

Risk-Aware Antenna Selection for Multiuser Massive MIMO under Incomplete CSI

Ke He, *Student Member, IEEE*, Thang X. Vu, *Senior Member, IEEE*, Dinh Thai Hoang, *Senior Member, IEEE*, Diep N. Nguyen, *Senior Member, IEEE*, Symeon Chatzinotas, *Fellow, IEEE*, and Björn Ottersten, *Fellow, IEEE*

Abstract—This paper investigates the antenna selection problem in massive multiple-input multiple-out (MIMO) systems under incomplete channel state information (CSI), with a particular interest on risk-aware planning subjected to practical constraints such as transmit power budgets and quality of services (QoS). Due to a very large number of antennas, obtaining complete channel measurements becomes a cost-prohibitive, energy-inefficient and spectral-inefficient task. To reduce pilot overhead, incomplete CSI and antenna selection (AS) are expected in practical massive MIMO systems. However, most existing AS algorithms heavily rely on the complete CSI, which imposes a high probability of violating the practical constraints in the scenarios of our interests. Motivated by this, we propose a joint channel prediction and antenna selection framework (JCPAS) which efficiently performs AS and is robust against the incomplete CSI and practical constraints. The proposed framework comprises i) a channel tracker which estimates the channel dynamics based on historical incomplete observations, and ii) a risk-aware Monte Carlo tree search (RA-MCTS) algorithm which utilizes the estimated channel dynamics to select antennas in a risk-aware manner. Simulation results show that the proposed RA-MCTS not only achieves much lower energy consumption compared to the existing typical algorithms, but also significantly reduces the probability of violating the practical constraints.

Index Terms—Massive MIMO, antenna selection, incomplete CSI, machine learning, risk-aware planning, Monte Carlo tree search

I. INTRODUCTION

MASSIVE multiple-input multiple-out (MIMO) has become the key technology to support the continuous development of future wireless network applications [2]–[5]. By deploying a very large number of antennas at the base station (BS), massive MIMO is capable of significantly improving the spectral efficiency via spatial multiplexing gain

Manuscript received August 25, 2023; revised December 30, 2023; accepted March 05, 2024. The associate editor handling this submission was Namyoon Lee. (*Corresponding author: Thang X. Vu*).

This work was funded in whole, or in part, by the Luxembourg National Research Fund (FNR), grant references FNR/IPBG19/14016225/INSTRUCT and FNR/C22/IS/17220888/RUTINE, and in part by the Australian Research Council under the DECRA project DE210100651. For the purpose of open access, and in fulfilment of the obligations arising from the grant agreement, the authors have applied a Creative Commons Attribution 4.0 (CC BY 4.0) license to any Author Accepted Manuscript version arising from this submission. Parts of this work were presented in the IEEE Globecom 2022 wks [1].

K. He, T. X. Vu, S. Chatzinotas, and B. Ottersten are with the Interdisciplinary Centre for Security, Reliability and Trust (SnT) - University of Luxembourg, L-1855 Luxembourg. E-mail: {ke.he, thang.vu, symeon.chatzinotas, bjorn.ottersten}@uni.lu.

D.T. Hoang and D. N. Nguyen are with the School of Electrical and Data Engineering, Faculty of Electrical and Engineering, University of Technology Sydney, Australia. Email: {hoang.dinh, diep.nguyen}@uts.edu.au.

[4]. However, the full potential of massive MIMO requires a huge number of dedicated radio frequency (RF) chains for every antenna, which results in not only increased capital expenditure (CAPEX) but also higher system energy consumption [3]–[5]. In practical massive MIMO systems, it is more cost-effective and energy-efficient to employ a number of RF chains less than the number of antennas, while the full spatial multiplexing gain can be preserved via antenna selection (AS) [4], [5]. Being a key component of hybrid signal processing techniques, AS aims to select the best subset of antennas for data transmission to reduce hardware cost and power consumption without losing the full potentials of antenna arrays [3]. During the last decades, the AS problem has been extensively studied in the presence of complete channel state information (CSI) (sometimes refer to full CSI), and it has been shown that AS can provide similar spectral efficiency with a lower energy consumption compared to the case without AS [4], [5].

A. Antenna Selection in Massive MIMO

In principle, the AS problem can be formulated as an integer programming problem under the assumption of complete CSI acquisition [5], [6]. The problem is NP-hard, thus solving it optimally imposes a prohibitive computational complexity which exponentially grows with the number of antennas in the worst case [6]. AS under complete CSI has been extensively investigated by researchers with the objective of designing a near-optimal algorithm with low computational complexities for massive MIMO [6]–[14]. However, it should be noted that most existing works were established in the presence of complete CSI, which implies that the channel coefficients of all antennas must be fully observed [15]–[17].

Unfortunately, this assumption does not strictly hold for practical scenarios, especially with very large-scale antenna arrays [18]. This is mainly because complete CSI acquisition turns out to be a time-consuming task in such scenarios. In practice, CSI acquisition is accomplished by pilot sequences which consume radio resources proportional to the numbers of active antennas and users. Since the number of RF chains is smaller than the number of antennas, only an incomplete CSI can be observed and estimated from each pilot transmission. This indicates that acquiring complete CSI eventually leads to extra pilot transmissions and reducing the effective transmission rate [1], [15], [17], [18]. For instance, acquiring complete CSI could occupy more than half of the frame duration for the downlink of a typical time-division duplexing (TDD) massive

MIMO system with 64 antennas, 16 RF chains and 4 single-antenna users [15], [19]. Therefore, it is of importance to take incomplete CSI acquisition into account in practical AS design, especially for very large antenna array configurations.

1) *Antenna Selection with Channel Prediction*: In order to address this issue, one promising approach would be the joint use of channel prediction and AS. It is straightforward that the extra pilot overhead can be reduced if we are able to predict the complete channel states. This approach works mainly because the real propagation environment is often time-varying and temporally correlated due to Doppler effects [20]–[22]. Based on this, researchers have developed different channel prediction algorithms in recent years [23]–[25]. In [23], the authors proposed a spatio-temporal autoregressive method for the prediction of the high mobility channel, where the prediction was performed utilizing the temporal correlation in the angle-delay domain. In [24], the authors predicted channel states by exploiting the channel correlations. The proposed method employed the convolutional neural network (CNN) and the long-short-term memory network (LSTM), which allows a multistep prediction of the channels. Similarly in [25], the authors predicted channel states by utilizing the spatio-temporal characteristics of CSI and a combination of CNN and convolutional LSTM. Although showing improved AS performance, these channel prediction methods are based on the historic complete CSI measurements [23]–[25]. It is noted that obtaining complete CSI is spectral-inefficient in the considered massive MIMO systems, especially when the number of antennas significantly exceeds the number of RF chains. In addition, the aforementioned channel predictions are deterministic methods, which limits the use of channel statistics for performance enhancement.

2) *Antenna Selection with Incomplete CSI*: Recently, researchers have focused on developing AS algorithms which can directly operate under incomplete CSI [1], [15], [17]. In [15], the authors formulated the AS problem as a combinatorial multi-armed bandit problem when only incomplete CSI is available, and proposed an online AS algorithm using Thompson sampling. However, it is assumed therein that each antenna contributes equally to the sum capacity. Since this assumption is not valid in most real propagation environments [18], the solution therein is not robust in practical scenarios. The authors of [17] considered the AS as a partially observable Markov decision process (POMDP) and proposed a myopic policy for selecting antennas under incomplete CSI. The myopic policy maintains a belief vector for the underlying channel states of each time slot, and updates this belief vector along with the system dynamics. The myopic policy therein, however, was only designed for single-user MIMO systems under general fading channels with a two-state coarse channel quantization. Since practical quality-of-service (QoS) constraints were not considered in [15], [17], their applicability is limited in practical massive MIMO systems.

Although several approaches have been proposed to address the antenna selection problem under incomplete CSI, further improvements are still needed to improve system performance and robustness. For the practical AS algorithm design, one crucial concern is how to maximize the system performance

while reducing the chance of violating the system' practical constraints when only incomplete CSIs are available. Risk-aware solutions using conditional value at risk (CVaR) have recently been proposed for resource management in ultra-reliable and low latency communications (URLLC) and the co-existence of eMBB and URLLC services [26]–[28]. However, these methods rely on the complete information of the system states, which are not applicable in the considered scenario.

B. Motivations and Contributions

As mentioned above, existing approaches cannot efficiently solve the antenna selection problem in multi-user massive MIMO and incomplete CSI. This motivates us to design a general antenna selection framework that can operate robustly against the complete CSI condition. Additionally, existing AS algorithms lack the capability to recognize the risk of violating the system constraints under incomplete CSI, which is essential to the required QoS. In general, risk awareness should be presented throughout the decision process, which implies that a desired antenna selection should be performed by jointly considering three important factors: practical system constraints, optimization objectives, and the uncertainties introduced by the incomplete CSI. Therefore, risk-aware planning remains a challenging issue for practical AS solutions, which is also the motivation of this work.

In this paper, we propose a joint channel prediction and antenna selection framework (JCPAS) for the antenna selection problem in multi-user massive MIMO under incomplete CSI and practical system constraints. The proposed JCPAS comprises a deep unsupervised learning-based conditional channel estimator and a risk-aware Monte Carlo tree search (RA-MCTS) algorithm. A risky event is identified when one of the system constraints cannot be satisfied. At each frame, the channel estimator maintains a belief distribution by estimating conditional channel statistics from the sequence of the past incomplete CSI measurements and estimates the posterior channel distribution. Based on the estimated posterior channel distribution, the RA-MCTS algorithm evaluates uncertain outcomes of each possible antenna combination through Monte Carlo simulations. In such a risk-aware manner, the chance of violating the system constraints can be reduced, and the corresponding negative consequences can also be mitigated when arisen. Simulation results show that the proposed RA-MCTS algorithm not only cuts the average power consumption by 50%, but also significantly reduces the probability of violating the system constraints by 90%.

To summarize, our main contributions are as follows:

- We introduce the JCPAS framework in massive MIMO systems under incomplete CSI. The proposed JCPAS framework does not require complete CSI measurements and is robust to conventional antenna selection algorithms. In addition, the proposed channel prediction method is a probabilistic model that can be used to enhance the performance of the integrated selection algorithm.
- We propose the RA-MCTS algorithm which enables efficient and robust antenna selection in massive MIMO

under incomplete CSI and practical system constraints. In contrast to the existing antenna selection algorithm, our proposed RA-MCTS is applicable to diverse optimization objectives and system constraints, and it is able to reduce the chance of violating the practical system constraints by leveraging channel statistics.

- We provide a new insight for risk-aware decision making with limited resources and insufficient information. In particular, a risk-aware system can be built by leveraging the historical incomplete observations to estimate a belief distribution over the underlying system dynamics and planning based on the statistics of the random outcomes introduced by incomplete observations accordingly.

The remainder of this paper is organized as follows. Section II describes the mathematical model of the considered massive MU-MIMO system, as well as the associated antenna selection problem. Section III presents the proposed deep unsupervised learning-based conditional channel estimator. Section IV presents the RA-MCTS algorithm, which is a risk-aware planning algorithm for selecting antennas with incomplete observations. Section V discusses the related simulation results. Finally, Section VI concludes the paper.

II. PRELIMINARIES

In this section, we introduce the system model of the considered massive MIMO system as well as the associated AS problem. After that, we will review the greedy search AS algorithm under the complete CSI assumption.

A. System Model

As shown in Fig. 1, we consider the downlink of a massive MU-MIMO system where a BS serves N_u single-antenna users. The BS is equipped with N_t transmit antennas and N_f ($0 < N_f \ll N_t$) RF chains. In addition, switches are also available at the BS such that an RF chain can connect with any antenna of interest. The channel between the BS and users is time-varying and temporally correlated, which is common in real propagation environments with Doppler effects [20]–[22]. Moreover, we assume that the CSI remains unchanged within each frame duration of T channel uses (c.u.), and the considered system operates in TDD mode, meaning that we have identical channels for both uplink and downlink transmission due to channel reciprocity [6], [9]. On the downlink transmission, the CSI acquisition is accomplished via uplink pilot-assisted channel measurement, and multi-user precoding is then adopted to mitigate inter-user interference. Under these settings, we further denote τ_{csi} as the number of c.u. consumed to acquire CSI, resulting in $T - \tau_{csi}$ c.u. for data transmission.

Since only N_f out of N_t transmit antennas can be activated at the same time, the BS needs to select the best N_f antennas in terms of performance metrics maximization. Let $\mathbf{a} = \{a_1, a_2, \dots, a_j, \dots, a_{N_f}\}$ be the set of the indices of the N_f selected antennas, and we denote \mathcal{A} as the set of all possible antenna combinations with a cardinality of $|\mathcal{A}| = \binom{N_t}{N_f}$. In addition, let $\mathbf{H} \in \mathbb{C}^{N_u \times N_t}$ be the complete CSI matrix, and we denote $\mathbf{H}(\mathbf{a})$ as the incomplete (or partial) CSI

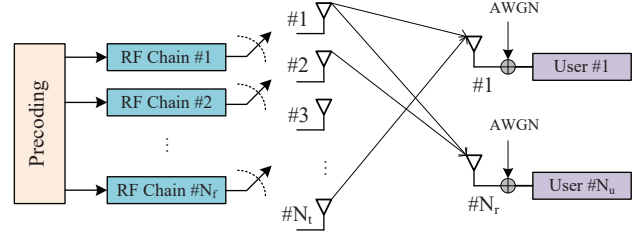


Fig. 1: Structure of the considered massive MIMO system.

from the chosen combination $\mathbf{a} \in \mathcal{A}$, meaning that the columns of $\mathbf{H}(\mathbf{a})$ are selected from \mathbf{H} with respect to the indices in \mathbf{a} . Besides, we denote x_k as the data symbol to be transmitted to user k , and $\mathbb{E}\{|x_k|^2\} = 1$. Then, the received signal $y_k(\mathbf{a})$ at user k is given by

$$y_k(\mathbf{a}) = \mathbf{h}_k(\mathbf{a})\mathbf{w}_k x_k + \sum_{j \neq k} \mathbf{h}_j(\mathbf{a})\mathbf{w}_j x_j + n_k, \quad (1)$$

where $\mathbf{h}_k(\mathbf{a}) \in \mathbb{C}^{1 \times N_f}$ is the channel vector for user k from the antenna combination \mathbf{a} , \mathbf{w}_k denotes the $N_f \times 1$ precoding vector for user k , and $n_k \sim \mathcal{CN}(0, \sigma_k^2)$ is the additive white Gaussian noise (AWGN) at user k . The second term of (1) is the inter-user interference at user k .

Assuming negligible processing time, the *effective spectral efficiency* for downlink transmission to user k with the selected antennas \mathbf{a} can be written as

$$R_k(\mathbf{a}) = \left(1 - \frac{\tau_{csi}}{T}\right) \log_2(1 + \text{SINR}_k(\mathbf{a})), \quad (2)$$

where $\text{SINR}_k(\mathbf{a}) = \frac{\|\mathbf{h}_k(\mathbf{a})\mathbf{w}_k(\mathbf{a})\|^2}{\sigma_k^2 + \sum_{j \neq k} \|\mathbf{h}_j(\mathbf{a})\mathbf{w}_j(\mathbf{a})\|^2}$ represents the signal-to-interference-noise ratio (SINR) of user k . Accordingly, the *effective sum spectral efficiency* with antenna combination \mathbf{a} can be bounded by $C(\mathbf{a}) = \sum_{k=1}^{N_u} R_k(\mathbf{a})$, and the total power consumption for transmitting data at each frame is given by $P(\mathbf{a}) = \sum_{k=1}^{N_u} \|\mathbf{w}_k(\mathbf{a})\|^2$.

B. Antenna Selection with Objective Maximization

For practical scenarios, a typical objective for selecting the best subset of antennas is to optimize a generic objective function $\mathcal{F}_k(\mathbf{a})$ under the constraints of total transmit power and minimum QoS requirements. Mathematically, the optimization problem can be formulated as

$$\underset{\mathbf{a} \in \mathcal{A}}{\text{maximize}} \quad \mathcal{F}(\mathbf{a}), \quad (3)$$

$$\text{subject to} \quad P(\mathbf{a}) \leq P_{tot}, \quad (4)$$

$$R_k(\mathbf{a}) \geq \eta_k, \forall k, \quad (5)$$

where η_k is the QoS requirement for user k , P_{tot} denotes the total transmit power, and $\mathcal{F}(\mathbf{a})$ is the objective function of interest. According to the specific problem, the objective function can be $\mathcal{F}(\mathbf{a}) = -\sum_{k=1}^{N_u} \|\mathbf{w}_k(\mathbf{a})\|^2$ when we want to minimize the energy consumption, and $\mathcal{F}(\mathbf{a}) = \sum_{k=1}^{N_u} R_k(\mathbf{a})$ if we want to maximize the sum-throughput for the system.

1) *Antenna Selection with Complete CSI*: Solutions to problem (3) have been well studied in the literature under the *complete CSI* assumption [10]–[14], that it is possible to fully observe the channel states with an affordable overhead. Among these solutions, the idea of greedy search is widely

Algorithm 1 Greedy Search Algorithm

Input: Full channel matrix \mathbf{H}

Output: Selected antenna combination \mathbf{a}

- 1: Let the set of all antennas be $\mathcal{N} = \{1, \dots, N_t\}$;
 - 2: Initialize the set of removed antennas as $\mathbf{r} = \emptyset$;
 - 3: Define $\mathbf{A} = (\mathbf{H}\mathbf{H}^H)^{-1}$;
 - 4: **while** $|\mathbf{r}| < N_t - N_f$ **do**
 - 5: $m = \arg \min_{r \in \mathbf{a}} \frac{\|\mathbf{h}_r^H \mathbf{A}\|^2}{1 - \mathbf{h}_r^H \mathbf{A} \mathbf{h}_r}$;
 - 6: $\mathbf{A} = \mathbf{A} + \frac{\mathbf{A} \mathbf{h}_m \mathbf{h}_m^H \mathbf{A}}{1 - \mathbf{h}_m^H \mathbf{A} \mathbf{h}_m}$;
 - 7: $\mathbf{r} = \mathbf{r} \cup \{m\}$;
 - 8: **return** $\mathcal{N} - \mathbf{r}$;
-

adopted due to its good performance and low-complexity [6]. Let \mathbf{a}^p be a set of p selected antenna indices, and let $\mathbf{a}^q \supset \mathbf{a}^p$ be a superset of \mathbf{a}^p with $q > p$. In addition, we denote $\mathbf{a}^{q-p} = \mathbf{a}^q \setminus \mathbf{a}^p$ as the difference between set \mathbf{a}^q and \mathbf{a}^p . According to the Proposition 1 of [9], the spectral efficiency loss of removing $q - p$ antennas from \mathbf{a}^q can be bound by

$$C(\mathbf{a}^{q-p}) = \log_2 \left(1 + \frac{P_{tot} \cdot \text{tr}(\mathbf{\Lambda}_{q-p})}{\text{tr}(\mathbf{Q}_q)^2 + \text{tr}(\mathbf{Q}_q) (P_{tot} + \text{tr}(\mathbf{\Lambda}_{q-p}))} \right),$$

with the following notations

$$\mathbf{\Lambda}_{q-p} = \mathbf{Q}_q \mathbf{H}(\mathbf{a}^{q-p}) \mathbf{A}_{q-p}^{-1} \mathbf{H}(\mathbf{a}^{q-p})^H \mathbf{Q}_q, \quad (6)$$

$$\mathbf{A}_{q-p} = (\mathbf{I} - \mathbf{H}(\mathbf{a}^{q-p})^H \mathbf{Q}_q \mathbf{H}(\mathbf{a}^{q-p})), \quad (7)$$

$$\mathbf{Q}_q = (\mathbf{H}(\mathbf{a}^q) \mathbf{H}(\mathbf{a}^q)^H)^{-1}, \quad (8)$$

where $\text{tr}(\cdot)$ denotes the matrix trace. Now, the sum-throughput maximization problem can be converted as removing $N_t - N_f$ antennas with the minimum capacity loss, which is given by

$$\mathbf{a}^* = \arg \min_{\mathbf{a} \in \mathcal{A}} \text{tr}(\mathbf{\Lambda}_{N_t - N_f}). \quad (9)$$

Unfortunately, the above equation still needs exhaustive search to find the best antenna combination to be removed. Nevertheless, it is possible to reduce the computational complexity by utilizing the concept of greedy search. Then, each time the greedy search suggests to remove the worst antenna that contributes least to the capacity, resulting in that the search space can be reduced significantly while still reserving a good performance. According to Proposition 2 of [9], the antenna to be removed at each iteration is given by

$$m = \arg \min_{r \in \mathbf{a}^p} \frac{\left\| \mathbf{h}_r^H (\mathbf{H}(\mathbf{a}^p) \mathbf{H}(\mathbf{a}^p)^H)^{-1} \right\|^2}{1 - \mathbf{h}_r^H (\mathbf{H}(\mathbf{a}^p) \mathbf{H}(\mathbf{a}^p)^H)^{-1} \mathbf{h}_r}, \quad (10)$$

where \mathbf{a}^p is currently selected antenna set. By repeating this strategy, the antenna combination which maximizes the sum-throughput will eventually be determined [9], and the pseudocodes of the resulting greedy search algorithm is shown in Algorithm 1. It is worth noting that Algorithm 1 is directly applied to generalized ZF by using $\mathbf{A} = (\alpha \mathbf{I} + \mathbf{H}\mathbf{H}^G)^{-1}$.

2) *Antenna Selection with Incomplete CSI:* Due to the limited number of RF chains at BS, only the partial CSI corresponding to the N_f selected antennas could be measured in each pilot transmission. In this case, we need $\tau_{csi} = N_u \lceil \frac{N_t}{N_f} \rceil$ c.u. to acquire the complete CSI [15], and obviously, τ_{csi} will

quickly become cost-prohibitive for large N_t , where extra pilot overhead grows rapidly in massive MIMO and results in a reduced effective transmission rate, as shown in (2). For this reason, acquiring complete CSI is thereby a very inefficient strategy for massive MIMO systems, which motivates us to study the antenna selection problem in the presence of incomplete (or partial) CSI.

III. PROPOSED ANTENNA SELECTION FRAMEWORK WITH INCOMPLETE CSI

In this section, we propose a joint channel prediction and antenna selection (JCPAS) framework operating without relying on a complete CSI assumption. JCPAS consists of two main blocks: *Channel tracking* and *Antenna selection*. The first block learns the belief of the complete channel matrix from historical incomplete channel estimates (of selected antennas) and predicts the complete channel matrix. Based on the predicted complete channel matrix, the second block selects the best antenna subset. To illustrate our proposed JCPAS framework, we present its workflow and time horizon structure in Fig. 2 for the ease of understanding.

A. Channel Tracking with Historical Incomplete Observations

In practical scenarios with large-scale antenna configurations, incomplete measurements of the current channel state are expected due to the limited number of RF chains and channel estimation duration. This limitation indicates that the antenna selection will be performed in the presence of incomplete channel information, in which existing complete CSI based antenna selection algorithms cannot be efficiently applied. Therefore, it is of vital importance to track the transition of channel states by exploiting the temporal correlation of real propagation environments, and thereby maintaining an accurate belief on the current channel state to help AS. We note that Kalman filtering based prediction is not applicable since it requires the complete historic CSI measurements

From this perspective, we hereby propose a deep conditional generative model (DCGM) to estimate the distribution of the belief state based on the sequence of past incomplete channel measurements. Mathematically, the channel tracking model can be considered as a probabilistic generative model $\mathcal{P}(\mathbf{H}_t | \Phi_t)$, which is conditioned on the history of incomplete CSI measurements, denoted by $\Phi_t \triangleq \{\mathbf{H}_{t-L+1}(\mathbf{a}_{t-L+1}), \dots, \mathbf{H}_{t-1}(\mathbf{a}_{t-1}), \mathbf{H}_t(\mathbf{a}_t)\}$ with a finite horizon L . In general, the value of L is determined by the storage and computing resources. Note that the missing entries will be replaced by zero entries in the sequence if the current time slot $t < L$. Although it is hard to estimate the exact posterior distribution $\mathcal{P}(\mathbf{H}_t | \Phi_t)$, we can still get its accurate approximation by the maximum likelihood approximation.

To this end, we begin by denoting the approximate distribution as $\mathcal{Q}_\theta(\mathbf{H}_t | \Phi_t)$ with parameters θ . Then, we train the model with the objective of maximizing the likelihood of the training samples on the chosen distribution, which is given by

$$\mathcal{N}(\theta) = \arg \min_{\theta} \frac{1}{N_s} \sum_{t=1}^{N_s} -\log \mathcal{Q}_\theta(\mathbf{H}_t | \Phi_t), \quad (11)$$

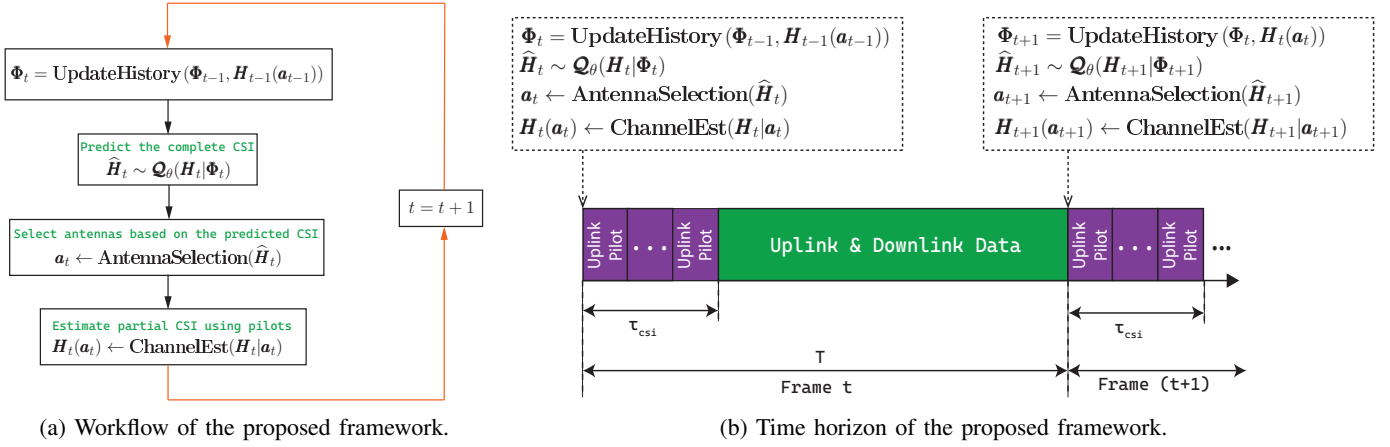


Fig. 2: Diagram of the proposed JCPAS framework.

where N_s denotes the size of the training sample set $\mathcal{D} = \{\mathbf{H}_t, \Phi_t\}_{t=1}^{N_s}$. Clearly, (11) quantifies the magnitude of fitness between the chosen distribution $Q_\theta(\mathbf{H}_t | \Phi_t)$ and the samples drawn from the real distribution $\mathbf{H}_t \sim \mathcal{P}(\mathbf{H}_t | \Phi_t)$. In particular, $Q_\theta(\mathbf{H}_t | \Phi_t)$ accurately approximates $\mathcal{P}(\mathbf{H}_t | \Phi_t)$ when (11) achieves its minimum. On the other hand, $Q_\theta(\mathbf{H}_t | \Phi_t)$ deviates from the real distribution when (11) enlarges. However, prior knowledge on the real distribution is needed to select an appropriate model for approximation, which is unpractical in the circumstances of our interest.

In order to solve this issue, we employ a deep normalizing flow (DNF) to construct $Q_\theta(\mathbf{H}_t | \Phi_t)$. Compared with other generative models, e.g., variational auto-encoders (VAEs) and generative adversarial networks (GANs), DNF is a fully probabilistic model with tractable exact density inference, which can accelerate the search efficiency of Monte-Carlo tree search [29]. Although in this paper the exact density of the predicted channel states is not utilized, it can help reduce the number of simulations in our future work by evaluating the certainty of current solutions. As a kind of generative model, DNF approximates the distribution by the change of latent distribution. This strategy allows us to sample complete observations from the latent space, while still being able to compute the corresponding log-likelihood by the law of change of variables [30], [31]. In principle, DNF assumes that complete observation \mathbf{H}_t depends on a latent random variable \mathbf{Z}_t following a tractable distribution $\mathcal{P}_\omega(\mathbf{Z}_t)$, where ω is the parameters of the latent distribution. It is also assumed that ω follows a tractable distribution, denoted by $\omega \sim \mathcal{P}_\psi(\omega)$ with parameters ψ . Besides, the parameters ψ can be determined based on the history, i.e., $\psi = \gamma_{\theta_1}(\Phi_t)$ in which $\gamma_{\theta_1}(\cdot)$ is a function represented by a deep neural network (DNN) with parameters θ_1 .

Intuitively, this approach takes the uncertainty of incomplete observations into channel tracking by considering ω as a random variable conditioned on the given history Φ_t . Therefore, the latent space is also conditioned on incomplete observations, denoted as $\mathbf{Z}_t \sim \mathcal{P}_\omega(\mathbf{Z}_t | \Phi_t)$. In conclusion, the

generative process can be described as

$$\begin{aligned} \psi_t &= \gamma_{\theta_1}(\Phi_t)'; & \omega_t &\sim \mathcal{P}_{\psi_t}(\omega_t) \\ \mathbf{Z}_t &\sim \mathcal{P}_{\omega_t}(\mathbf{Z}_t); & \mathbf{H}_t &= g_{\theta_2}(\mathbf{Z}_t), \end{aligned} \quad (12)$$

where $g_{\theta_2}(\cdot)$ represents an *invertible (or bijective)* function with parameters θ_2 . Since $g_{\theta_2}(\cdot)$ is an invertible function, the associated latent variable can be effectively inferred by $\mathbf{Z}_t = f_{\theta_2}(\mathbf{H}_t) \triangleq g_{\theta_2}^{-1}(\mathbf{H}_t)$. By setting $\theta = \{\theta_1, \theta_2\}$, the log-likelihood of the complete observation \mathbf{H}_t can be approximately computed from

$$\log Q_\theta(\mathbf{H}_t | \Phi_t) = \log \mathcal{P}_\omega(f(\mathbf{H}_t) | \Phi_t) + \log \left| \det \left(\frac{d\mathbf{f}}{d\mathbf{H}_t} \right) \right|, \quad (13)$$

where $\det \left(\frac{d\mathbf{f}}{d\mathbf{H}_t} \right)$ is the determinant of the Jacobian. In order to construct a flexible model $Q_\theta(\mathbf{H}_t | \Phi_t)$, we assume that the invertible function $f(\cdot)$ is composed by N_I invertible sub-functions, given by

$$f(\cdot) = f_1(\cdot) \otimes f_2(\cdot) \otimes \cdots \otimes f_n(\cdot) \cdots \otimes f_{N_I}(\cdot). \quad (14)$$

Based on the above factorization, we can infer the corresponding latent variable \mathbf{Z} accordingly,

$$\mathbf{H} \xrightarrow{f_1} \mathbf{V}_1 \xrightarrow{f_2} \mathbf{V}_2 \cdots \xrightarrow{f_n} \mathbf{V}_n \cdots \xrightarrow{f_{N_I}} \mathbf{Z}. \quad (15)$$

Now, using the notations $\mathbf{V}_0 \triangleq \mathbf{H}_t$ and $\mathbf{V}_{N_i} \triangleq \mathbf{Z}_t$, we can rewrite (13) as

$$\begin{aligned} \log Q_\theta(\mathbf{H}_t | \Phi_t) &= \log \mathcal{P}_\omega(f(\mathbf{H}_t) | \Phi_t) \\ &+ \sum_{n=1}^{N_I} \log \left| \det \left(\frac{d\mathbf{V}_n}{d\mathbf{V}_{n-1}} \right) \right|. \end{aligned} \quad (16)$$

Thus, we construct the approximate model $Q_\theta(\mathbf{H}_t | \Phi_t)$ with N sub-functions, with each sub-function being a small flow step of the complete flows. In this case, it is straightforward to train the model Q_θ by recalling the principle of maximum likelihood approximation in (13).

B. Network Implementation

The applied network structure of the proposed DCGM is illustrated in Fig. 3. As mentioned in Sec. III-A, it is important to ensure that each flow step is fully invertible for the implementation of the proposed DCGM. In our implementation, we

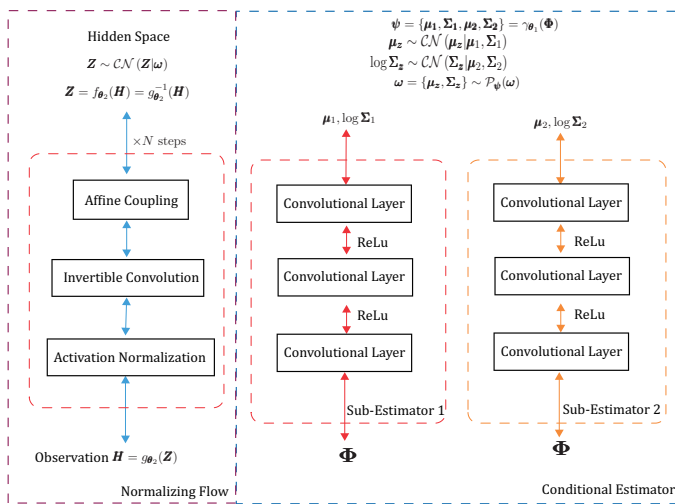


Fig. 3: Diagram of the structure of the proposed DCGM.

implement the flow steps via normalization layers, invertible convolutional layers and affine coupling layers, where the details of these layers can be found in [1], [30], [32], [33]. Using the above invertible layers, an invertible network can be constructed to track channel transitions from the incomplete history. In particular, the constructed invertible network is composed by N flow steps, and each flow step containing only three layers: activation normalization layer, invertible convolutional layer and affine coupling layer.

Because the latent variable \mathbf{Z}_t relies on the incomplete history Φ_t , a conditional distribution $\mathcal{P}_{\mathbf{Z}}(\mathbf{Z}_t|\Phi_t)$ should also be implemented. Consider $\mathbf{Z}_t \sim \mathcal{CN}(\mathbf{Z}_t|\boldsymbol{\mu}_z(t), \boldsymbol{\Sigma}_z(t))$ and denote $\boldsymbol{\omega}_t = \{\boldsymbol{\mu}_z(t), \boldsymbol{\Sigma}_z(t)\}$ as the parameters of the distribution. Then, we sample $\boldsymbol{\omega}_t$ by the reparameterization steps

$$\boldsymbol{\mu}_z(t) = \boldsymbol{\mu}_1(t) + \boldsymbol{\nu}_1 \odot \boldsymbol{\Sigma}_1(t), \quad (17)$$

$$\boldsymbol{\Sigma}_z(t) = \boldsymbol{\mu}_2(t) + \boldsymbol{\nu}_2 \odot \boldsymbol{\Sigma}_2(t), \quad (18)$$

where $\boldsymbol{\nu}_1$ and $\boldsymbol{\nu}_2$ are two standard complex Gaussian samples, \odot denotes the element-wise multiplication, and the associated parameters $\boldsymbol{\psi}_t = \{\boldsymbol{\mu}_1(t), \boldsymbol{\mu}_2(t), \boldsymbol{\Sigma}_1(t), \boldsymbol{\Sigma}_2(t)\}$ are determined by the incomplete history, i.e., $\boldsymbol{\psi}_t = \gamma_{\theta_1}(\Phi_t)$. Specifically, $\gamma_{\theta_1}(\Phi_t)$ is represented by two independent convolutional networks, which can be expressed as $\{\boldsymbol{\mu}_1(t), \boldsymbol{\Sigma}_1(t)\} = \text{CNN}_1(\Phi_t)$ and $\{\boldsymbol{\mu}_2(t), \boldsymbol{\Sigma}_2(t)\} = \text{CNN}_2(\Phi_t)$. $\text{CNN}_1(\cdot)$ and $\text{CNN}_2(\cdot)$ may both have exactly the same network structure composed by multiple convolutional layers and rectified linear units (ReLU) [30]. In order to retain the spatial information, zero-padding is used to keep each incomplete observation $\mathbf{H}_t(\mathbf{a}_t)$ within Φ_t having the same shape of $N_u \times N_t$.

C. The Proposed Antenna Selection Framework

Given the prediction of complete CSI from the incomplete history, the antenna selection becomes a straightforward task. The proposed JCPAS framework utilizes the proposed channel tracking model $\mathcal{Q}_{\theta}(\mathbf{H}_t|\Phi_t)$ to select the antennas in practical environments. The workflow of the proposed JCPAS framework is presented in Fig. 2.

At the beginning of each frame, we estimate the belief of the current channel state via the well-trained DCGM, i.e., $\hat{\mathbf{H}}_t \sim$

Algorithm 2 Proposed Joint Channel Prediction and Antenna Selection (JCPAS) Framework

- 1: Set the current frame index $t = 0$;
- 2: Initialize history as $\Phi_t = \{\mathbf{0}_{N_u \times N_t}, \dots, \mathbf{0}_{N_u \times N_t}\}$;
- 3: **loop**
- 4: $\hat{\mathbf{H}}_t \sim \mathcal{Q}_{\theta}(\mathbf{H}_t|\Phi_t)$
- 5: $\mathbf{a}_t = \text{ANTENNA_SELECTION}(\hat{\mathbf{H}}_t)$ (e.g., Algorithm 1)
- 6: Obtain the incomplete channel measurements $\mathbf{H}_t(\mathbf{a}_t)$ with \mathbf{a}_t ;
- 7: $\text{PRECODING}(\mathbf{H}_t(\mathbf{a}_t))$
- 8: $\text{DATA_TRANSMISSION}(\mathbf{H}_t(\mathbf{a}_t))$
- 9: **for** $i = 0$ to $\min(t, L)$ **do**
- 10: $\Phi_t[L - i] = \mathbf{H}_{t-i}(\mathbf{a}_{t-i})$;
- 11: $t = t + 1$;

$\mathcal{Q}_{\theta}(\mathbf{H}_t|\Phi_t)$. It should be noted that we fill in the history by zero entries for initialization. After estimating the belief state, we employ an antenna selection algorithm (e.g., Algorithm 1) to select the antennas subset for acquiring the incomplete CSI as well as data transmission. When the data transmission of the current frame is completed, we update the history for the next frame by pushing the last incomplete observation $\mathbf{H}_t(\mathbf{a}_t)$ into Φ_t .

To summarize the process of our proposed JCPAS framework, the associated pseudo-code is detailed in Algorithm 2. It is worth noting that the proposed framework is a general framework with the purpose of reducing the channel estimation overhead for massive MIMO systems. Thus, the choice of antenna selection algorithms is not limited, and hereby we use the greedy search algorithm for illustration. In principle, the choice of antenna selection algorithms should be determined based on the available resources and different needs of environments.

IV. PROPOSED RISK-AWARE PLANING ANTENNA SELECTION ALGORITHM

This section will introduce the proposed risk-aware planning algorithm, which can be integrated into the JCPAS framework to further improve system robustness, and meanwhile reducing the chance of violating practical constraints. After that, we will discuss the performance-complexity tradeoff of the proposed algorithm.

A. Risk-Aware Monte Carlo Tree Search

Due to incomplete CSI and imperfect belief estimation $\mathcal{Q}_{\theta}(\mathbf{H}_t|\Phi_t)$, antenna selection often carries the risk of violating practical constraints of problem (3), as the information is not perfect for making decisions. In principle, the outcome $\mathcal{F}(\mathbf{a}_t)$ of the selected antennas \mathbf{a}_t can be regarded as a random variable conditioned on the history Φ_t . This indicates that making decisions based on the inspection of a single sample from the estimated belief distribution is obviously not sufficient and risky.

To be specific, such potential risk results from two parts: the probability of selecting antennas based on a sampled belief deviating from the truth, and the negative consequences if it

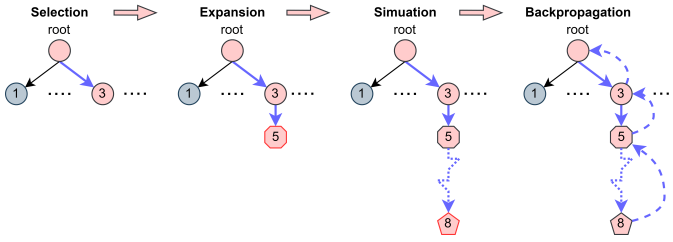


Fig. 4: Workflow of MCTS. Nodes to be visited are highlighted in red, and the search path is highlighted as blue arrows. The simulation and backpropagation paths are illustrated by blue dashed curves.

does. Unfortunately, the existing algorithms depend on the known channel coefficients rather than the channel statistics. To solve this issue, planning based on the expected outcomes should play a crucial role in risk-aware AS algorithm design. This realization is a key to heuristic risk-aware planning in the presence of incomplete observations and to guarantee the constraints.

From this perspective, we propose to select antennas and manage risks (of violating the system constraints) by learning a posterior distribution over the expected outcomes of each selected antennas combination. This approach can be accomplished based on the concepts of MCTS [34] and bootstrap Thompson sampling (BTS) [35]–[38]. As a best-first search strategy, MCTS employs a heuristic exploration to iteratively explore the combinatorial search space. In general, the search space can be regarded as a decision tree consisting of decision nodes, and each of which has a number of child nodes, and each child node corresponds to an available decision of removing the associated antenna index. As illustrated in Fig. 4, the typical routine of MCTS includes the following four steps [34]:

- **Selection:** We traverse the search tree in accordance to the estimated statistics of each node until encountering a node that has not been fully expanded, which is also called as a leaf node.
- **Expansion:** Whenever a leaf node is selected, it must be expanded. The expansion is done by randomly generating a child node and then initializing the prior information for the newly generated node.
- **Simulation:** We execute a random rollout through Monte Carlo simulations until a complete selection is reached. We then simulate the outcome of the explored complete selection by Monte Carlo simulation.
- **Backpropagation:** After receiving the simulated outcome of a complete selection, the results will be back-propagated to all ascendant nodes, in which a set of predefined algorithm statistics should be updated accordingly.

During the search, the above routine will be repeated a number of times such that the combinatorial search space will be incrementally explored and the search tree will be simultaneously expanded. After a number of simulations, the statistics of each node will be sufficient for making decisions.

As described in Algorithm 3 and Algorithm 4, we jointly use $\mathcal{Q}_\theta(\mathbf{H}_t|\Phi_t)$ and MCTS to build the proposed RA-MCTS

Algorithm 3 Risk-Aware Monte Carlo Tree Search (RA-MCTS)

Input: History of the past incomplete measurements Φ_t
Output: Selected antenna combination \mathbf{a}_t

- 1: Let the set of all antennas be $\mathcal{N} = \{1, \dots, N_t\}$;
- 2: Initialize the set of removed antennas as $\mathbf{r} = \emptyset$;
- 3: **while** $|\mathbf{r}| < N_t - N_f$ **do**
- 4: $m = \text{RISKAWAREPLANNING}(\mathbf{r}, \Phi_t)$
- 5: $\mathbf{r} = \mathbf{r} \cup \{m\}$;
- 6: $\mathbf{a}_t \leftarrow \{1, 2, \dots, N_t\} - \mathbf{r}$;
- 7: **return** \mathbf{a} ;

- 8: **function** $\text{RISKAWAREPLANNING}(\mathbf{r}, \Phi_t)$
- 9: **while** within computational budget **do**
- 10: $\hat{\mathbf{H}}_t \sim \mathcal{Q}_\theta(\mathbf{H}_t|\Phi_t)$;
- 11: $\mathbf{A} \leftarrow (\hat{\mathbf{H}}_t \hat{\mathbf{H}}_t^H)^{-1}$;
- 12: $n \leftarrow \text{RETRIEVENODE}(\mathbf{r}, \emptyset, \mathbf{A}, \hat{\mathbf{H}}_t)$;
- 13: $v = \text{TREEPOLICY}(n)$;
- 14: $\delta = \text{ROLLOUT}(v)$;
- 15: $\text{BACKPROPAGATION}(\delta, v)$;
- 16: **return** $\text{BESTCHILD}(n)$;

- 17: **function** $\text{RETRIEVENODE}(\mathbf{r}, \mathbf{r}', \mathbf{A}, \hat{\mathbf{H}})$
- 18: **if** the node corresponding \mathbf{r} has not been generated **then**
- 19: **for** $m \in \mathbf{r} - \mathbf{r}'$ **do**
- 20: $\mathbf{A} \leftarrow \mathbf{A} + \frac{\mathbf{A} \hat{\mathbf{h}}_m \hat{\mathbf{h}}_m^H \mathbf{A}}{1 - \hat{\mathbf{h}}_m^H \mathbf{A} \hat{\mathbf{h}}_m}$;
- 21: **if** $|\mathbf{r}| < N_t - N_f$ **then**
- 22: $\lambda_i \leftarrow \frac{\|\hat{\mathbf{h}}_i^H \mathbf{A}\|^2}{1 - \hat{\mathbf{h}}_i^H \mathbf{A} \hat{\mathbf{h}}_i}$, $\forall i \in \mathcal{N} - \mathbf{r}$;
- 23: $\alpha_i \leftarrow \frac{e^{-\lambda_i}}{\sum_{j \neq i} e^{-\lambda_j}}$, $\forall i \in \mathcal{N} - \mathbf{r}$;
- 24: **for** each child node $i \in \mathcal{N} - \mathbf{r}$ **do**
- 25: $\alpha_{ij} \leftarrow \alpha_i$, $\forall j \in \{1, \dots, J\}$;
- 26: $\beta_{ij} \leftarrow 1$, $\forall j \in \{1, \dots, J\}$;
- 27: **return** a node with the above associated statistics;
- 28: **else**
- 29: **return** the already generated node;

algorithm. The proposed RA-MCTS works in a similar manner as the greedy search, which removes antennas one by one iteratively. Each rollout is simulated based on the sample drawn from the estimated belief distribution $\mathcal{Q}_\theta(\mathbf{H}_t|\Phi_t)$. Besides, we maintain an approximate posterior distribution over the expected outcome of simulations, and utilize BTS to compute policies in risk-aware and multi-constraints settings.

For each simulation, the proposed RA-MCTS always starts at a root node with a belief sample $\hat{\mathbf{H}}_t \sim \mathcal{Q}_\theta(\mathbf{H}_t|\Phi_t)$ drawn from the estimated belief distribution. It is worth nothing that the belief sample $\hat{\mathbf{H}}_t$ is utilized to compute the prior probability of each node. Specifically, we start with the antenna set $\mathbf{a} = \{1, 2, \dots, N_t\}$ and define $\mathbf{A} = (\hat{\mathbf{H}}_t \hat{\mathbf{H}}_t^H)^{-1}$. Note that we will sometimes omit the time index t since the selection is done within the current frame. Then, we update the antenna set as $\mathbf{a} = \mathbf{a} \setminus \{m\}$ whenever an antenna index m is removed. Similar to greedy search, \mathbf{A} can also be successively updated

Algorithm 4 Utility Functions of RA-MCTS

```

1: function TREEPOLICY( $n$ )
2:   while node  $n$  is not a terminal node do
3:     if node  $n$  is not fully expanded then
4:       return EXPAND( $n$ );
5:     else
6:        $n \leftarrow$  BESTCHILD( $n$ );
7:   return  $n$ 

8: function EXPAND( $n$ )
9:    $\mathbf{r}, \mathbf{A}, \hat{\mathbf{H}} \leftarrow n$ 
10:   $\mathcal{C} = \emptyset$ 
11:  for each untried child node  $i \in \mathcal{N} - \mathbf{r}$  do
12:     $\mathcal{C} = \mathcal{C} \cup \{i\}$ 
13:   $m \leftarrow \arg \max_{i \in \mathcal{C}} \alpha_i$ ;
14:  return RETRIEVENODE( $\mathbf{r} \cup \{m\}, \mathbf{r}, \mathbf{A}, \hat{\mathbf{H}}$ )

15: function BESTCHILD( $n$ )
16:   $\mathbf{r}, \mathbf{A}, \hat{\mathbf{H}} \leftarrow n$ 
17:  for each child node  $i \in \mathcal{N} - \mathbf{r}$  do
18:    Sample uniform replicate  $j \in \{1, 2, \dots, J\}$ ;
19:    Retrieve  $\alpha_{ij}, \beta_{ij}$  according to  $j$ ;
20:   $m \leftarrow \arg \max_i \frac{\alpha_{ij}}{\beta_{ij}}$ ;
21:  return RETRIEVENODE( $\mathbf{r} \cup \{m\}, \mathbf{r}, \mathbf{A}, \hat{\mathbf{H}}$ )

22: function ROLLOUT( $n$ )
23:  while node  $n$  is not a terminal node do
24:     $n \leftarrow$  BESTCHILD( $n$ );
25:   $\mathbf{r}, \mathbf{A}, \hat{\mathbf{H}} \leftarrow n$ 
26:   $\mathbf{a} \leftarrow \mathcal{N} - \mathbf{r}$ 
27:  return  $\mathcal{U}(\mathbf{a})$ 

28: function BACKPROPAGATION( $\delta, n$ )
29:  while node  $n$  is not null do
30:    UPDATEDISTRIBUTION( $\delta, n$ );
31:   $n \leftarrow$  parent of  $n$ ;

32: function UPDATEDISTRIBUTION( $\delta, i$ )
33:  for  $j \in \{1, 2, \dots, J\}$  do
34:    Sample  $\epsilon_j \sim \text{Bernoulli}(\frac{1}{2})$ ;
35:    if  $\epsilon_j = 1$  then
36:       $\alpha_{ij} \leftarrow \alpha_{ij} + \delta$ ;
37:       $\beta_{ij} \leftarrow \beta_{ij} + 1$ ;

```

as $\mathbf{A} = \mathbf{A} + \frac{\hat{\mathbf{A}}_m \hat{\mathbf{h}}_m^H \mathbf{A}}{1 - \hat{\mathbf{h}}_m^H \hat{\mathbf{A}}_m}$. When expanding a new child node i , we compute the corresponding potential spectral loss as $\lambda_i = \frac{\|\mathbf{h}_i^H \mathbf{A}\|^2}{1 - \mathbf{h}_i^H \mathbf{A} \mathbf{h}_i}$. In addition, we have $\alpha_i = \frac{e^{-\lambda_i}}{\sum_{j \neq i} e^{-\lambda_j}}$, which denotes the prior probability of exploring child node i .

By considering each selection as a multi-arm bandit problem, we use the concept of BTS to select child nodes while balancing between exploration and exploitation. The intuition behind BTS is simple and intuitive. The algorithm randomly selects a child node at each step with the probability of being optimal according to current beliefs and, in the meantime, continues to sample all possible child nodes that

could plausibly be optimal [35], [37]. As more information is collected, beliefs about the expected utility of each node are carefully tracked to balance exploration and exploitation [38]. In contrast to the existing Thompson sampling based AS scheme introduced in [15], our proposed approach does not rely on the assumption of equal antenna contribution, as well as the Beta-Bernoulli posterior specification. Indeed, these advantages can help improve the robustness under the circumstance of model misspecification.

Specifically, we adopt a bootstrap distribution to approximate the posterior distribution over the expected outcome of each node. The bootstrap distribution is parameterized by a number of bootstrap replicates, $j \in \{1, \dots, J\}$. In the initialization of a new node i , for each bootstrap replicate, j , we store a set of parameters with $\alpha_{ij} = \alpha_i$ and $\beta_{ij} = 1$ by default, and these parameters will be updated during the backpropagation of each simulation. Specifically, at node i , for each bootstrap replicate j , we update α_{ij} and β_{ij} depending on the result of a coin flip $\text{Bernoulli}(\frac{1}{2})$. If a coin flip is equal to 1, we update the parameters by

$$\alpha_{ij} = \alpha_{ij} + \mathcal{U}(\mathbf{a}); \quad \beta_{ij} = \beta_{ij} + 1, \quad (19)$$

where $\mathcal{U}(\mathbf{a})$ denotes a utility function which evaluates the normalized outcome of a complete antenna selection \mathbf{a} (i.e., $|\mathbf{a}| = N_f$) in the simulation phase. For instance, a utility function in the energy minimization problem is given by

$$\mathcal{U}(\mathbf{a}) = \begin{cases} 1 - \frac{\sum_{k=1}^{N_u} \|\mathbf{w}_k(\mathbf{a})\|^2}{P_{tot}} & \text{if constraints satisfied,} \\ 1, & \text{else,} \end{cases} \quad (20)$$

which normalizes the risk of violating the QoS requirements in the outcome evaluation. This implies that if it cannot satisfy all the constraints, the BS will try its best to allocate all the available power budgets to improve QoS. To decide which antenna should be removed, the previously computed statistics is utilized. For each child node i , we first uniformly sample j from J bootstrap replicates, and then a node with the largest point estimate is selected by $i^* = \arg \max_i \frac{\alpha_{ij}}{\beta_{ij}}$, which allows us to break ties randomly.

B. Complexity Analysis

Intuitively, our proposed RA-MCTS degrades to greedy search (see Algorithm 1) if the computational budget of each iteration is restricted to one single simulation. On the other hand, RA-MCTS is capable of intelligently managing risks and allocating exploration efforts with sufficient computational budget, while the greedy search ignores risks and does not actively explore. This implies that our proposed RA-MCTS achieves a trade-off between performance and computational complexity.

In general, the computational complexity of our proposed RA-MCTS depends on two aspects: the number of simulations and the computational cost of each simulation. For the latter, it can be controlled by the number of bootstrap replicates, as the choice of J obviously limits the number of samples we have from the bootstrap distribution. For a smaller J , RA-MCTS is expected to become greedy. This is because the probability of choosing some nodes that do not have the largest point

estimate will show a trend of being zero. On the other hand, a larger J will involve more exploration, at an expense of extra computation complexity.

To characterize the computational complexity of our proposed algorithms, we consider the total number of floating-point operations (a.k.a. FLOPs) to quantify the complexity order. Note that we consider a FLOP to be either a complex-valued multiplication or a complex-valued summation. In fact, a complex-valued multiplication requires 4 real-valued multiplications and 2 real-valued summations, whereas a complex-valued summation requires only 2 real summations. However, each operation will be counted as one FLOP for simplicity. Given $\mathbf{A} \in \mathbb{C}^{q \times p}$ and $\mathbf{X} \in \mathbb{C}^{p \times r}$, the arithmetic order of FLOPs for the matrix multiplication of $\mathbf{AX} \in \mathbb{C}^{q \times r}$ is of $\mathcal{O}(pqr)$. Given a nonsingular $\mathbf{Y} \in \mathbb{C}^{n \times n}$, the computation complexity for the matrix inversion \mathbf{Y}^{-1} is of $\mathcal{O}(n^3)$.

For the RA-MCTS algorithm, it computes and stores the result of $(\mathbf{HH}^H)^{-1}$ only once, which contains one matrix multiplication and one matrix inversion. The complexity order of $(\mathbf{HH}^H)^{-1}$ is $\mathcal{O}(N_u^2 N_t)$. For each loop of RA-MCTS, despite the fact that it works in a similar manner as greedy search, however, it should be noted that RA-MCTS may explore a path which starts from an intermediate node and has been partially explored before, while greedy search always explore a brand new path starting from the root. This difference implies that for each loop of RA-MCTS, the complexity of greedy search can be regarded as an upper bound in the worst cases, which is given by $\mathcal{O}(N_u^2(N_t^2 - N_f^2))$ [9]. Since RA-MCTS has to backpropagate the simulation results up to $N_t - N_f$ ascendant nodes, the computational complexity for the tree policy and rollout procedures is bounded by $\mathcal{O}(N_t(N_t - N_f)J + N_u^2(N_t^2 - N_f^2))$.

Besides, RA-MCTS also draws one belief state from the belief distribution at each loop. As to the neural network, the normalizing flow is constructed by three kinds of invertible layers, and their computational complexities are determined by element-wise operations and log-determinants [30], [32], [33]. Hence, for a normalizing flow with N_I layers, the computational complexity depends on the input size, which is given by $\mathcal{O}((N_I L N_t N_u))$. For a CNN with L_{Conv} layers, we denote the kernel size and the number of kernels at the i -th layer as $S_{ker}(i)$ and $N_{ker}(i)$, respectively. Then, the computational complexity of CNN is given by $\mathcal{O}(\sum_{i=1}^{L_{Conv}} N_{ker}(i-1)S_{ker}(i)^2 N_t N_u N_{ker}(i))$ [39]. The total computational complexity of drawing belief samples is given by:

$$P_{Net} = \mathcal{O}(N_I L N_t N_u + \sum_{i=1}^{L_{Conv}} N_{ker}(i-1)S_{ker}(i)^2 N_t N_u N_{ker}(i)). \quad (21)$$

Consider RA-MCTS(N, J) with N rollouts and J replicas in total, its computational complexity can be bounded by:

$$\mathcal{O}(NP_{Net} + NN_t(N_t - N_f)J + NN_u^2(N_t^2 - N_f^2)). \quad (22)$$

For comparison, we summarize the computational complexities of some antenna selection algorithms in Table I. It is clear that JCPAS-Basic is a special case of RA-MCTS(N, J) with $N = 1$ and $J = 1$. Moreover, JCPAS-Basic works in the

same way as greedy search, except that the complete CSI is predicted from the belief distribution.

V. NUMERICAL RESULTS

A. Environment Setup

We perform simulations considering the energy minimization problem in (3), and the utility function (20) is adopted to evaluate the system energy efficiency subjected to limited transmit power and QoS constraints. The users are assumed to be randomly located around the BS, and the channels between the BS and users are time-varying. In order to simulate the real propagation environment, we consider temporally correlated fading channels. By the maximum entropy principle, it is common to characterize the channel evolution by the Gaussian-Markov process with a one-step correlation coefficient given by Jakes' model [20]–[22]. Mathematically, the first-order Gaussian-Markov channel model is described by $\mathbf{h}_{k,t} = \zeta_k \mathbf{h}_{k,t-1} + \sqrt{1 - \zeta_k^2} \Delta_t$, where $\zeta_k \in [0, 1]$ denotes the temporal correlation coefficient for user k , and Δ_t is the innovation process with unit-variance complex Gaussian i.i.d. in time. The value of ζ_k is determined by the maximum Doppler frequency and is inversely proportional to the terminal speed [40], in which $\zeta_k = 1$ represents a static channel and $\zeta_k = 0$ implies that the channel is i.i.d. over time. The fading correlation coefficient can be obtained from Jakes' model given by $\zeta_k = J_0\left(2\pi \frac{v_k f_c T}{C}\right)$, where $J_0(\cdot)$ denotes the zeroth order Bessel function of the first kind, v_k is the speed of user k , C is the speed of light, and T is the frame duration [40]. For example, we present the correlation coefficients of some typical scenarios within a wide range of speeds from 3.6 km/h to 290 km/h in Table II. Note that we consider the following two scenarios in our simulations:

- **Scenario I**, low-mobility users: Each user has a uniformly distributed random fading correlation coefficient as $\zeta_k \sim \text{Uni}(0.997, 0.999)$, which represents pedestrians and runners.
- **Scenario II**, high-mobility users: Each user has a uniformly distributed random fading correlation coefficient as $\zeta_k \sim \text{Uni}(0.932, 0.982)$, which represents the vehicular terminals.

In terms of the structure of the neural network, we employ a normalizing flow with $N_i = 16$ flow steps, and each CNN contains 6 layers where the number of convolutional kernels and the kernel size of each layer are $\{64, 32, 32, 16, 64, 128\}$ and $\{3, 9, 3, 3, 3, 9\}$, respectively. Note that the parameters of the neural network are chosen on the basis of experimental experiences. "The DNF model is trained based on the principle of maximum likelihood approximation given in (14) using Pytorch machine learning framework and Adam optimizer." Other common parameters are as follows: $N_f = 8, N_u = 6, P_{tot} = 20$ dBW, $L = 128$.

B. Competing Algorithms

In order to verify the effectiveness of the proposed framework, we compare the proposed algorithms with various competitive algorithms. The proposed algorithms are summarized below:

TABLE I: Computational Complexity Comparisons of Antenna Selection Algorithms with Zero-Forcing Precoding.

Name of Methods	Computational Complexity
Online [15]	$\mathcal{O}(N_t \log N_t + N_f)$
Greedy Search [9]	$\mathcal{O}(N_u^2(N_t^2 - N_f^2))$
JCPAS-Basic	$\mathcal{O}(N_u^2(N_t^2 - N_f^2) + P_{Net})$
RA-MCTS(N, J)	$\mathcal{O}(NN_u^2(N_t^2 - N_f^2) + NP_{Net} + NN_t(N_t - N_f)J)$
Exhaustive Search [9]	$\mathcal{O}(N_u^2 N_t + \frac{N_u^{N_f}}{N_f!}(N_t - N_f) + 2N_u(N_t - N_f)^2 + (N_t - N_f)^3)$

TABLE II: Temporal Correlation Coefficients of Typical Scenarios [17].

Band	Scenarios	Speed (km/h)	Bandwidth (kHz)	Frame duration (ms)	Correlation coefficient
WLAN 802.11@2.4 GHz	Pedestrian user	3.6	15	~ 300	0.999
LTE network@2.6 GHz	Driving in residential area	27.0	15	~ 15.3	0.990
LTE network@2.6 GHz	Driving in residential area	36.0	15	~ 11.5	0.986
LTE network@2.6 GHz	Driving in highway	140.0	15	~ 2.2	0.950
LTE network@2.6 GHz	High speed railway	290.0	15	~ 1.4	0.900

- **JCPAS-Basic:** The proposed JCPAS framework introduced in Section III-A, where the greedy search algorithm is used to select antennas.
- **RA-MCTS(N, J):** The proposed RA-MCTS algorithm introduced in Section IV-A, where N is the maximum number of simulations and J is the number of bootstrap replicates. When compared to JCPAS-Basic, the only difference is that the greedy search algorithm is replaced by RA-MCTS in the JCPAS framework.

We compare our proposed algorithms with three following benchmark schemes:

- **Random:** This scheme randomly selects antennas for data transmission, which is the most naive solution.
- **Online [15]:** The online antenna selection algorithm uses Thompson sampling to select antennas with incomplete CSI.
- **Full CSI:** This scheme uses maximum channel estimation overhead to obtain the complete CSI at each frame and then employs the greedy search [9] for selecting antennas.

For a fair comparison, zero-forcing based precoding is employed in all schemes. We do not compare with the hybrid beamforming (HB) technique since it also requires complete CSI as Full CSI reference and two schemes achieve comparable performance [41]. It is worth noting that the *Full CSI* scheme spends $\tau_{csi} = N_u \lceil \frac{N_t}{N_f} \rceil$ c.u. for CSI estimation, while other schemes use $\tau_{csi} = N_u$ c.u..

C. Performance Comparison and Discussions

Fig. 5 plots the average power consumption of different schemes versus the QoS requirements with $N_t = 32$, $N_f = 8$, $N_u = 6$ and $T = 256$ c.u.. It is observed from Fig. 5 that both the proposed JCPAS-Basic and RA-MCTS algorithms outperform the competing solutions in a wide range of QoS requirements from 5 to 6.6 bps/Hz. Specifically, when $\eta_k = 6.2$ bps/Hz, $\forall k$, the JCPAS-Basic algorithm can reduce the power consumption by about 77% and 79% compared with the “Full CSI” and “Online” schemes, respectively; while the RA-MCTS algorithm can reduce these numbers to approximately 80% and 83%. This is because the RA-MCTS

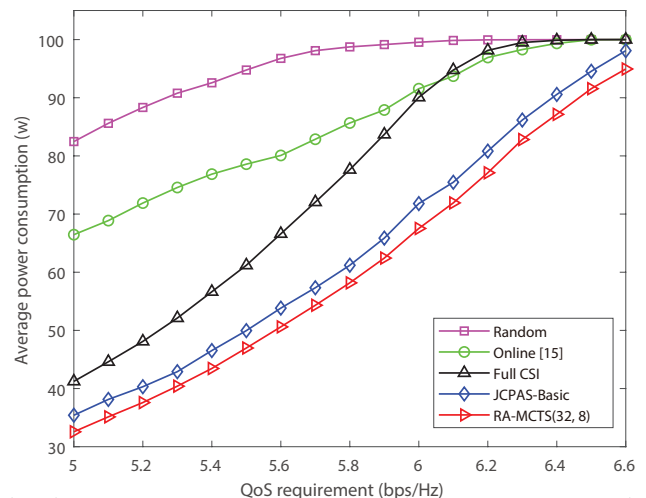


Fig. 5: SCENARIO I, power consumption versus QoS requirements, where $N_t = 32$, $N_f = 8$, $N_u = 6$ and $T = 256$ c.u..

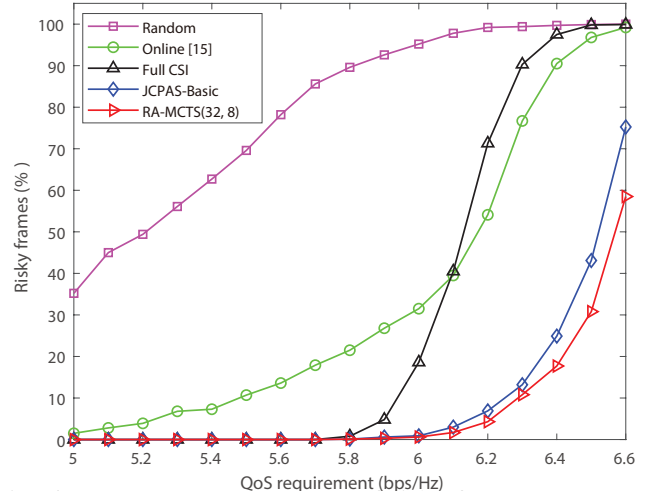


Fig. 6: SCENARIO I, percentage of risky frames versus QoS requirements, where $N_t = 32$, $N_f = 8$, $N_u = 6$ and $T = 256$ c.u..

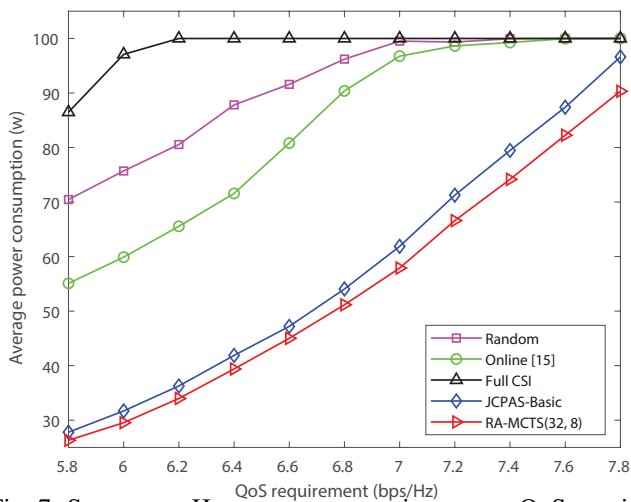


Fig. 7: SCENARIO II, power consumption versus QoS requirements, where $N_t = 64$, $N_f = 8$, $N_u = 4$ and $T = 128$ c.u..

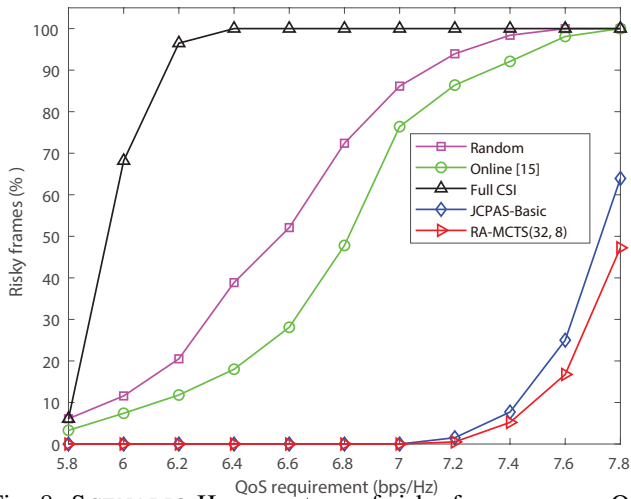


Fig. 8: SCENARIO II, percentage of risky frames versus QoS requirements, where $N_t = 64$, $N_f = 8$, $N_u = 4$ and $T = 128$ c.u..

selects the antennas based on the estimated posterior over the expected outcome of each antennas combination, while the JCPAS-Basic selects antennas only based on a single sample from the estimated belief distribution. This difference makes the proposed RA-MCTS more robust under incomplete CSI measurements.

To demonstrate the robustness of the proposed framework, we compare the proposed algorithms with the reference schemes using a new performance metric of the percentage of risky frames in Fig. 6, where the simulation settings are same as in Fig. 5. A frame is considered as risky if any users' QoS requirement is not satisfied. It is clearly shown that the proposed algorithms, JCPAS-Basic and RA-MCTS, significantly reduce the number of risky frames compared with the references. At the spectral efficiency of 6.1 bps/Hz, all three reference schemes have more than 40% of the frames that are risky, while the proposed JCPAS-Basic and RA-MCTS algorithms have only 4% and 2% risky frames, respectively. One interesting observation is that the "Full CSI" scheme performs well at medium QoS values η_k , but its performance

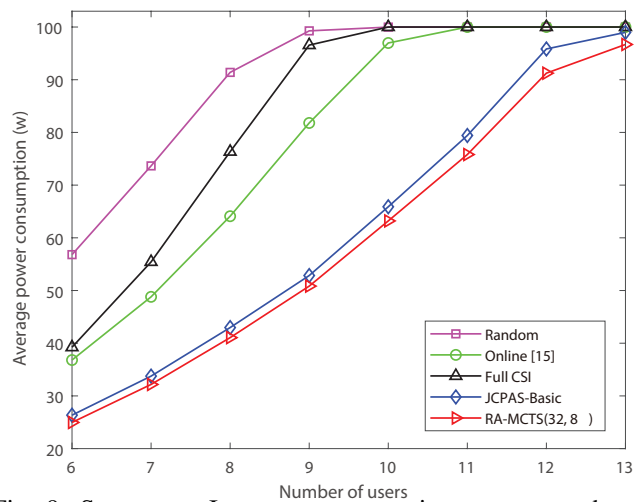


Fig. 9: SCENARIO I, power consumption versus number of users, where $N_t = 128$, $N_f = 16$, $\eta_k = 6.5$ bps/Hz and $T = 512$ c.u..

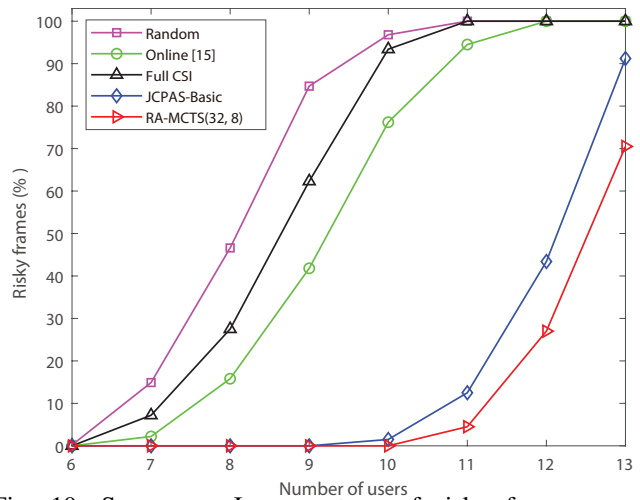


Fig. 10: SCENARIO I, percentage of risky frames versus number of users, where $N_t = 128$, $N_f = 16$, $\eta_k = 6.5$ bps/Hz and $T = 512$ c.u..

quickly drops for higher η_k , to be even worse than the "Online" scheme. This is because the "Full CSI" solution spends a large number of c.u. for channel estimation, hence has the least time for data transmission. As the QoS increases, the "Full CSI" could not satisfy the high QoS requirements in the limited data transmission time even using the maximum transmit power, which results in high number of risky frames. Whereas, the "Online" scheme does not need to estimate the complete channel states. Even when the transmit power of RA-MCTS approaches the power budget (as shown in Fig. 5 for high η_k), the resulting percentage of risky frames is still much lower than other schemes, which demonstrates the robustness of the proposed risk-ware planning framework.

In order to further verify the effectiveness of our proposed algorithms, we present simulation results for SCENARIO II in Figs. 7 and 8 with $N_t = 64$, $N_f = 8$, $N_u = 4$ and $T = 128$ c.u.. Note that we use the same neural network as in Fig. 5. It can be observed from Fig. 8 that in a higher

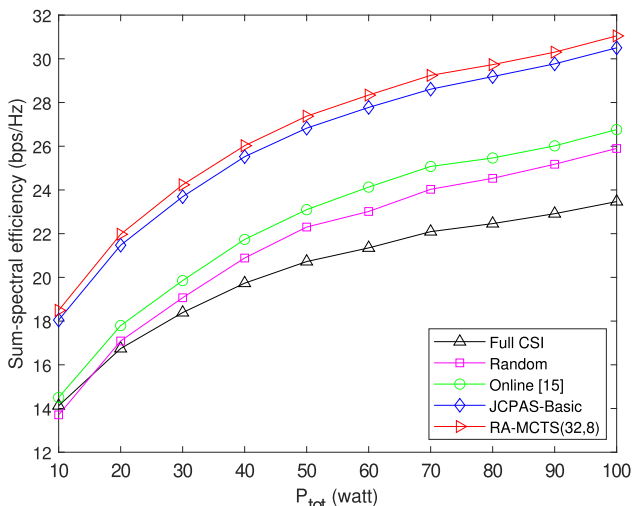


Fig. 11: Sum-spectral efficiency versus power budget, where $N_t = 64$, $N_f = 8$, $N_u = 4$, $\eta_k = 5.5$ bps/Hz and $T = 128$ c.u..

mobility scenario (reduced the frame duration) and with a larger antenna array, the average energy consumption of “Full CSI” scheme is even more than “Random”, which indicates that estimating complete CSI in this case is cost-prohibitive. In general, the superior performance of the proposed algorithms are preserved compared with the references. Specifically, at $\eta_k = 6.4$ bps/Hz, the RA-MCTS scheme can reduce about 60%, 45% and 6% energy consumption compared to the “Full CSI”, “Online” and JCPAS-Basic counterparts. In addition, with a very high QoS requirement of $\eta_k = 7.6$ bps/Hz, the RA-MCTS can reduce 85%, 80% and 8% risky frames compared to “Full CSI”, “Online” and JCPAS-Basic. These results further verify the robustness and effectiveness of our proposed RA-MCTS.

In Figs. 9 and 10, we respectively evaluate the power consumption and the percentage of risky frames as a function of the number of users N_u , where $N_t = 128$, $N_f = 16$, $\eta_k = 6.5$ bpz/HZ and the coherence time $T = 512$ c.u.. In general, serving more users requires more transmit power at all schemes, however, the proposed JCPAS-Basic and RA-MCTS algorithms only consume about 50% of the transmit power of other schemes in most cases, as shown in Figs. 9. The robustness of the proposed framework is clearly shown in Figs. 10, in which the proposed RA-MCTS do not have any risky frame for $N_u \leq 10$, while the percentage of risky frames of three reference schemes grows quickly up to about 80% when N_u varies from 6 to 10. When $N_u = 12$, all three reference schemes have all the frames risky, while the proposed JCPAS-Basic and RA-MCTS algorithms impose a percentage of 45% and 25% of risky frames, respectively. This result confirms the robustness of the proposed RA-MCTS in highly-loaded systems with limited resources.

In Figs. 11-12, we present simulation results regarding the sum-throughput maximization optimization problem, where $N_t = 64$, $N_f = 8$, $N_u = 4$ and the user’s velocity varies from 3.6 km/h to 72 km/h. In addition, the coherent time $T = 128$ c.u. and each user has a QoS requirement of 5.5 bps/Hz. Noted

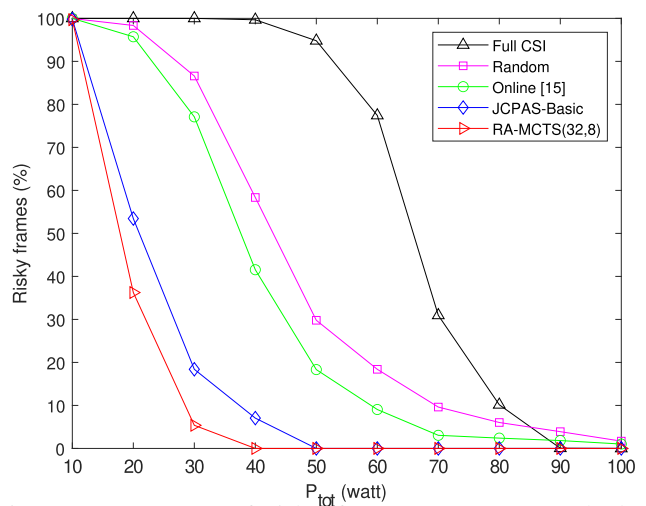


Fig. 12: Percentage of risky frames versus power budget, where $N_t = 64$, $N_f = 8$, $N_u = 4$, $\eta_k = 5.5$ bps/Hz and $T = 128$ c.u..

that if there is no feasible solution to satisfy all the constraints when optimizing the system’s sum-throughput, the frame will be marked as “risky frame” and the QoS constraint will be neglected. In other words, we maximize the sum-throughput by neglecting the QoS constraints for risky frames in the simulations. From the two figures, it can be observed that the proposed algorithms outperform the other three reference schemes in terms of not only the sum-throughput but also the chance of violating user’s QoS requirement. Specifically, when the power budget grows to 30 watt, JCPAS achieves 3.73 bps/Hz more than the “Online” scheme proposed in [15], while the corresponding sum-throughput obtained by RA-MCTS is about 4.29 bps/Hz. Meanwhile, “Online” scheme has 77.10% of risky frames, while JCPAS and RA-MCTS have about only 18.40% and 5.38% risky frames. On the contrary, the proposed JCPAS and RA-MCTS are much more robust to the scenarios of our interests, as the two schemes have much higher sum-throughputs and impose much lower chance to violate the user’s QoS constraint. These results show that the proposed algorithms are robust for both the power minimization problem and the sum-throughput maximization problem.

VI. CONCLUSION

In this paper, we have proposed a robust joint channel prediction and antenna selection framework JCPAS for massive MIMO systems under practical incomplete CSI, which results from the limited number of RF chains, and limited transmit power and QoS requirements. In order to address this problem, we first proposed a deep neural network to estimate the posterior belief distribution of the current channel states. After that, we developed a joint channel tracking and antenna selection algorithm to select the best antennas based on the estimated belief. To improve system robustness, we proposed a risk-aware planning framework, namely RA-MCTS, which employs Monte Carlo tree search and bootstrap Thompson sampling to approximate a posterior distribution over the random objectives. Simulation results showed that the proposed RA-MCTS not only achieves a lower energy

consumption but also significantly reduces the risk, quantified as the probability of violating the system constraints.

For the future work, one interesting topic is to further improve the computational efficiency of the proposed RA-MCTS. A feasible solution is to directly estimate the rollout results through a DNN such that the rollout overhead can be significantly reduced. In this case, the posterior distribution over the rollout results is also estimated by DNN, rather than the bootstrap distribution parameterized by the J replicates. The developed framework can be easily applied to cell-free MIMO systems to optimize the number of active access points under limited system resources. Another promising topic is to consider the confidence of the predicted channel coefficient in the selection process. In this case, the best antennas subset should be selected based on not only the channel gain but also the prediction confidence.

REFERENCES

- [1] K. He, T. X. Vu, S. Chatzinotas, and B. Ottersten, "Learning-based joint channel prediction and antenna selection for massive MIMO with partial CSI," in *IEEE GLOBECOM Workshops 2022*, pp. 178–183.
- [2] L. Lu, G. Y. Li, A. L. Swindlehurst, A. Ashikhmin, and R. Zhang, "An overview of massive MIMO: Benefits and challenges," *IEEE J. Sel. Top. Signal Process.*, vol. 8, no. 5, pp. 742–758, 2014.
- [3] J. Wang, X. Zhang, X. Shi, and J. Song, "Higher spectral efficiency for mmwave MIMO: Enabling techniques and precoder designs," *IEEE Commun. Mag.*, vol. 59, no. 4, pp. 116–122, 2021.
- [4] W.-B. Sun, W.-X. Meng, J.-C. Guo, and C. Li, "Multiuser MIMO opportunistic beamforming communications: State-of-the-Art and perspectives," *IEEE Wirel. Commun.*, vol. 29, no. 3, pp. 95–101, 2022.
- [5] A. F. Molisch, V. V. Ratnam, S. Han, Z. Li, S. L. H. Nguyen, L. Li, and K. Haneda, "Hybrid beamforming for massive MIMO: A survey," *IEEE Commun. Mag.*, vol. 55, no. 9, pp. 134–141, 2017.
- [6] Y. Gao, H. Vinck, and T. Kaiser, "Massive MIMO antenna selection: Switching architectures, capacity bounds, and optimal antenna selection algorithms," *IEEE Trans. Sig. Proc.*, vol. 66, no. 5, pp. 1346–1360, 2017.
- [7] A. Gorokhov, "Antenna selection algorithms for MEA transmission systems," in *Proc. ICASSP 2002*, vol. 3, pp. III–2857.
- [8] M. Gharavi-Alkhanjari and A. B. Gershman, "Fast antenna subset selection in MIMO systems," *IEEE Trans. Sig. Proc.*, vol. 52, no. 2, pp. 339–347, 2004.
- [9] P.-H. Lin and S.-H. Tsai, "Performance analysis and algorithm designs for transmit antenna selection in linearly precoded multiuser MIMO systems," *IEEE Trans. Veh. Technol.*, vol. 61, no. 4, pp. 1698–1708, 2012.
- [10] J. Chen, S. Chen, Y. Qi, and S. Fu, "Intelligent massive MIMO antenna selection using monte carlo tree search," *IEEE Trans. Sig. Proc.*, vol. 67, no. 20, pp. 5380–5390, 2019.
- [11] S. Zhang, S. Zhang, F. Gao, J. Ma, and O. A. Dobre, "Deep learning optimized sparse antenna activation for reconfigurable intelligent surface assisted communication," *IEEE Trans. Commun.*, vol. 69, no. 10, pp. 6691–6705, 2021.
- [12] T. X. Vu, S. Chatzinotas, V.-D. Nguyen, D. T. Hoang, D. N. Nguyen, M. Di Renzo, and B. Ottersten, "Machine learning-enabled joint antenna selection and precoding design: From offline complexity to online performance," *IEEE Trans. Wirel. Commun.*, vol. 20, no. 6, pp. 3710–3722, 2021.
- [13] W. Yu, T. Wang, and S. Wang, "Multi-label learning based antenna selection in massive MIMO systems," *IEEE Trans. Veh. Technol.*, vol. 70, no. 7, pp. 7255–7260, 2021.
- [14] Z. Liu, Y. Yang, F. Gao, T. Zhou, and H. Ma, "Deep unsupervised learning for joint antenna selection and hybrid beamforming," *IEEE Trans. Commun.*, vol. 70, no. 3, pp. 1697–1710, 2022.
- [15] Z. Kuai and S. Wang, "Thompson sampling-based antenna selection with partial CSI for tdd massive MIMO systems," *IEEE Trans. Commun.*, vol. 68, no. 12, pp. 7533–7546, 2020.
- [16] O. Raaeesi, A. Gokceoglu, Y. Zou, E. Björnson, and M. Valkama, "Performance analysis of multi-user massive MIMO downlink under channel non-reciprocity and imperfect CSI," *IEEE Trans. Commun.*, vol. 66, no. 6, pp. 2456–2471, 2018.
- [17] S. Sharifi, S. Shahbazpanahi, and M. Dong, "A POMDP-based antenna selection for massive MIMO communication," *IEEE Trans. Commun.*, vol. 70, no. 3, pp. 2025–2041, 2021.
- [18] X. Gao, O. Edfors, F. Tufvesson, and E. G. Larsson, "Massive MIMO in real propagation environments: Do all antennas contribute equally?" *IEEE Trans. Commun.*, vol. 63, no. 11, pp. 3917–3928, 2015.
- [19] J. Vieira, S. Malkowsky, K. Nieman, Z. Miers, N. Kundargi, L. Liu, I. Wong, V. Öwall, O. Edfors, and F. Tufvesson, "A flexible 100-antenna testbed for massive MIMO," in *IEEE GLOBECOM Workshops 2014*, pp. 287–293.
- [20] H. S. Wang and P.-C. Chang, "On verifying the first-order Markovian assumption for a Rayleigh fading channel model," *IEEE Trans. Veh. Technol.*, vol. 45, no. 2, pp. 353–357, 1996.
- [21] C. C. Tan and N. C. Beaulieu, "On first-order Markov modeling for the rayleigh fading channel," *IEEE Trans. Commun.*, vol. 48, no. 12, pp. 2032–2040, 2000.
- [22] G. J. Byers and F. Takawira, "Spatially and temporally correlated MIMO channels: Modeling and capacity analysis," *IEEE Trans. Veh. Technol.*, vol. 53, no. 3, pp. 634–643, 2004.
- [23] C. Wu, X. Yi, Y. Zhu, W. Wang, L. You, and X. Gao, "Channel prediction in high-mobility massive MIMO: From spatio-temporal autoregression to deep learning," *IEEE J. Sel. Areas Commun.*, vol. 39, no. 7, pp. 1915–1930, 2021.
- [24] T. Zhou, H. Zhang, B. Ai, C. Xue, and L. Liu, "Deep-learning based spatial-temporal channel prediction for smart high-speed railway communication networks," *IEEE Trans. Wirel. Commun.*, 2022.
- [25] G. Liu, Z. Hu, L. Wang, J. Xue, H. Yin, and D. Gesbert, "Spatio-temporal neural network for channel prediction in massive MIMO-OFDM systems," *IEEE Trans. Commun.*, vol. 70, no. 12, pp. 8003–8016, 2022.
- [26] T. K. Vu, M. Bennis, M. Debbah, M. Latva-aho, and C. S. Hong, "Ultra-reliable communication in 5G mmwave networks: A risk-sensitive approach," *IEEE Commun. Lett.*, vol. 22, no. 4, pp. 708–711, 2018.
- [27] N. B. Khalifa, M. Assaad, and M. Debbah, "Risk-sensitive reinforcement learning for urllc traffic in wireless networks," in *Proc. IEEE WCNC 2019*, pp. 1–7.
- [28] M. Alsenwi, N. H. Tran, M. Bennis, S. R. Pandey, A. K. Bairagi, and C. S. Hong, "Intelligent resource slicing for embb and URLLC coexistence in 5G and beyond: A deep reinforcement learning based approach," *IEEE Trans. Wirel. Commun.*, vol. 20, no. 7, pp. 4585–4600, 2021.
- [29] L.-C. Lan, T.-R. W. Wu, I.-C. W. Wu, and C.-J. Hsieh, "Learning to stop: Dynamic simulation monte-carlo tree search," in *AAAI Conference on Artificial Intelligence*, vol. 35, no. 1, 2021, pp. 259–267.
- [30] D. P. Kingma and P. Dhariwal, "Glow: Generative flow with invertible 1x1 convolutions," in *Proc. NeurIPS 2018*, vol. 31.
- [31] A. Abdelhamed, M. A. Brubaker, and M. S. Brown, "Noise flow: Noise modeling with conditional normalizing flows," in *Proc. ICCV 2019*, pp. 3165–3173.
- [32] L. Dinh, D. Krueger, and Y. Bengio, "NICE: Non-linear independent components estimation," *arXiv preprint arXiv:1410.8516*, 2014.
- [33] K. He, L. He, L. Fan, Y. Deng, G. K. Karagiannidis, and A. Nallanathan, "Learning-based signal detection for MIMO systems with unknown noise statistics," *IEEE Trans. Commun.*, vol. 69, no. 5, pp. 3025–3038, 2021.
- [34] C. B. Browne, E. Powley, D. Whitehouse, S. M. Lucas, P. I. Cowling, P. Rohlfshagen, S. Tavener, D. Perez, S. Samothrakis, and S. Colton, "A survey of Monte Carlo tree search methods," *IEEE Trans. Comput. Intell. AI Games*, vol. 4, no. 1, pp. 1–43, 2012.
- [35] D. J. Russo, B. Van Roy, A. Kazerouni, I. Osband, Z. Wen *et al.*, "A tutorial on Thompson sampling," *Found. Trends Mach. Learn.*, vol. 11, no. 1, pp. 1–96, 2018.
- [36] M. Phan, Y. Abbasi Yadkori, and J. Domke, "Thompson sampling and approximate inference," in *Proc. NeurIPS 2019*, vol. 32.
- [37] D. Eckles and M. Kaptein, "Thompson sampling with the online bootstrap," *arXiv preprint arXiv:1410.4009*, 2014.
- [38] C. F. Hayes, M. Reymond, D. M. Roijers, E. Howley, and P. Mannon, "Distributional monte carlo tree search for risk-aware and multi-objective reinforcement learning," in *Proc. AAMAS 2021*, pp. 1530–1532.
- [39] K. He and J. Sun, "Convolutional neural networks at constrained time cost," in *Proc. CVPR*, 2015, pp. 5353–5360.
- [40] G. Caire, N. Jindal, M. Kobayashi, and N. Ravindran, "Multiuser MIMO achievable rates with downlink training and channel state feedback," *IEEE Trans. Inf. Theory*, vol. 56, no. 6, pp. 2845–2866, 2010.

- [41] E. Zöchmann, S. Schwarz, and M. Rupp, "Comparing antenna selection and hybrid precoding for millimeter wave wireless communications," in *IEEE Sensor Array and Multichannel Signal Processing Workshop (SAM)*, 2016, pp. 1–5.



Ke He (Student Member, IEEE) received his bachelor's degree from the Wuhan University of Technology in 2015 and the master's degree in computer science from Guangzhou University in 2021. He is currently pursuing his PhD degree with the Interdisciplinary Centre for Security, Reliability and Trust (SnT), University of Luxembourg. His current research interests include signal processing, resources management and statistical machine learning.



Thang X. Vu (Senior Member, IEEE) received the B.S. and the M.Sc., both in Electronics and Telecommunications Engineering, from the VNU University of Engineering and Technology, Vietnam, in 2007 and 2009, respectively, and the Ph.D. in Electrical Engineering from the University Paris-Sud, France, in 2014. In 2010, he received the Allocation de Recherche fellowship to study Ph.D. in France. From July 2014 to January 2016, he was a postdoctoral researcher with the Singapore University of Technology and Design (SUTD), Singapore. Currently, he is

a research scientist at the Interdisciplinary Centre for Security, Reliability and Trust (SnT), University of Luxembourg. His research interest includes wireless communications, with particular interests of applications of optimization and machine learning on design and analyze the multi-layer 6G networks. He has successfully acquired, as the PI and vice PI, several Luxembourg national and ESA projects with a total funding of 2.6 MEURs. He was a recipient of the SigTelCom 2019 Best Paper Award. He is also serving as an Associate Editor for the IEEE Communications Letters.



Dinh Thai Hoang (M'16, SM'22) is currently a faculty member at the School of Electrical and Data Engineering, University of Technology Sydney, Australia. He received his Ph.D. in Computer Science and Engineering from the Nanyang Technological University, Singapore 2016. His research interests include emerging wireless communications and networking topics, especially machine learning applications in networking, edge computing, and cybersecurity. He has received several precious awards, including the Australian Research Council

Discovery Early Career Researcher Award, IEEE TCSC Award for Excellence in Scalable Computing for Contributions on "Intelligent Mobile Edge Computing Systems" (Early Career Researcher), IEEE Asia-Pacific Board (APB) Outstanding Paper Award 2022, and IEEE Communications Society Best Survey Paper Award 2023. He is currently an Editor of IEEE TMC, IEEE TWC, IEEE TCCN, IEEE TVT, and IEEE COMST.



Diep N. Nguyen (Senior Member, IEEE) received the M.E. degree in electrical and computer engineering from the University of California at San Diego (UCSD), La Jolla, CA, USA, in 2008, and the Ph.D. degree in electrical and computer engineering from The University of Arizona (UA), Tucson, AZ, USA, in 2013. He is currently a faculty member with the Faculty of Engineering and Information Technology, University of Technology Sydney (UTS), Sydney, NSW, Australia. Before joining UTS, he was a DECRA Research Fellow with Macquarie University, Macquarie Park, NSW, Australia, and a Member of the Technical Staff with Broadcom Corporation, San Jose, CA, USA, and ARCON Corporation, Boston, MA, USA, and consulting the Federal Administration of Aviation, Washington, DC, USA, on turning detection of UAVs and aircraft, and the U.S. Air Force Research Laboratory, USA, on anti-jamming. His research interests include computer networking, wireless communications, and machine learning application, with emphasis on systems' performance and security/privacy. Dr. Nguyen received several awards from LG Electronics, UCSD, UA, the U.S. National Science Foundation, and the Australian Research Council. He has served on the Editorial Boards of the IEEE Transactions on Mobile Computing, IEEE Communications Surveys & Tutorials (COMST), IEEE Open Journal of the Communications Society, and Scientific Reports (Nature's).



Symeon Chatzinotas (Fellow, IEEE) received the M.Eng. degree in telecommunications from Aristotle University of Thessaloniki, Greece, in 2003, and the M.Sc. and Ph.D. degrees in electronic engineering from the University of Surrey, U.K., in 2006 and 2009, respectively. He is currently a Full Professor/Chief Scientist I and the Head of the Research Group SIGCOM, Interdisciplinary Centre for Security, Reliability and Trust, University of Luxembourg. In the past, he has lectured as a Visiting Professor with the University of Parma, Italy, and

contributed in numerous research and development projects for the Institute of Informatics and Telecommunications, National Center for Scientific Research "Demokritos," the Institute of Telematics and Informatics, Center of Research and Technology Hellas, and Mobile Communications Research Group, Center of Communication Systems Research, University of Surrey. He has authored more than 700 technical papers in refereed international journals, conferences, and scientific books. Prof. Chatzinotas received numerous awards and recognitions, including the IEEE Fellowship and an IEEE Distinguished Contributions Award. He is currently in the editorial board of the IEEE TRANSACTIONS ON COMMUNICATIONS, IEEE OPEN JOURNAL OF VEHICULAR TECHNOLOGY, and the International Journal of Satellite Communications and Networking.



Bjorn Ottersten (S'87–M'89–SM'99–F'04) received the M.S. degree in electrical engineering and applied physics from Linköping University, Linköping, Sweden, in 1986, and the Ph.D. degree in electrical engineering from Stanford University, Stanford, CA, USA, in 1990. He has held research positions with the Department of Electrical Engineering, Linköping University, the Information Systems Laboratory, Stanford University, the Katholieke Universiteit Leuven, Leuven, Belgium, and the University of Luxembourg, Luxembourg. From 1996 to

1997, he was the Director of Research with ArrayComm, Inc., a start-up in San Jose, CA, USA, based on his patented technology. In 1991, he was appointed Professor of signal processing with the Royal Institute of Technology (KTH), Stockholm, Sweden. Dr. Ottersten has been Head of the Department for Signals, Sensors, and Systems, KTH, and Dean of the School of Electrical Engineering, KTH. He is currently the Director for the Interdisciplinary Centre for Security, Reliability and Trust, University of Luxembourg.

He is a recipient of the IEEE Signal Processing Society Technical Achievement Award, the EURASIP Group Technical Achievement Award, and the European Research Council advanced research grant twice. He has co-authored journal papers that received the IEEE Signal Processing Society Best Paper Award in 1993, 2001, 2006, 2013, and 2019, and 8 IEEE conference papers best paper awards. He has been a board member of IEEE Signal Processing Society, the Swedish Research Council and currently serves of the boards of EURASIP and the Swedish Foundation for Strategic Research. Dr. Ottersten has served as Editor in Chief of EURASIP Signal Processing, and acted on the editorial boards of IEEE Transactions on Signal Processing, IEEE Signal Processing Magazine, IEEE Open Journal for Signal Processing, EURASIP Journal of Advances in Signal Processing and Foundations and Trends in Signal Processing. He is a fellow of EURASIP. For more information, see <https://creativecommons.org/licenses/by/4.0/>