(a) Expected Free Energy

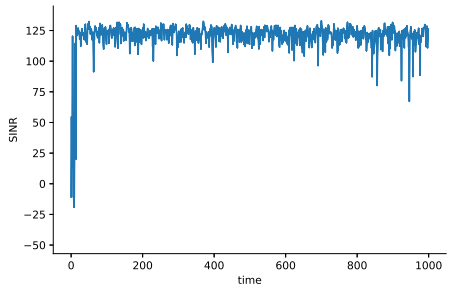


(b) SINR

Fig. 2. Temporal variations of expected free energy and SINR at the UE (without vertical collaboration)



(a) Expected Free Energy



(b) SINR

Fig. 3. Temporal variations of expected free energy and SINR at the UE (with vertical collaboration)

appropriate beams are selected. However, without vertical collaboration, temporary decreases in SINR are observed even after it initially becomes high. These temporary decreases are caused by the response of each base station to fine variations in fading, which affects the changes in beams of other base stations. In particular, without vertical collaboration, the independent adaptation of each base station to fine variations in fading leads to oscillations and further changes in the situation. In contrast, the sharing of state information through vertical collaboration allows the upper layers to predict the overall situation and suppress oscillations in control responses to such fine variations.

D. Multi-User Switching

Next, we consider the scenario of switching the transmission target among multiple UEs according to a given schedule. The schedule switches the UE every 300 time units.

Figures 4 and 5 show the results of the expected free energy and SINR over time for multi-user switching without and with vertical collaboration, respectively. Each UE is color-coded in the figures.

Fig. 4 shows the results without vertical collaboration, where each base station independently performs beamforming for the switching UEs according to the given schedule. Fig. 5 illustrates the results with vertical collaboration, where the base stations share the estimated states through higher layers.

From both figures, it can be observed that the expected free energy decreases for each UE during the switching process, indicating the improvement in beamforming performance.

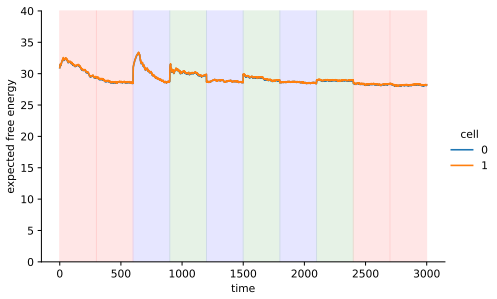
The SINR, in both cases, reaches a high level after a certain number of iterations and is maintained thereafter. However, without vertical collaboration, temporary decreases in SINR are observed even after it initially becomes high. These temporary decreases are caused by the response of each base station to fine variations in fading, which affects the changes in beams of other base stations. Nevertheless, the use of vertical collaboration enables stable control by predicting the overall situation based on higher-level information and transmitting it to lower-level layers.

Although each UE requires iterations to improve SINR, once the SINR improves, it remains at a high level. Even when a UE is switched during the process, the learned models and state estimation results are maintained, allowing the setting of beams tailored to the situation of each UE. In practical operations, since the models and states are continuously utilized, the time required for the initial iterations is not a significant issue.

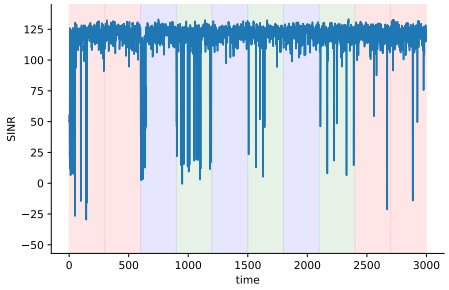
VI. Conclusion

In this paper, we proposed a collaborative beamforming approach, utilizing vertical collaboration between base stations to enhance the overall performance. The proposed approach leverages the sharing of estimated channel states and learned models among base stations to improve beamforming adaptation and mitigate the effects of fading variations.

Through evaluations in single-user operation scenarios, we observed that the expected free energy, which represents the beamforming performance, decreased over time in both cases with and without vertical collaboration. However, without verti-

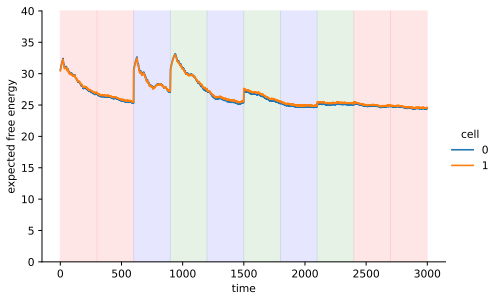


(a) Expected Free Energy

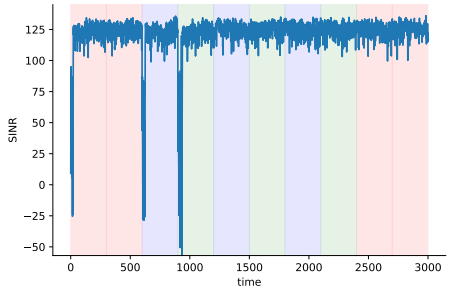


(b) SINR

Fig. 4. Temporal variations of expected free energy and SINR at the UE (without vertical collaboration, with 3 UE switching)



(a) Expected Free Energy



(b) SINR

Fig. 5. Temporal variations of expected free energy and SINR at the UE (with vertical collaboration, with 3 UE switching)

cal collaboration, the SINR showed temporary decreases due to the independent adaptation of each base station to fine variations in fading. In contrast, with vertical collaboration, the sharing of state information allowed for the prediction of the overall situation and suppression of oscillations in control responses, resulting in improved stability.

Furthermore, we investigated the multi-user switching scenario, where the transmission target switched among multiple users according to a predefined schedule.

Future work can explore further enhancements to the collaborative beamforming approach. One potential direction is to investigate horizontal collaboration between beamforming agents, enabling coordinated actions and information exchange to achieve better overall performance. Additionally, incorporating predictions of environmental time variations could improve the adaptability and robustness of the beamforming system.

Acknowledgment

This work was supported by JSPS KAKENHI Grant Number JP20K14735.

References

[1] L. Rao, M. Pant, L. Malviya, A. Parmar, and S. V. Charhate, "5G beamforming techniques for the coverage of intended directions in modern wireless communication: in-depth review," *International Journal of Microwave and Wireless Technologies*, vol. 13, no. 10, pp. 1039–1062, 2021.

[2] F. A. P. de Figueiredo, "An overview of massive MIMO for 5G and 6G," *IEEE Latin America Transactions*, vol. 20, no. 6, pp. 931–940, 2022.

[3] M. S. J. Solaija, H. Salman, A. B. Kihero, M. İ. Sağlam, and H. Arslan, "Generalized coordinated multipoint framework for 5G and beyond," *IEEE Access*, vol. 9, pp. 72 499–72 515, 2021.

[4] T. K. Tandra, F. Tajrian, A. Hossain, M. T. Kawser, M. R. Akram, and A. B. Shams, "Joint transmission coordinated multipoint on mobile users in 5G heterogeneous network," in *2022 IEEE 2nd Conference on Information Technology and Data Science (CITDS)*. IEEE, 2022, pp. 273–278.

[5] K. Friston, "The free-energy principle: a unified brain theory?" *Nature reviews neuroscience*, vol. 11, no. 2, pp. 127–138, 2010.

[6] G. Bartoli, R. Fantacci, K. B. Letaief, D. Marabissi, N. Privitera, M. Pucci, and J. Zhang, "Beamforming for small cell deployment in lte-advanced and beyond," *IEEE Wireless Communications*, vol. 21, no. 2, pp. 50–56, 2014.

[7] M.-L. Ku, L.-C. Wang, and Y.-L. Liu, "Joint antenna beamforming, multiuser scheduling, and power allocation for hierarchical cellular systems," *IEEE Journal on Selected Areas in Communications*, vol. 33, no. 5, pp. 896–909, 2014.

[8] D. Castanheira, P. Lopes, A. Silva, and A. Gameiro, "Hybrid beamforming designs for massive MIMO millimeter-wave heterogeneous systems," *IEEE Access*, vol. 5, pp. 21 806–21 817, 2017.

[9] K. Yu, G. Wu, S. Li, and G. Y. Li, "Local observations-based energy-efficient multi-cell beamforming via multi-agent reinforcement learning," *Journal of Communications and Information Networks*, vol. 7, no. 2, pp. 170–180, 2022.

[10] G. Pezzulo, F. Rigoli, and K. J. Friston, "Hierarchical active inference: a theory of motivated control," *Trends in cognitive sciences*, vol. 22, no. 4, pp. 294–306, 2018.

[11] J. Bauermeister and P. Lanillos, "The role of valence and meta-awareness in mirror self-recognition using hierarchical active inference," *arXiv preprint arXiv:2208.13213*, 2022.

[12] O. Çatal, T. Verbelen, T. Van de Maele, B. Dhoedt, and A. Safron, "Robot navigation as hierarchical active inference," *Neural Networks*, vol. 142, pp. 192–204, 2021.

[13] J. Pöppel, S. Kahl, and S. Kopp, "Resonating minds—emergent collaboration through hierarchical active inference," *Cognitive Computation*, vol. 14, no. 2, pp. 581–601, 2022.

[14] H. Al-Obiedollah, K. Cumanan, J. Thiyagalingam, A. G. Burr, Z. Ding, and O. A. Dobre, "Energy efficiency fairness beamforming designs for MISO NOMA systems," in *2019 IEEE Wireless Communications and Networking Conference (WCNC)*. IEEE, 2019, pp. 1–6.

[15] C. Heins, B. Millidge, D. Demekas, B. Klein, K. Friston, I. Couzin, and A. Tschantz, "pymdp: A python library for active inference in discrete state spaces," *arXiv preprint arXiv:2201.03904*, 2022.