Transmit Antenna Selection in Massive MIMO Systems: An Online Learning Framework

Zhenran Kuai*, Tianyu Wang*†, and Shaowei Wang*
* School of Electronic Science and Engineering, Nanjing University, Nanjing 210023, China
† National Mobile Communications Research Laboratory, Southeast University, Nanjing 210096, China
Email: 151180058@smail.nju.edu.cn, {tianyu.alex.wang, wangsw}@nju.edu.cn

Abstract—**Antenna selection (AS) is a signal processing technique that activates a selected subset of available antennas in multi-antenna systems, based on which a performance-hardware tradeoff can be achieved by reducing the number of costly radio-frequency (RF) chains. The biggest challenge of AS is the combinatorial complexity that arises from the classic K-out-of-N problem, which makes it more challenging for massive MIMO systems equipped with large-scale antenna arrays. In addition, for massive MIMO systems with limited RF chains, the amount of radio resources dedicated to channel state information (CSI) measurement will increase tremendously, which may highly degrade the overall performance of AS. In this paper, we consider the transmit AS problem in time division duplexing (TDD) massive MIMO systems, where K out of N transmit antennas are selected to maximize the total throughput of M single-antenna users in the downlink. We propose an online learning scheme and introduce Thompson sampling techniques to update the set of active antennas with partial CSI. The idea behind is to find an efficient tradeoff between the exploitation of high-performance antennas and the exploration of antennas with uncertain CSI with low complexity. Our proposed scheme is validated by using COST 2100 channel model, and simulation results show that it greatly outperforms the conventional power-based and convex relaxation based schemes, in terms of the total downlink throughput.**

Index Terms—**Antenna selection, massive MIMO, online learning, Thompson sampling.**

I. Introduction

Massive MIMO has been recognized as one of the key technologies in the upcoming 5G networks, which serves multiple single-antenna users in the same time-frequency block by equipping base stations with a large-scale antenna array. Many studies have shown that it can greatly improve the performance in spectral and energy efficiency by exploiting the spacial diversity brought by massive number of transmit antennas. However, the amount of corresponding radio-frequency (RF) hardware (e.g., low noise amplifiers, frequency up/down converters, and analog-to-digital/digital-to-analog converters) that increases linearly with the number of antennas can become a challenging issue for both implementation complexity and financial cost [1]. In fact, for millimeter-wave (mm-Wave) massive MIMO systems, a fully digital signal processing structure is physically impractical due to the tight space between mm-Wave antennas [2]. Therefore, signal processing techniques using reduced number of RF chains have gained great attentions in the massive MIMO literature.

Antenna selection (AS) is a signal processing technique that activates a selected subset of available antennas in multi-antenna systems, which achieves a performance-hardware tradeoff by reducing the number of required RF chains. It has been been widely studied in conventional MIMO systems, which shows it significantly improves the performance in terms of diversity, capacity and energy efficiency, compared with systems using the same number of antennas and RF chains [3]–[5]. In [3], the authors studied the receive AS problem in MIMO for spatial multiplexing from a theoretic point of view, and proposed two greedy AS algorithms with low computational complexity. In [4], the authors studied the transmit AS problem in single-stream MIMO systems, which shows that it can be a competitive technology compared to transmit beamforming in terms of the energy efficiency performance. In [5], the authors proposed an AS scheme under the multi-armed bandit framework to maximize the system energy efficiency. However, this work considers a static scenario where the channel parameters are time-invariant, which makes it completely different from our work that considers a dynamic setting of wireless channels from an online point of view. And their proposed algorithm based on upper confidence bound is different from our proposed online algorithm based on Thompson sampling.

AS is then naturally introduced in massive MIMO, which suffers more in the issue of RF hardware implementation. The challenge comes from two aspects. The first is the combinatorial complexity that arises from the classic K-out-of-N problem, which increases exponentially with the number of both RF chains and available antennas. The second is the amount of radio resources dedicated to channel state information (CSI) measurement, which grows linearly with the number of simultaneous users and available antennas, for uplink and downlink pilots, respectively. This challenge is aggravated by the use of AS, as it may require multiple time of pilot transmissions to achieve full CSI due to the reduced number of RF chains. Most studies in the literatures have focused on the first challenge, for which a variety of low-complexity AS algorithms have been proposed. In [6], the authors proposed a low-complexity algorithm based on sorting and shifting

This work was partially supported by the National Natural Science Foundation of China (61671233, 61801208), the Jiangsu Science Foundation (BK20170650), the Postdoctoral Science Foundation of China (BX201700118, 2017M621712), and the Jiangsu Postdoctoral Science Foundation(1701118B), and the open research fund of National Mobile Communications Research Laboratory (2019D02).

techniques for transmit AS in massive MIMO, which shows great performance in spectral efficiency. In [7], the authors transformed the transmit AS problem in massive MIMO into a convex optimization problem by using relaxation techniques, and proposed two AS algorithms based on convex relaxation and power distribution. In [8], the authors formulated the AS problem with fixed power allocation at receivers as a maximum entropy sampling problem by leveraging its sub-modularity, and proposed a greedy AS algorithm with low computational complexity.

We note that the extra pilot overhead brought by AS may significantly degrade the performance of massive MIMO systems. In this paper, we consider the transmit antenna selection problem in time division duplexing (TDD) massive MIMO systems, in which K out of N transmit antennas are selected to maximize the total throughput of M single-antenna users in the downlink. In order to reduce pilot overhead, we introduce an online paradigm using partial CSI measurement, in which the AS strategy is updated by using the partial CSI measured in previous time slots. Specifically, we formulate the online AS scheme as a combinatorial multi-armed bandit (CMAB) problem, for which we propose a low-complexity algorithm by using Thompson sampling techniques. The idea behind online AS is to find an efficient tradeoff between the exploitation of high-performance antennas and the exploration of antennas with uncertain CSI by using low-complexity techniques. And its effectiveness is guaranteed by the temporal correlation of wireless channels. We validate our proposed online AS scheme by using the COST 2100 channel model, and simulation results show that it greatly outperforms the conventional power-based and convex relaxation based schemes, in terms of the total downlink throughput.

The rest of the paper is organized as follows: In Section II, we give the system model. In Section III, we show how it can be formulated as a CMAB problem. In Section IV, we propose the online AS scheme using Thompson sampling. In Section V, the proposed scheme is validated by using the COST 2100 channel model, and compared with conventional schemes. And in Section VI, we conclude our work.

II. System Model

We consider the downlink transmission of a TDD multi-user massive MIMO system with one base station and M mobile users. As shown in Fig. 1, the base station is equipped with N antennas and K RF chains. The number of antennas is much greater than the number of mobile users, i.e., $N \gg M$, and the number of RF chains is smaller than the number of antennas, i.e., $K < N$. We denote $\mathcal{N} = \{1, 2, ...N\}$ as the set of all available antennas at the base station. For each transmission slot, an antenna subset $\mathcal{K} = \{a_1, a_2, ...a_K\}$ with exact K antennas, are selected and connected to K RF chains, respectively, through the RF switch shown in Fig. 1. We denote $\mathcal{S}(\mathcal{N})$ as the set of all feasible antenna sets, i.e., $\mathcal{S}(\mathcal{N}) = \{s \subset \mathcal{N} \mid \ |s| = K\}$.

For each subcarrier l, we denote $\boldsymbol{x}_l \in \mathbb{C}^K$ as the transmit signal vector of K antennas at the base station, and $\boldsymbol{y}_l \in \mathbb{C}^M$ as the received signal vector of M users. The channel matrix between K antennas and M users is denoted by $\boldsymbol{H}_l^{\mathcal{K}} \in \mathbb{C}^{M\times K}$, which is normalized such that each element $h_{m,k}^{\mathcal{K}}$ has unit energy in average. The superscript $\mathcal{K}$ indicates the set of selected antennas. Thus, we have

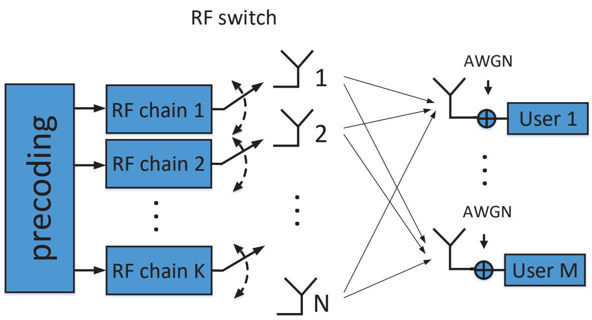


Fig. 1. Transmit antenna selection in massive MIMO systems.

$$\boldsymbol{y}_l = \sqrt{M\rho}\boldsymbol{H}_l^{\mathcal{K}}\boldsymbol{x}_l + \boldsymbol{n}_l, \tag{1}$$

where ρ represents the normalized transmit signal-to-noise ratio (SNR) per user and $\boldsymbol{n}_l \in \mathbb{C}^K$ is the complex Gaussian noise with zero mean and unit variance. Note that the total transmit power $M\rho$ is independent of the number of selected antennas K, while the average received SNR per-user $K\rho$ increases as the number of selected antennas increases.

We denote $\boldsymbol{W}^{\mathcal{K}} \in \mathbb{C}^{K\times M}$ as the precoding matrix of the base station, and $\boldsymbol{s} \in \mathbb{C}^M$ as the transmit signal vector for M users. Thus, we have

$$\boldsymbol{x}_l = \boldsymbol{W}^{\mathcal{K}}\boldsymbol{s}, \tag{2}$$

where $\boldsymbol{s}$ is normalized as $\mathbb{E}\{\|\boldsymbol{s}\|^2\} = 1$ and $\boldsymbol{W}^{\mathcal{K}}$ is normalized as $\mathbb{E}\{\|\boldsymbol{x}_l\|^2\} = 1$. $\|\cdot\|$ represents 2-norm of a vector. In this paper, we consider the zero-forcing precoding matrix [9]

$$\boldsymbol{W}^{\mathcal{K}} = \left(\boldsymbol{H}_l^{\mathcal{K}}\right)^H \left(\boldsymbol{H}_l^{\mathcal{K}}\left(\boldsymbol{H}_l^{\mathcal{K}}\right)^H\right)^{-1}, \tag{3}$$

where $(\cdot)^H$ represents the conjugate transpose of a matrix.

Therefore, the received signal-to-interference-plus-noise ratio of user m is given by

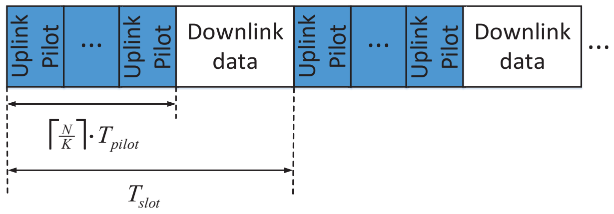
$$SINR_m(\mathcal{K}) = \frac{M\rho\left|\boldsymbol{h}_m^{\mathcal{K}}\boldsymbol{w}_m^{\mathcal{K}}\right|^2}{M\rho\sum_{m'\neq m}^{M}\left|\boldsymbol{h}_m^{\mathcal{K}}\boldsymbol{w}_{m'}^{\mathcal{K}}\right|^2 + 1}, \tag{4}$$

where $\boldsymbol{h}_m^{\mathcal{K}}$ is the m-th row of the channel matrix $\boldsymbol{H}_l^{\mathcal{K}}$, and $\boldsymbol{w}_m^{\mathcal{K}}$ is the m-th column of the precoding matrix $\boldsymbol{W}^{\mathcal{K}}$. And the system capacity with AS $\mathcal{K} \in \mathcal{S}(\mathcal{N})$ is then given by

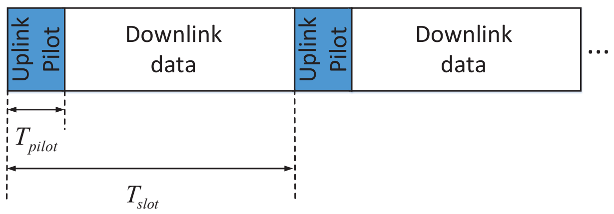
$$C(\mathcal{K}) = \sum_{m=1}^{M}\log_2\left(1 + SINR_m(\mathcal{K})\right). \tag{5}$$

III. Antenna Selection as Combinatorial Multi-Armed Bandit Problem

In TDD massive MIMO systems, according to the reciprocity characteristics of wireless channels, the downlink CSI



(a) Full CSI



(b) Partial CSI

Fig. 2. Slot structure of full and partial CSI AS in TDD massive MIMO.

is acquired with the received uplink pilots. The overhead of uplink pilots is determined by the number of simultaneous users M. However, when the RF chains are less than the available antennas, i.e., $K < N$, the base station requires $\lceil N/K \rceil$ times channel estimation to achieve full CSI of all antennas, which greatly increases the pilot overhead as shown in Fig. 2(a). Thus, the actual throughput of AS $\mathcal{K} \in \mathcal{S}(\mathcal{N})$ with full CSI is given by

$$R_f(\mathcal{K}) = \frac{T_{slot} - \left\lceil \frac{N}{K} \right\rceil T_{pilot}}{T_{slot}} C(\mathcal{K}), \tag{6}$$

where T_{slot} is the slot length and T_{pilot} is the length of a single pilot sequence in the uplink.

To reduce the pilot overhead, we consider an online framework in which only partial CSI is updated by using single channel estimation in each transmission slot, as shown in Fig. 2(b). The corresponding online strategy needs to decide the current AS $\mathcal{K}(t) \in \mathcal{S}(\mathcal{N})$ based on partially updated CSI. Thus, the average throughput of an online AS sequence $\mathcal{K}(1)$, $\mathcal{K}(2)$, $\cdots$ is given by

$$R_p\left(\{K(t)\}_{t\geq 1}\right) = \frac{T_{slot} - T_{pilot}}{T_{slot}} \lim_{T\to\infty} \frac{1}{T} \sum_{t=1}^{T} C(\mathcal{K}(t)). \tag{7}$$

To maximize $R_p\left(\{K(t)\}_{t\geq 1}\right)$, the online strategy needs to decide the antenna set $\mathcal{K}(t) \in \mathcal{S}(\mathcal{N})$ that not only helps to maximize the current system capacity $C(\mathcal{K})$, but also helps the AS decision in the next transmission slot by updating CSI of uncertain channels. This online problem can be formulated as a CMAB problem, in which the expected reward in the long run is maximized by making a combinatorial choice in each round, when the properties of each choice are only partially revealed from each of its previous rewards [10], [11].

Specifically, in each time slot t, the reward of choice $\mathcal{K}(t) \in \mathcal{S}(\mathcal{N})$ is the corresponding system capacity $C(\mathcal{K}(t))$. We denote $\pi = \{\pi(t)\}_{t\geq 1}$ as an online strategy of the considered CMAB problem, in which each element $\pi(t) : \mathbb{R}^{t-1} \to \mathcal{S}(\mathcal{N})$ determines the AS $\mathcal{K}(t)$ at time slot t by using the previous rewards $C(\mathcal{K}(1))$, $C(\mathcal{K}(2))$,... $C(\mathcal{K}(t-1))$. Therefore, the optimal online strategy is given by

$$\pi^* = \arg\max_{\pi} \lim_{T\to\infty} \frac{1}{T} \sum_{t=1}^{T} C\left(\mathcal{K}^{\pi(t)}(t)\right). \tag{8}$$

We note that the optimal solution is a combinatorial optimization problem which is NP-hard in general. Also, due to the dynamic nature of wireless channels, the capacity function $C : \mathcal{S}(\mathcal{N}) \to \mathbb{R}$ is also time-varing, which makes it an even harder dynamic combinatorial optimization problem.

IV. Online Antenna Selection

We develop an online AS algorithm by using Thompson sampling techniques. The key idea is to balance the benefit of exploiting high-performance antennas, and the cost of exploring antennas with uncertain performance. Also, we introduce discount factors to track the time-varying CSI of wireless channels.

A. Thompson Sampling

Thompson sampling is based on Bayesian posterior credible intervals. It targets on choosing the arm with the maximum reward by using sampling techniques in multi-armed bandit problems. The key idea is to approximate an unknown variable by using of a general distribution with tunable parameters which are iteratively updated according to the sampling results in each round [12].

Assume an antenna-specific reward exists, which is denoted by $r_n(t)$ for each antenna n at each slot t, and the total reward $R(\mathcal{K}(t))$ is the sum of the corresponding antenna-specific rewards, i.e., $R(\mathcal{K}(t)) = \sum_{n\in\mathcal{K}(t)} r_n(t)$. The antenna-specific reward is normalized by an interval $[r_{min},\ r_{max}]$, in which r_{min} and r_{max} are lower and upper bounds of $r_n(t)$, respectively, for all $n \in \mathcal{N}$. The normalized reward is then given by

$$\hat{r}_n(t) = \frac{r_n(t) - r_{min}}{r_{max} - r_{min}}. \tag{9}$$

Note that $\hat{r}_n(t)$ is within interval $[0,\ 1]$.

For each antenna n, $\hat{r}_n(t)$ is approximated by sampling a random variable with Beta distribution, the probability distribution of which is given by

$$p_n\left(\Theta_n\right) = \frac{\Gamma\left(S_n + F_n\right)}{\Gamma(S_n)\Gamma(F_n)} \Theta_n^{S_n - 1} \left(1 - \Theta_n\right)^{F_n - 1}, \tag{10}$$

where Γ is the gamma function, S_n and F_n are Beta distribution parameters. The value of Θ_n is between $[0,\ 1]$ with mean $S_n/(S_n + F_n)$ and its variance decreases with $S_n + F_n$. Thus, from a statistic point of view, the sampling result of Θ_n increases with S_n and decreases with F_n, and it will be concentrated around its mean value if $S_n + F_n$ increase. We denote $\theta_n(t)$ as the sampling result of Θ_n for time slot t. Then, the set of antennas that maximizes the approximated reward is selected as the current transmit antenna set, i.e.,

$$\mathcal{K}(t) = \arg\max_{\mathcal{K} \in \mathcal{S}(\mathcal{N})} \sum_{a_k \in \mathcal{K}} \theta_{a_k}(t). \tag{11}$$

Note that this optimization problem can be easily solved by selecting the antennas with the largest samples. The corresponding system capacity $C(\mathcal{K}(t))$ is determined after the selection of antenna set $\mathcal{K}(t)$. We distribute it equally to the K antennas in $\mathcal{K}(t)$ and the antenna-specific reward $r_n(t)$ is then uniformly given by

$$r(t) = r_n(t) = \frac{1}{K} C(\mathcal{K}(t)). \tag{12}$$

The normalized antenna-specific reward $\hat{r}(t)$ is then utilized to update the corresponding parameters of the Beta distribution $Beta(S_n, F_n)$. Specifically, for each antenna n, we take a sample $b_n(t)$ from a Bernoulli variable with success probability $\hat{r}_n(t)$. If $b_n(t)$=1, then S_n increases by 1 while F_n remains unchanged. Otherwise, F_n increases by 1 while S_n remains unchanged. Thus, the iterative strategy of S_n and F_n are given by

$$\begin{aligned} S_n(t+1) &= S_n(t) + b_n(t), \\ F_n(t+1) &= F_n(t) + 1 - b_n(t). \end{aligned} \tag{13}$$

As we can see from (13), if the antenna-specific reward $r(t)$ is large, then there is a high probability to increase S_n by 1, which increases the mean value of Θ_n that approximates r_n. If the antenna-specific reward $r(t)$ is small, then there is a high probability to increase F_n by 1, which decreases the mean value of Θ_n that approximates r_n. As S_n and F_n increase, the variance of Θ_n decreases and the sampling result θ_n will be concentrated within a small interval around its mean value, until it approaches the antenna-specific reward $\hat{r}_n$. The exploitation and exploration tradeoff is natrually captured by the sampling process, which prefers to select antennas with stable high-performance after sufficient times of channel estimation, as well as antennas with uncertain performance when channel estimation is not enough.

B. Discount Factors

In practical systems, the time-varying nature of wireless channels will make the achieved CSI in previous slots outdated after a certain period of time, which implies the antenna-specific reward $r(t)$ as well as the corresponding parameters r_{min} and r_{max} should be dynamically updated. Specifically, we introduce two discount factors $\alpha_u \in (0,\ 1)$ and $\alpha_l > 1$ to update the value of r_{min} and r_{max}, respectively. If the antenna specific reward $r(t)$ is within interval $[r_{min},\ r_{max}]$, i.e., $r_{min} \le r(t) \le r_{max}$, we shrink the interval as follows:

$$\begin{aligned} r_{max} &= \alpha_u r_{max}, \\ r_{min} &= \alpha_l r_{min}. \end{aligned} \tag{14}$$

Otherwise, if the antenna specific reward $r(t)$ is greater than r_{max}, i.e., $r(t) > r_{max}$, we keep the lower bound unchanged and update the upper bound as

$$r_{max} = r(t), \tag{15}$$

Algorithm 1 Online Antenna Selection

Initialize the Beta distribution factors S_n, F_n for each antenna $n \in \mathcal{N}$, and the reward bounds r_{max} and r_{min}.
for $t = 1, 2, 3 \cdots;$ **do**
  Sample $\theta_n(t)$ from Beta distribution $Beta(S_n, F_n)$ for each antenna $n \in \mathcal{N}$;
  Select antenna set $\mathcal{K}(t) \in \mathcal{S}(\mathcal{N})$ as given in (11);
  Calculate $r(t)$ as given in (12) and $\hat{r}(t)$ based on (9);
  if $r_{min} \le r(t) \le r_{max}$ **then**
    Update r_{max}, r_{min} using (14);
  else if $r(t) > r_{max}$ **then**
    Update r_{max} using (15);
  else if $r(t) < r_{min}$ **then**
    Update r_{min} using (16);
  end if
  for each antenna $n \in \mathcal{K}(t)$, **do**
    Sample $b_n(t)$ from Bernoulli distribution with success probability $\hat{r}(t)$;
    Update S_n, F_n using (17);
  end for
end for

and if the antenna specific reward $r(t)$ is less than r_{min}, i.e., $r(t) < r_{min}$, we keep the upper bound unchanged and update the lower bound as

$$r_{min} = r(t). \tag{16}$$

Therefore, the interval $[r_{min},\ r_{max}]$ can track the changes of antenna-specific reward $r(t)$.

Also, we note that the previous CSI outside channel coherence time is irrelevant for estimating the current CSI. Thus, we introduce a discount factor γ to eliminate the influence of outdated CSI. The iterative strategy given by (13) is then rewritten as

$$\begin{aligned} S_n(t+1) &= \gamma S_n(t) + b_n(t), \\ F_n(t+1) &= \gamma F_n(t) + 1 - b_n(t). \end{aligned} \tag{17}$$

The overall online AS algorithm is given in **Algorithm 1**.

V. Simulation Results

In this section, we validate the proposed online AS algorithm by using the COST 2100 channel model [13], and compare it with two conventional AS algorithms, i.e., the power-based and the convex relaxation based algorithms [7]. In the power-based algorithm, the antennas with the highest K received power are selected as the transmit antennas in each slot. In the convex relaxation based algorithm, the optimal antenna set is calculated by using convex relaxation techniques. For the proposed online algorithm, the supper parameters are given by $\alpha_u = 0.95$, $\alpha_l = 1.05$ and $\gamma = 0.999$.

The COST 2100 specification defines a geometry-based stochastic channel model, in which the environment is static and the users are moving at a uniform speed. Therefore, the dynamic and time-variant channel behavior can be evaluated by setting different speeds of mobile users. Specifically, we consider a system with a bandwidth of 10 MHz at central

TABLE I
SIMULATION PARAMETERS

Parameter	Value
Slot length	0.5 ms
Pilot length	0.071 ms
Central frequency	1 GHz
Total bandwidth	10 MHz
Interference-free per-user SNR	$\rho = [-5,\ 20]$ dB
Number of base station antennas	$N = 128$
Number of RF chains	$K = [24,\ 80]$
Number of mobile users	$M = 4$
Velocity of users	$v = [0,\ 40]$ m/s

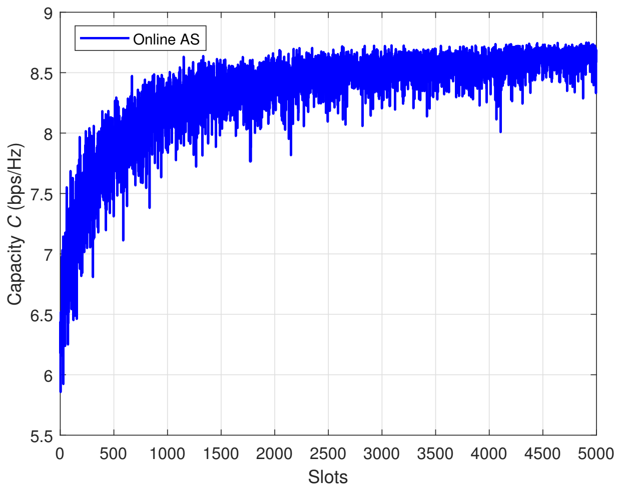


Fig. 3. Capacity C as a function of time slots. The number of RF chains is $K = 32$, the user velocity is $v = 0$ m/s, and the interference-free per-user SNR is $\rho = -5$ dB.

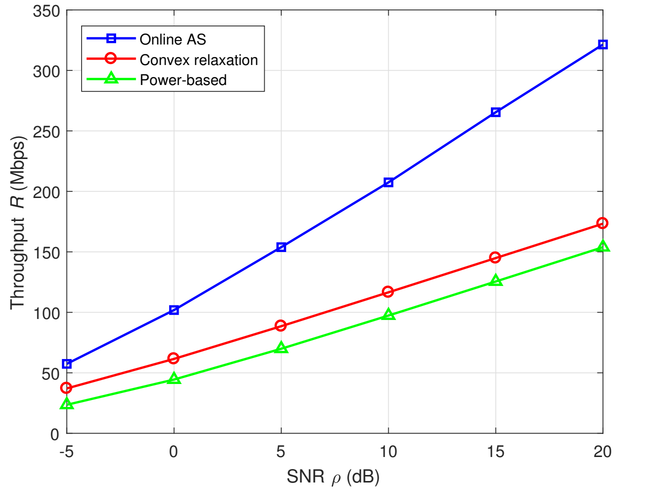


Fig. 4. Average throughput R as a function of the interference-free per-user SNR ρ. The user velocity is $v = 1$ m/s and the number of RF chains is $K = 32$.

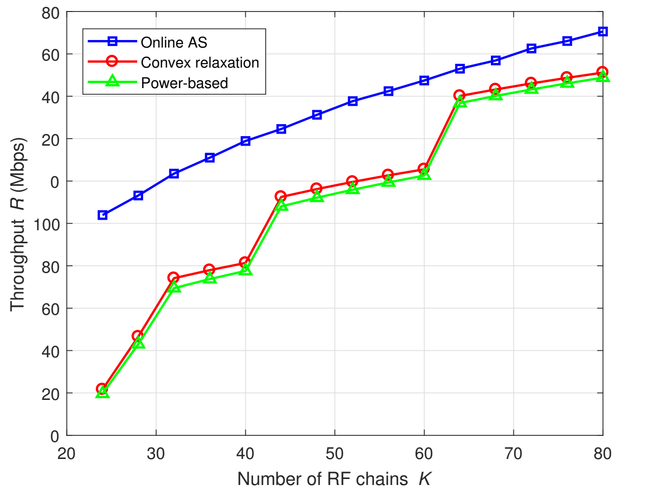


Fig. 5. Average throughput R as a function of the number RF chains K. The user velocity is $v = 1$ m/s and the interference-free per-user SNR is $\rho = -5$ dB.

frequency 1 GHz in an outdoor scenario. The interference-free per-user SNR ρ ranges from -5 to 20 dB. The number of available antennas is $N = 128$. The number of RF chains ranges from 24 to 80 so that the length of pilots does not occupy the entire slot of full CSI and the advantages of AS can be guaranteed. The number of mobile users is $M = 4$, and their velocities are uniformly given between 0 and 40 m/s to simulate pedestrian and vehicular movement. We adopt the slot structure of a massive MIMO testbed proposed in [14], in which the slot length is $T_{slot} = 0.5$ ms and each pilot consists of 7 symbols, i.e., $T_{pilot} = T_{slot}/7$. The parameters are summarised in Table I.

In Fig. 3, we show the system capacity as a function of time slots, in which the number RF chains is $K = 32$, the interference-free per-user SNR is $\rho = -5$ dB and the users are assumed to be static, i.e., $v = 0$ m/s. As we can see, the curve converges after about 2000 times iterative channel estimations, which implies that the antenna-specific reward can be well approximated by the corresponding Beta distribution and the selected antenna set that maximizes the system throughput becomes stable.

In Fig. 4, we show the average throughput as a function of the interference-free per-user SNR when the user velocity is $v = 1$ m/s and the number of RF chains is $K = 32$. As we can see, the proposed online algorithm outperforms the power-based and the convex relaxation based algorithms by $61\%-83\%$ and $100\%-110\%$, respectively. The reason for the improvement is that the pilot overhead is highly reduced in the proposed online algorithm. In the power-based and the convex relaxation based algorithms, in order to achieve full CSI of all available antennas, $\lceil N/K \rceil = 4$ symbols are utilized for uplink pilot transmission in each slot, while in our proposed online algorithm, only partial CSI is required, and 1 symbol is utilized for uplink pilot transmission in each slot.

In Fig. 5, we show the average throughput as a function of the number of RF chains when the user velocity is $v = 1$ m/s

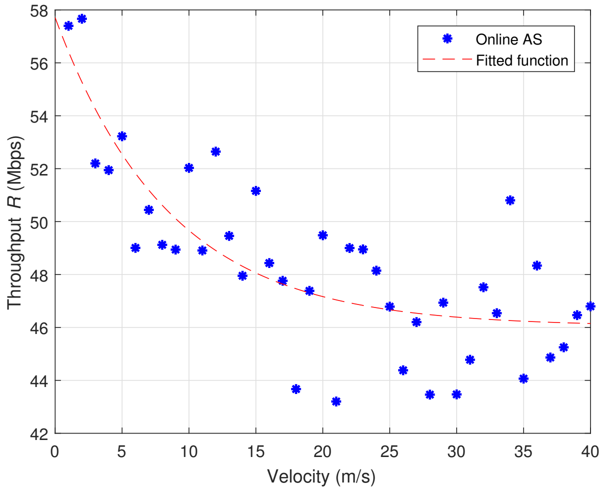


Fig. 6. Average throughput R as a function of user velocity with $K = 32$, $\rho = -5$ dB. The data is fitted as an exponential function $f(x) = a \cdot exp(-b \cdot x) + c$ with $a = 11.7$, $b = 0.12$ and $c = 46.0$.

and the interference-free per-user SNR is $\rho = -5$ dB. The number of RF chains ranges from 24 to 80. As we can see, The throughput increases with the number of RF chains. For the proposed online algorithm, the system capacity increases with the number of transmit antennas, and thus, it increases smoothly with the number of RF chains. However, for the algorithms using full CSI, the throughput is influenced not only by the real-time capacity, but also the data transmission time of each slot, which is a step function of RF chains. Thus, we can see the throughput hops at $K = 24$, 28, 40, and 60 as in Fig. 5. As the number of RF chains K approaches the number of available antennas N, the pilot overhead of full CSI algorithms decreases and the average throughput approaches the performance of massive MIMO without AS. Meanwhile, the proposed online scheme can achieve a nearly full CSI update in each slot, and it can track the capacity change of the channels with a competitive performance.

In Fig. 6, we show the average throughput as a function of the user velocity when the number of RF chains is $K = 32$ and the interference-free per-user SNR is $\rho = -5$dB. The velocity of mobile users increases from 1 m/s to 40 m/s in the COST 2100 channel model. As we can see, the average throughput is fluctuating with increasing velocity, and we fit a curve for these points. We see that the average throughput as a function of the user velocity can be fitted into an exponential function $f(x) = 11.7exp(-0.12x) + 46.0$. The fitted curve shows that the throughput performance is degraded by the increase of user mobility. The average throughput is decreased by 20% at $v = 20$ m/s, compared with the scenario when the users are static. The reason for the degradation is that the increase of user mobility will reduce the channel coherence time, which makes it more difficult to track the dynamics of wireless channels by using our proposed online algorithm.

VI. Conclusion

In this paper, we have considered the transmit AS in TDD massive MIMO systems, in which K out of N transmit antennas are selected to maximize the total throughput of M single-antenna users in the downlink. We have formulated the AS problem as a combinatorial multi-armed bandit problem and proposed an online algorithm by using Thompson sampling techniques, which highly reduces the pilot overhead by tracking the dynamics of wireless channels with partial CSI. The proposed algorithm can provide a tradeoff between the benefits of exploiting high-performance antennas and the cost of exploring antennas with uncertain performance through the sampling process. Simulation results have shown that the proposed online AS algorithm has a significant performance improvement in system throughput, as compared to power-based and convex relaxation based algorithms. For a system with 60 out of 128 antennas selected, 4 pedestrian users with velocity 1 m/s and interference-free per-user SNR -5 dB, the throughput improvement is 23% and 19%, respectively, compared to the power-based and convex relaxation based algorithms. The improvement is higher when the number of RF chains is small and the user mobility is low.

References

[1] S. Sanayei and A. Nosratinia, "Antenna selection in MIMO systems," *IEEE Commun. Mag.*, vol. 42, no. 10, pp. 68–73, Oct. 2004.

[2] T. S. Rappaport, S. Sun, R. Mayzus, H. Zhao, Y. Azar, K. Wang, G. N. Wong, J. K. Schulz, M. Samimi, and F. Gutierrez, "Millimeter wave mobile communications for 5g cellular: It will work!" *IEEE Access*, vol. 1, pp. 335–349, 2013.

[3] A. Gorokhov, D. A. Gore, and A. J. Paulraj, "Receive antenna selection for MIMO spatial multiplexing: theory and algorithms," *IEEE Trans. Signal Process.*, vol. 51, no. 11, pp. 2796–2807, Nov. 2003.

[4] M. T. Kakitani, G. Brante, R. D. Souza, and M. A. Imran, "Energy efficiency of transmit diversity systems under a realistic power consumption model," *IEEE Commun. Lett.*, vol. 17, no. 1, pp. 119–122, Jan. 2013.

[5] A. Mukherjee and A. Hottinen, "Learning algorithms for energy-efficient mimo antenna subset selection: Multi-armed bandit framework," in *Proc. EUSIPCO'12*, Aug. 2012, pp. 659–663.

[6] T.-W. Ban and B. C. Jung, "A practical antenna selection technique in multiuser massive MIMO networks," *IEICE Trans. Commun.*, vol. 96, no. 11, pp. 2901–2905, Nov. 2013.

[7] 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, Nov. 2015.

[8] A. Konar and N. D. Sidiropoulos, "A simple and effective approach for transmit antenna selection in multiuser massive MIMO leveraging submodularity," *IEEE Trans. Signal Process.*, vol. 66, no. 18, pp. 4869–4883, Sep. 2018.

[9] H. Q. Ngo, *Massive MIMO: Fundamentals and system designs*. Linköping University Electronic Press, 2015, vol. 1642.

[10] N. Cesa-Bianchi and G. Lugosi, "Combinatorial bandits," *J. Comp. Sys. Sci.*, vol. 78, no. 5, pp. 1404–1422, Sep. 2012.

[11] Y. Lin, T. Wang, and S. Wang, "UAV-assisted emergency communications: An extended multi-armed bandit perspective," *IEEE Commun. Lett.*, vol. 23, no. 5, pp. 938–941, May 2019.

[12] D. J. Russo, B. Van Roy, A. Kazerouni, I. Osband, Z. Wen *et al.*, "A tutorial on thompson sampling," *Foundations and Trends® in Machine Learning*, vol. 11, no. 1, pp. 1–96, Jul. 2018.

[13] L. Liu, C. Oestges, J. Poutanen, K. Haneda, P. Vainikainen, F. Quitin, F. Tufvesson, and P. D. Doncker, "The cost 2100 MIMO channel model," *IEEE Wireless Commun.*, vol. 19, no. 6, pp. 92–99, Dec. 2012.

[14] J. Vieira, S. Malkowsky, K. Nieman, Z. Miers, N. Kundargi, L. Liu, I. Wong, V. wall, O. Edfors, and F. Tufvesson, "A flexible 100-antenna testbed for massive MIMO," in *Proc. IEEE Globecom'14 (Wkshps)*, Dec. 2014, pp. 287–293.