# Predictive Resource Allocation in mmWave Systems with Rotation Detection

Yifei Sun[†‡], Bojie Lv[†], Rui Wang[†], Haisheng Tan[§], and Francis C. M. Lau[‡]

[†]Southern University of Science and Technology, Shenzhen, China
[‡]The University of Hong Kong, Hong Kong, China
[§]University of Science and Technology of China, Hefei, China

*Abstract*—**Millimeter wave (MmWave) has been regarded as a promising technology to support high-capacity communications in 5G era. However, its high-layer performance such as latency and packet drop rate in the long term highly depends on resource allocation because mmWave channel suffers significant fluctuation with rotating users due to mmWave sparse channel property and limited field-of-view (FoV) of antenna arrays. In this paper, downlink transmission scheduling considering rotation of user equipments (UE) and limited antenna FoV in an mmWave system is optimized via a novel approximate Markov decision process (MDP) method. Specifically, we consider the joint downlink UE selection and power allocation in a number of frames where future orientations of rotating UEs can be predicted via embedded motion sensors. The problem is formulated as a finite-horizon MDP with non-stationary state transition probabilities. A novel low-complexity solution framework is proposed via one iteration step over a base policy whose average future cost can be predicted with analytical expressions. It is demonstrated by simulations that compared with existing benchmarks, the proposed scheme can schedule the downlink transmission and suppress the packet drop rate efficiently in non-stationary mmWave links.**

## I. INTRODUCTION

Millimeter wave (MmWave) communications is one of the key technologies in 5G and beyond systems [1]. In order to achieve desired signal-to-noise ratio (SNR) and high data rate, phased arrays with highly directional beams are adopted to overcome huge pathloss. However, because of the sparse propagation paths in mmWave channel between the base station (BS) and user equipments (UEs), pencil-shaped beams and limited field-of-view (FoV) of antenna arrays, UE mobility especially rotation will cause significant SNR fluctuation [2]. Ignorance of such SNR fluctuation may lead to severe high-layer performance degradation such as large latency, buffer overflow and packet drop. Fortunately, exploiting statistical channel model and motion information measured from embedded motion sensors at the UEs makes future SNR fluctuation predictable. This raises a new design issue of joint scheduling in a large time-scale with non-stationary but predictable channel statistics caused by UE rotation.

There has been a number of works considering resource allocation in mmWave MIMO systems either within channel coherent time [3] or in a larger time-scale with stationary channel statistics [4]–[7]. Scheduling in the former scenario leads to deterministic optimization problems, and infinite-horizon Markov decision process (MDP) or Lyapunov optimization is usually applied for the latter scenario. However, both methods are not applicable for large-time-scale scheduling with non-stationary channel statistics. Scheduling over non-stationary channel statistics with prediction of future channel statistics shall be addressed. For example, the authors in [8] proposed an adaptive design for beam alignment, data transmission and handover by exploiting vehicle mobility in mmWave vehicular networks. In [9], [10], predictive beamforming is investigated by sensing vehicle mobility with radar to improve the efficiency of resource allocation. However, these works focus on vehicles driving along straight lanes considering only translation while UE rotation may raise more stringent requirement [2]. Moreover, current works investigating temporal correlations of angle of departure (AoD) and angle of arrival (AoA) [11] or exploiting motion sensors to capture device rotation [2], [12] to assist beam alignment neglect the effect of rapid SNR fluctuation to high-layer performance.

Channel fluctuation due to limited antenna FoV and UE rotation shall be considered in mmWave scheduling. Fortunately, future UE orientation can be predicted according to current angular velocity or angular acceleration measured by motion sensors [13], [14] or integrated sensing and communications (ISAC) techniques [15]. Moreover, due to the small-scale fading, it might not be efficient to determine all the transmission parameters in a large time-scale at its beginning as in [8]–[10]. Instead, dynamic programming might be a better framework to address the predictive scheduling algorithm design with random channel fading. In this paper, we consider the downlink scheduling with UE rotation where the angular velocity can be measured by motion sensors.

In this paper, we would like to shed some light on the predictive transmission scheduling in mmWave systems with channel fluctuation due to UE rotation and limited antenna FoV. We utilize the motion sensors embedded in UEs to track the orientation of phased arrays. Moreover, by exploiting temporal correlations of mmWave channel, the non-stationary channel statistics can be predicted. We formulate the delay-aware transmission scheduling with non-stationary mmWave channel statistics as a finite-horizon MDP. Finally, a low-complexity approximate MDP solution framework, applying one iteration step over analytically approximated value function, is proposed to reduce the computational complexity.

To the best of our knowledge, this is the first paper on mmWave queue-aware scheduling in a large time-scale with consideration of UE rotation.

## II. SYSTEM MODEL

### A. mmWave System with UE Rotation

We consider the downlink transmission in an mmWave communication system with one BS and $K$ UEs, where the set of UEs is denoted by $\mathcal{K} \triangleq \{1, 2, \ldots, K\}$. The analog MIMO transceiver with a single radio frequency (RF) chain and a half-wavelength uniform linear phased array (ULA) is adopted at both the BS and UEs. The linear phased arrays at the BS and UEs are with $N_{\mathrm{T}}$ and $N_{\mathrm{R}}$ antenna elements, respectively. Hence, analog precoder and combiners can be adopted at the BS and UEs respectively to enhance the receiving SNR. For elaboration convenience, we consider the mmWave communication in a two-dimensional plane as illustrated in Fig. 1(a).

We focus on mmWave communication scenarios for UEs with rotation but no translation, e.g., playing games with virtual reality (VR) headsets or watching videos when sitting on a rotating chair. Due to UE rotation and limited antenna FoVs, channel statistics such as the number of in-FoV signal scatterers and their directions with respect to the phased arrays will change over time. Fortunately, embedded motion sensors such as magnetometers and gyroscopes are able to periodically detect the orientation and rotation of UEs, respectively [2]. Moreover, future UE orientations can be predicted by current angular velocity [13]. With the direction knowledge of scatterers and phased arrays, the channel statistics in the future can also be predicted to assist predictive scheduling by jointly considering current and future transmission costs.

Specifically, the transmission time is organized by frames, and the wireless channel is assumed to be quasi-static within one frame. UEs are rotating with predictable angular velocities in a number of frames. The period during which all the UEs are rotating with predictable angular velocities is referred to as one *scheduling period* consisting of $T$ frames. For the elaboration convenience, we assume that all UEs are predicted to rotate with constant angular velocity as in [14], [16] in a scheduling period. We shall focus on the joint UE selection and power allocation within one scheduling period.

In the considered scheduling period, the angular velocity of the $k$-th UE is denoted as $\omega_k$. The boresight direction of the $k$-th UE in the 1-st frame of the scheduling period is denoted as $\mathbf{n}_{1,k}$, and the boresight direction of the BS's array is denoted as $\mathbf{n}_{BS}$. Then the rotation angle during the $(t-1)$ frames and the boresight direction of the $k$-th UE in the $t$-th frame can be predicted as $\Delta\phi_{t,k} \triangleq (t-1)\omega_k T_{\mathrm{F}}$ and $\mathbf{n}_{t,k} = \begin{bmatrix} \cos(\Delta\phi_{t,k}) & -\sin(\Delta\phi_{t,k}) \\ \sin(\Delta\phi_{t,k}) & \cos(\Delta\phi_{t,k}) \end{bmatrix} \mathbf{n}_{1,k}$, respectively, where $t \le T$ and $T_{\mathrm{F}}$ denotes the frame duration.

### B. Channel Model

The extended Saleh-Valenzuela channel model illustrated in Fig. 1(b) is adopted. Specifically, there are $N_k^{\mathrm{cl}}$ quasi-static scattering clusters in the propagation channel from the BS to
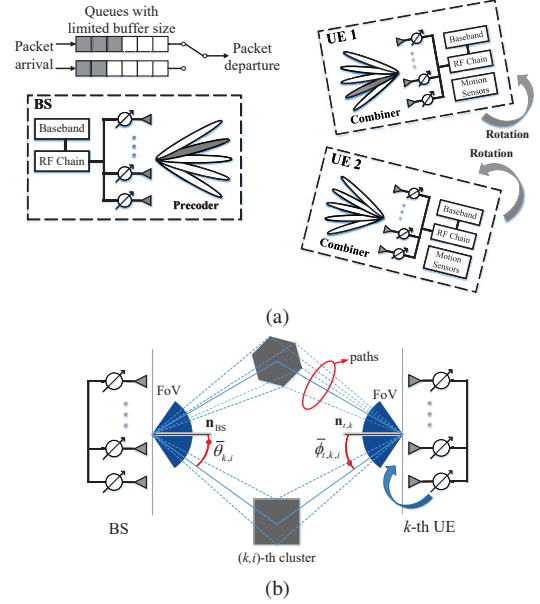


Fig. 1. Illustration of (a) system model and (b) channel model.

the $k$-th UE, and $N_{k,i}^{\mathrm{ray}}$ propagation paths in the $i$-th cluster. For the exposition convenience, the $i$-th cluster in the channel between the BS and the $k$-th UE is referred to as the $(k,i)$-th cluster, and the $\ell$-th path of the $(k,i)$-th cluster is referred to as the $(k,i,\ell)$-th path. Denoting the AoA and AoD of the $(k,i,\ell)$-th path in the $t$-th frame as $\phi_{t,k,i,\ell}$ and $\theta_{t,k,i,\ell}$ respectively, the channel matrix from the BS to the $k$-th UE in the $t$-th frame $\mathbf{H}_{t,k} \in \mathbb{C}^{N_{\mathrm{R}} \times N_{\mathrm{T}}}$ can be represented by

$$\mathbf{H}_{t,k} = \sum_{i=1}^{N_k^{\mathrm{cl}}} \sum_{\ell=1}^{N_{k,i}^{\mathrm{ray}}} \alpha_{t,k,i,\ell} \mathbf{a}_{\mathrm{R}}(\phi_{t,k,i,\ell}) \mathbf{a}_{\mathrm{T}}^{\mathsf{H}}(\theta_{t,k,i,\ell})$$
$$\times \Lambda_{\mathrm{R}}(\phi_{t,k,i,\ell}) \Lambda_{\mathrm{T}}(\theta_{t,k,i,\ell}), \quad (1)$$

where $\alpha_{t,k,i,\ell}$ is the instantaneous complex gain of the $(k,i,\ell)$-th path in the $t$-th frame, $\Lambda_{\mathrm{R}}(\phi_{t,k,i,\ell})$ and $\Lambda_{\mathrm{T}}(\theta_{t,k,i,\ell})$ refer to the receiving and transmission antenna gains at $\phi_{t,k,i,\ell}$ and $\theta_{t,k,i,\ell}$ respectively, $\mathbf{a}_{\mathrm{R}}$ and $\mathbf{a}_{\mathrm{T}}$ represent the array response vectors of the ULAs at UEs and the BS, which can be expressed by $\mathbf{a}_{\mathrm{R}}(\phi) = \frac{1}{\sqrt{N_{\mathrm{R}}}} \left[ 1, e^{-j\pi \sin(\phi)}, \ldots, e^{-j\pi(N_{\mathrm{R}}-1)\sin(\phi)} \right]^{\mathsf{T}}$, and $\mathbf{a}_{\mathrm{T}}(\theta) = \frac{1}{\sqrt{N_{\mathrm{T}}}} \left[ 1, e^{-j\pi \sin(\theta)}, \ldots, e^{-j\pi(N_{\mathrm{T}}-1)\sin(\theta)} \right]^{\mathsf{T}}$, respectively. The antenna patterns are modeled as identical and ideal sectored, hence,

$$\Lambda_{\mathrm{R}}(\phi) = \begin{cases} 1 & \phi \in [\phi_{\min}, \phi_{\max}] \\ 0 & \text{otherwise} \end{cases}, \Lambda_{\mathrm{T}}(\theta) = \begin{cases} 1 & \theta \in [\theta_{\min}, \theta_{\max}] \\ 0 & \text{otherwise} \end{cases}.$$

Due to the limited FoVs of receiving antennas, i.e., $\Lambda_{\mathrm{R}}$, the phased array can not capture all propagation paths in $360°$ azimuth. Thus, the in-FoV scattering clusters may vary during UE rotation. For example, the paths via the scattering cluster represented by black square in Fig. 1(b) may be out of the FoV in a few frames when the $k$-th UE is rotating clockwise.

The instantaneous gains $\{\alpha_{t,k,i,\ell}\}$, AoAs $\{\phi_{t,k,i,\ell}\}$ and AoDs $\{\theta_{t,k,i,\ell}\}$ are drawn from independent distributions in each frame. Specifically, $\alpha_{t,k,i,\ell}$ follows the complex Gaussian distribution with zero mean and variance $\sigma_{\alpha;k,i}^2$. The angular distribution of propagation paths $\phi_{t,k,i,\ell}$ and $\theta_{k,i,\ell}$ follow the truncated Laplacian distributions with cluster means $\bar{\phi}_{t,k,i}$ and

$\bar{\theta}_{k,i}$, and variances $\sigma^2_{\phi;k,i}$ and $\sigma^2_{\theta;k,i}$ respectively as in [17]. Due to UE rotation, the boresight directions of their phased arrays change in each frame. Since the scattering clusters are quasi-static, the AoA means of the $(k,i)$-th cluster in one scheduling period satisfy the following temporal correlation,

$$\bar{\phi}_{t,k,i} = \bar{\phi}_{1,k,i} + \Delta\phi_{t,k}, \forall t, k, i. \tag{2}$$

As a result, the channel statistics (distribution of $\mathbf{H}_{t,k}$) are non-stationary in one scheduling period, due to the limited FoVs at the UEs and time-varying AoA cluster means $\bar{\phi}_{t,k,i}$. However, with the AoA cluster means and the angular velocity of the $k$-th UE in the 1-st frame, i.e., $\{\bar{\phi}_{1,k,i}|\forall i\}$ and $\omega_k$, the AoA cluster means of all clusters in the $t$-th frame can be predicted according to (2).

For the elaboration convenience, we define the statistical channel state information (SCSI) as the tuple of parameters sufficiently characterizing the distribution of channel matrix in (1) as follows.

**Definition 1.** *The SCSI in $t$-th frame is defined by $\mathcal{I}_t^{\text{SCSI}} \triangleq (\mathcal{I}_{\text{sta}}^{\text{SCSI}}, \{\bar{\phi}_{t,k,i}|\forall k, i\})$, where $\mathcal{I}_{\text{sta}}^{\text{SCSI}} \triangleq (\{N_k^{\text{cl}}|\forall k\}, \{N_{k,i}^{\text{ray}}|\forall k, i\}, \{\sigma^2_{\alpha;k,i}|\forall k, i\}, \{\bar{\theta}_{k,i}|\forall k, i\}, \{\sigma^2_{\theta;k,i}|\forall k, i\}, \{\sigma^2_{\phi;k,i}|\forall k, i\})$ is the tuple of quasi-static channel statistical parameters.*

With SCSI in the 1-st frame $\mathcal{I}_1^{\text{SCSI}}$ and $\{\omega_k\}$ measured in the 1-st frame, the SCSI and thus the distribution of $\mathbf{H}_{t,k}$ of all the frames within the scheduling period can be predicted according to the statistical channel model in (1).

### C. Sensing-based Beam Alignment

Due to the single RF chain at the BS, one UE is selected for downlink transmission in each frame. Let $\mathbf{w}_{t,k} \in \mathbb{C}^{N_R \times 1}$ and $\mathbf{f}_{t,k} \in \mathbb{C}^{N_T \times 1}$ be the analog combiner at the $k$-th UE and the analog precoder at the BS in the $t$-th frame respectively if the $k$-th UE is selected. In practice, $\mathbf{w}_{t,k}$ and $\mathbf{f}_{t,k}$ are selected from the pre-defined finite codebooks $\mathbf{w}_{t,k} \in \mathcal{W} \triangleq \{\mathbf{a}_R(\phi_q)|q = 1, 2, \ldots, N_R\}$ and $\mathbf{f}_{t,k} \in \mathcal{F} \triangleq \{\mathbf{a}_T(\theta_p)|p = 1, 2, \ldots, N_T\}$, respectively, where $\phi_q = \arcsin\left(\frac{2(q-1)}{N_R} - 1\right)$ and $\theta_p = \arcsin\left(\frac{2(p-1)}{N_T} - 1\right)$. Hence, the spectral efficiency achieved by the $k$-th UE in the $t$-th frame is given by $R_{t,k} = \log_2\left(1 + \frac{P_{t,k}Y_{t,k}}{N_0 W}\right)$, where $P_{t,k}$ is the transmission power of the BS, $Y_{t,k} \triangleq |\mathbf{w}_{t,k}^H \mathbf{H}_{t,k} \mathbf{f}_{t,k}|^2$ is the channel power gain in baseband, $N_0$ is the noise power spectral density, and $W$ is the bandwidth.

Instead of beamforming with instantaneous CSI feedback which raises significant overhead, we exploit the SCSI prediction and adopt the following statistical beam alignment scheme, which maximizes the average baseband SNR without instantaneous CSI feedback.

**Scheme 1** (SCSI-Based Beam Alignment). *Given $\mathcal{I}_1^{\text{SCSI}}$ and $\{\omega_k\}$, the analog combiner and precoder for the $k$-th UE in the $t$-th frame are selected by $\mathbf{w}_{t,k} = \mathbf{a}_R(\phi_{q_t^\dagger})$ and $\mathbf{f}_{t,k} = \mathbf{a}_T(\theta_{p_t^\dagger})$, respectively, where $(q_t^\dagger, p_t^\dagger)$ is given by (3).*

Since the integrals in (3) depends on SCSI in the $t$-th frame $\mathcal{I}_t^{\text{SCSI}}$ which can be predicted from $\mathcal{I}_1^{\text{SCSI}}$ and $\{\omega_k\}$, the precoders and combiners for all the frames within one

scheduling period can be pre-designed in the beginning of 1-st frame. Note that different from the statistical beamforming proposed in [18], we consider non-stationary AoA/AoD cluster means and limited FoVs of ULA. Therefore, beam may switch towards another cluster in advance when previously steered cluster becomes out of FoV due to UE rotation. Moreover, given Scheme 1, the cumulative distribution function (CDF) of the baseband channel power gain can be derived as follows.

**Lemma 1** (CDF of $Y_{t,k}$). *With Scheme 1, when $N_R$ and $N_T$ are sufficiently large, the CDF of $Y_{t,k}$ is given by (4), where $\Phi_{t,k,i}$ follows the binomial distribution given by (5), and*

$$\mathcal{P}_{R;q} = \left\{\phi \,\middle|\, \left|\sin(\phi) - \frac{2(q-1) - N_R}{N_R}\right| \le \frac{1}{N_R}\right\}, \tag{6}$$

$$\mathcal{P}_{T;p} = \left\{\theta \,\middle|\, \left|\sin(\theta) - \frac{2(p-1) - N_T}{N_T}\right| \le \frac{1}{N_T}\right\}. \tag{7}$$

*Proof.* Please refer to Appendix A. $\square$

### D. System Queue Dynamics

There are $K$ queues for downlink transmission at the BS, each for one UE. The arrival data of each queue is organized by packets, each with $B$ information bits. It is assumed that the number of arrival packets at the $k$-th UE in the $t$-th frame, denoted as $A_{t,k}$, follows independent Poisson distribution with expectation $\lambda_k$ across UEs as in [19], i.e., $\Pr[A_{t,k} = n] = (\lambda_k^n/n!)e^{-\lambda_k}$. Let $\mathcal{A}_t \triangleq \{A_{t,k}|\forall k \in \mathcal{K}\}$ represent the aggregated packet arrivals in the $t$-th frame. Without loss of generality, it is assumed that all packets arrive at the end of each frame.

Suppose the $d_t$-th UE is selected in the $t$-th frame, the number of departure packets from the $d_t$-th queue in the $t$-th frame is given by $D_{t,d_t} = \lfloor W R_{t,d_t} T_F / B \rfloor$. Denote $Q_{t,k}$ as the queue length of the $k$-th queue at the beginning of $t$-th frame and $Q_{\max}$ as the buffer size for each queue both in terms of packets. The queue dynamics can be expressed as $Q_{t+1,k} = \min\{Q_{t,k}^D + A_{t,k}, Q_{\max}\}$, where the arrival packets will be discarded if the buffer is full. The post-decision queue length $Q_{t,k}^D$ is defined as

$$Q_{t,k}^D = \begin{cases} (Q_{t,k} - D_{t,k})^+ & k = d_t, \\ Q_{t,k} & k \ne d_t, \end{cases} \tag{8}$$

where $(\cdot)^+ \triangleq \max(0, \cdot)$.

## III. PROBLEM FORMULATION

Given the precoder and combiner design in Scheme 1, we shall formulate the UE selection and power allocation for all the frames in one scheduling period as a finite-horizon MDP. Note that this is because the system adopts the same prediction for UE rotations in all the frames of one scheduling period. In order to facilitate the MDP formulation, the system state, scheduling action and policy, and post-decision scheduling policy are first elaborated as follows.

**Definition 2** (System State). *At the beginning of the $t$-th frame, the system state is represented by $\mathcal{S}_t \triangleq (\mathcal{Q}_t, \mathcal{Y}_t)$, consisting of queuing state information (QSI) of all the UEs $\mathcal{Q}_t \triangleq \{Q_{t,1}, Q_{t,2}, \ldots, Q_{t,K}\}$, and baseband channel power gains to all the UEs $\mathcal{Y}_t \triangleq \{Y_{t,1}, Y_{t,2}, \ldots, Y_{t,K}\}$.*

$$(q_t^\dagger, p_t^\dagger) = \arg\max_{q,p} \mathbb{E}_{\mathbf{H}_{t,k}} \left[ \left| \mathbf{a}_R^H (\phi_q) \mathbf{H}_{t,k} \mathbf{a}_T (\theta_p) \right|^2 \right]$$

$$= \arg\max_{q,p} \sum_{i=1}^{N_k^{cl}} N_{k,i}^{ray} \sigma_{\alpha;k,i}^2 \left\{ \int_{\phi_{\min}}^{\phi_{\max}} \left| \mathbf{a}_R^H (\phi_q) \mathbf{a}_R(\phi_{t,k,i,1}) \right|^2 f_{\phi;k,i}(\phi_{t,k,i,1}) \mathrm{d}\phi_{t,k,i,1} \right\}$$

$$\times \left\{ \int_{\theta_{\min}}^{\theta_{\max}} \left| \mathbf{a}_T^H (\theta_{t,k,i,1}) \mathbf{a}_T (\theta_p) \right|^2 f_{\theta;k,i}(\theta_{t,k,i,1}) \mathrm{d}\theta_{t,k,i,1} \right\}. \tag{3}$$

$$F_{Y_{t,k}}(x) \triangleq \Pr[Y_{t,k} \leq x] = \mathbb{E}_{\{\Phi_{t,k,i}|\forall i\}} \left\{ 1 - \exp(-x/\sum_{i=1}^{N_k^{cl}} \Phi_{t,k,i} \sigma_{\alpha;k,i}^2) \right\}, \quad x > 0, \tag{4}$$

$$\Phi_{t,k,i} \sim \mathtt{Binomial}\big( N_{k,i}^{ray}, \Pr\left[\phi_{t,k,i,\ell} \in \mathcal{P}_{R;q_t^\dagger} \cap [\phi_{\min}, \phi_{\max}]\right] \Pr\left[\theta_{t,k,i,\ell} \in \mathcal{P}_{T;p_t^\dagger} \cap [\theta_{\min}, \theta_{\max}]\right] \big) \tag{5}$$

---

**Definition 3** (Scheduling Action and Policy). *At the beginning of the $t$-th frame, the scheduling actions include the UE selection for downlink transmission $d_t \in \mathcal{K}$ and the downlink transmission power $P_t \triangleq P_{t,d_t}$, where the following instantaneous power constraint at the BS should be satisfied,*

$$P_t \leq P_{\max}, \quad \forall t. \tag{9}$$

*Hence, the scheduling policy of the BS, denoted as $\Omega_t$, is a mapping from the system state $\mathcal{S}_t$ to the scheduling actions. Thus, $\Omega_t(\mathcal{S}_t) \triangleq (d_t, P_t)$.*

**Definition 4** (Post-Decision State). *At the beginning of the $t$-th frame, the post-decision system state is defined by $\mathcal{S}_t^D \triangleq (\mathcal{Q}_t^D, \mathcal{Y}_t)$, where $\mathcal{Q}_t^D \triangleq \{Q_{t,1}^D, Q_{t,2}^D, \ldots, Q_{t,K}^D\}$ denotes post-decision QSI of all the UEs.*

The *post-decision system state* is the system state after packet transmission but before packet arrivals. Given the above definition of post-decision system state and scheduling policy, the transition probability is given by $\Pr(\mathcal{S}_{t+1}^D | \mathcal{S}_t^D, \Omega_{t+1}) = \Pr(\mathcal{Q}_{t+1}^D | \mathcal{Q}_t^D, \Omega_{t+1}) \Pr(\mathcal{Y}_{t+1}) = \prod_{k=1}^K \Pr(Q_{t+1,k}^D | Q_{t,k}^D, \Omega_{t+1,k}) \prod_{k=1}^K \Pr(Y_{t+1,k})$.

In this paper, the scheduling policies are designed to optimize the system queuing performance, while saving the average total energy consumption. Specifically, the transmission energy, average packet transmission delay and penalty of packet drop are considered as the costs of scheduling. According to Little's law, the sum of queuing packet numbers of all the frames in the scheduling period can be used as an equivalent measurement of average transmission delay. Hence, we define the following weighted sum of the transmission power consumption, the number of queuing packets for all UE and full buffer penalty as the system cost in the $t$-th frame,

$$g_t\big(\mathcal{S}_t, \Omega_t(\mathcal{S}_t)\big) \triangleq w_P P_t + \sum_{k \in \mathcal{K}} \big( Q_{t,k} + w_Q \mathbb{I}[Q_{t,k} = Q_{\max}]\big),$$

where $w_P$ and $w_Q$ are the weights of power consumption and full buffer penalty respectively, and $\mathbb{I}[\cdot]$ denotes an indicator function, which is 1 when the event is true and 0 otherwise.

The overall minimization objective of one scheduling period with the initial system state $\mathcal{S}_1$ is then given by

$$G(\mathcal{S}_1, \Omega) \triangleq \mathbb{E}_{\mathcal{A}, \mathcal{Y}}^\Omega \big[ \sum_{t=1}^T g_t(\mathcal{S}_t, \Omega_t(\mathcal{S}_t)) + \varrho(\mathcal{Q}_{T+1}) \big| \mathcal{S}_1 \big],$$

where $\mathcal{A} \triangleq \{\mathcal{A}_t | \forall t\}$, $\mathcal{Y} \triangleq \{\mathcal{Y}_t | \forall t\}$, $\Omega \triangleq \{\Omega_t | \forall t\}$, and $\varrho(\mathcal{Q}_{T+1}) \triangleq \sum_{k \in \mathcal{K}} Q_{T+1,k}$ counts for the remaining packet number at the end of one scheduling period. The expectation is taken with respect to the randomness of packet arrivals $\mathcal{A}$ and baseband channel power gains $\mathcal{Y}$. As a result, the transmission design in this paper can be formulated as the

following dynamic programming problem.

**P1 :** $\Omega^\star \triangleq \{\Omega_t^\star | \forall t\} = \arg\min_\Omega G(\mathcal{S}_1, \Omega)$

$$\text{s.t.} \quad 0 \leq P_t \leq P_{\max}, \quad \forall t.$$

We shall adopt the Bellman's equations with *post-decision value functions* to solve P1. Note that the baseband channel power gains are independently distributed in different frames, it can be averaged as in [19], and the Bellman's equation based on the post-decision QSI only can be written as

$$W_t(\mathcal{Q}_t^D) = \min_{\Omega_{t+1}(\mathcal{S}_{t+1})} \mathbb{E}_{\mathcal{A}_t, \mathcal{Y}_{t+1}} \big[ g_{t+1}\big(\mathcal{S}_{t+1}, \Omega_{t+1}(\mathcal{S}_{t+1})\big)$$
$$+ W_{t+1}\big(\mathcal{Q}_{t+1}^D\big) \big| \mathcal{Q}_t^D \big] \tag{10}$$

where $W_t(\mathcal{Q}_t^D)$ is the post-decision value function of the optimal policy (referred to as optimal value function for short), and $W_T(\mathcal{Q}_T^D) \triangleq \mathbb{E}_{\mathcal{A}_T} \varrho(\mathcal{Q}_{T+1} | \mathcal{Q}_T^D)$ for the notation convenience. The Bellman's equations with post-decision system state can avoid complicated calculation of transition matrix and expectation of value functions [20]. Moreover, the optimal scheduling policy in the $t$-th frame ($\forall t = 1, 2, \ldots, T$) for P1 can be obtained by

$$\Omega_t^\star(\mathcal{S}_t) = \arg\min_{\Omega_t(\mathcal{S}_t)} \big\{ g_t\big(\mathcal{S}_t, \Omega_t(\mathcal{S}_t)\big) + W_t(\mathcal{Q}_t^D) \big\} \tag{11}$$
$$\text{s.t.} \quad 0 \leq P_t \leq P_{\max}, \quad \forall t.$$

### IV. Low-Complexity Scheduling

It can be observed from (11) that the optimal value function for the $t$-th frame $W_t(\mathcal{Q}_t^D)$ should be calculated before the derivation of the optimal scheduling policy for the $t$-th frame $\Omega_t(\mathcal{S}_t)$. However, because of the minimization in (11), it is difficult to derive the closed-form expression for the optimal value function in each frame $\{W_t(\mathcal{Q}_t^D) | \forall t\}$. It is also difficult to calculate the value functions for all possible system states numerically due to the huge system state space. In this section, a low-complexity solution framework is proposed. Specifically, we first adopt the backpressure algorithm [21] as the base policy and derive the closed-form expressions of its value functions (referred to as approximate value function for short). Then the approximate value functions are used to approximate the optimal value functions to obtain the improved scheduling policy by one-step policy improvement.

#### A. Base Policy

As in (10), the base policy provides an approximation of average future cost to improve current scheduling actions. It should have a good scheduling performance and a simple structure for analysis. Hence, as the base policy, we predetermine the transmission powers and UE selection for all frames at the very beginning of one scheduling period. Particularly, given the system QSI in the 1-st frame, we use the average

downlink throughput and packet arrival rate to approximate the queue dynamics, and adopt the *backpressure algorithm* [21] to select the downlink UE. Let $P_{\mathrm{BSL}}$ be the predetermined transmission power and $Q_{t,k}^{\dagger}$ be the approximate queue length of the $k$-th queue at the beginning of the $t$-th frame, then the index of the selected UE is given by

$$d_t^{\Pi} = \arg\max_{k\in\mathcal{K}} \left[ R_{t,k}^{\dagger}(P_{\mathrm{BSL}})Q_{t,k}^{\dagger} \right], \ \forall t, \quad (12)$$

where $R_{t,k}^{\dagger}$ is the predicted average spectral efficiency of the $k$-th UE in the $t$-th frame, i.e., $R_{t,k}^{\dagger}(P_{\mathrm{BSL}}) \triangleq \int \log_2\left(1+\frac{P_{\mathrm{BSL}}x}{N_0 W}\right)\frac{\mathrm{d}F_{Y_{t,k}}(x)}{\mathrm{d}x}\mathrm{d}x, \ \forall t,k\in\mathcal{K}$. Moreover, given UE selection in the $t$-th frame, the QSI in the $(t+1)$-th frame can be approximated by

$$Q_{t+1,k}^{\dagger} = \begin{cases} \min\{(Q_{t,k}^{\dagger}-D_{t,k}^{\dagger})^++\lambda_k, Q_{\max}\} & k=d_t^{\Pi}, \\ \min\{Q_{t,k}^{\dagger}+\lambda_k, Q_{\max}\} & k\neq d_t^{\Pi}, \end{cases} \quad (13)$$

where $Q_{1,k}^{\dagger}=Q_{1,k}$ and $D_{t,k}^{\dagger} \triangleq \lfloor WR_{t,k}^{\dagger}(P_{\mathrm{BSL}})N_{\mathrm{F}}/B \rfloor, \ \forall t,k\in\mathcal{K}$. By applying (12) and (13) iteratively, the UE selection of the base policy can be determined. Hence, the base policy can be summarized as following.

**Policy 1** (Base Policy $\Pi$). *The transmission power to each selected UE for base policy is fixed, i.e., $P_t = P_{\mathrm{BSL}}, \forall t$. Moreover, the UE selection is determined by applying (12) and (13) iteratively, which is denoted as $\{d_1^{\Pi}, d_2^{\Pi}, \ldots, d_T^{\Pi}\}$.*

Although the base policy is obtained by approximating queue dynamics with averaging, the approximate value function should be evaluated with actual distribution of packet arrivals and departures. Hence, the approximate value function, which approximately measures the average system cost for all the UEs from the $(t+1)$-th frame, can be written as

$$W_t^{\Pi}(\mathcal{Q}_t^{\mathrm{D}}) = (T-t)w_{\mathrm{P}}P_{\mathrm{BSL}} + \sum_{k\in\mathcal{K}} W_{t,k}^{\Pi}(Q_{t,k}^{\mathrm{D}}), \quad (14)$$

where the *local value function* $W_{t,k}^{\Pi}(Q_{t,k}^{\mathrm{D}})$ is defined as the average queuing cost raised by the $k$-th UE from the $(t+1)$-th frame to the end of the current scheduling period given the base policy $\Pi$ and its *local post-decision QSI $Q_{t,k}^{\mathrm{D}}$*. Thus,

$$W_{t,k}^{\Pi}(Q_{t,k}^{\mathrm{D}}) \triangleq \mathbb{E}_{\mathcal{A},\mathcal{Y}}^{\Pi}\Big[\sum_{\tau=t+1}^{T}\big(Q_{\tau,k}+w_{\mathrm{Q}}\mathbb{I}[Q_{\tau,k}=Q_{\max}]\big) +Q_{T+1,k}\big|Q_{t,k}^{\mathrm{D}}\Big], \ \forall k\in\mathcal{K}. \quad (15)$$

In order to derive the analytical expression of $W_{t,k}^{\Pi}(Q_{t,k}^{\mathrm{D}})$, denote the number of departure packets from the $k$-th queue in the $t$-th frame under the base policy as $D_{t,k}^{\Pi}$ if the $k$-th UE is selected, then we have the following lemma.

**Lemma 2.** *With Scheme 1, the probability mass function (PMF) of $D_{t,k}^{\Pi}$ is given by $\Pr[D_{t,k}^{\Pi} = n] = F_{Y_{t,k}}\Big[\big(2^{\frac{(n+1)B}{WT_{\mathrm{F}}}}-1\big)\frac{N_0 W}{P_{\mathrm{BSL}}}\Big] - F_{Y_{t,k}}\Big[\big(2^{\frac{nB}{WT_{\mathrm{F}}}}-1\big)\frac{N_0 W}{P_{\mathrm{BSL}}}\Big]$.*

*Proof.* The proof is straightforward based on the definition of $D_{t,k}$ and $R_{t,k}$. $\square$

Hence, the local approximate value function $W_{t,k}^{\Pi}(Q_{t,k}^{\mathrm{D}})$ ($\forall t,k$) is derived in the following lemma.

**Lemma 3** (Analytical Expression of $W_{t,k}^{\Pi}(Q_{t,k}^{\mathrm{D}})$). *Let $\mathbf{u}_{t,k}, \mathbf{s}_{t,\tau,k}, \mathbf{c}^{(1)}, \mathbf{c}^{(2)} \in \mathbb{R}^{(Q_{\max}+1)\times 1}$, which are defined by $\mathbf{u}_{t,k} \triangleq \mathbf{1}_{Q_{t,k}^{\mathrm{D}}+1}, \mathbf{s}_{t,\tau,k} \triangleq \mathbf{1}_{Q_{\tau,k}+1}, [\mathbf{c}^{(1)}]_i \triangleq i-1+w_{\mathrm{Q}}\mathbb{I}[i=Q_{\max}+1]$, and $[\mathbf{c}^{(2)}]_i \triangleq i-1$, respectively. $\mathbf{1}_i$ denotes the column vector whose*

$i$-*th element is* 1 *and other elements are* 0. *Let $\mathbf{P}_k, \mathbf{M}_{t,k} \in \mathbb{R}^{(Q_{\max}+1)\times(Q_{\max}+1)}$, whose entries are given by (17) and (18), respectively. Then $W_{t,k}^{\Pi}(Q_{t,k}^{\mathrm{D}})$ can be represented by*

$$W_{t,k}^{\Pi}(Q_{t,k}^{\mathrm{D}}) = \sum_{\tau=t+1}^{T}\mathbf{s}_{t,\tau,k}^{\mathsf{T}}\mathbf{c}^{(1)} + \mathbf{s}_{t,T+1,k}^{\mathsf{T}}\mathbf{c}^{(2)}, \quad (16)$$

*where $\mathbf{s}_{t,\tau,k} = \mathbf{X}_k^{\mathsf{T}}(t,\tau)\mathbf{P}_k^{\mathsf{T}}\mathbf{u}_{t,k}$, and $\mathbf{X}_k(t,\tau) = \prod_{n=t+1}^{\tau-1}\mathbf{M}_{n,k}^{\mathbb{I}(d_n^{\Pi}=k)}\mathbf{P}_k^{\mathbb{I}(d_n^{\Pi}\neq k)}, \tau = t+1, \ldots, T$.*

*Proof.* Please refer to Appendix B. $\square$

### B. Scheduling with Approximate Value Function

In this part, we use approximate value functions $\{W_t^{\Pi}(\mathcal{Q}_t^{\mathrm{D}})|\forall t, \mathcal{Q}_t^{\mathrm{D}}\}$ to approximate optimal value functions $\{W_t(\mathcal{Q}_t^{\mathrm{D}})|\forall t, \mathcal{Q}_t^{\mathrm{D}}\}$. Because the approximate value function is analytically expressed, conventional value iteration to evaluate the value function can be avoided, which can significantly reduce the computational complexity. With the approximate value function, the scheduling actions in the $t$-th ($\forall t$) frame could be obtained by solving the following problem.

**P2** (One-Step Policy Iteration):

$$\Psi_t(\mathcal{S}_t) \triangleq (d_t^{\Psi}, P_t^{\Psi})$$
$$=\arg\min_{\Omega_t(\mathcal{S}_t)} \big\{ g_t\big(\mathcal{S}_t, \Omega_t(\mathcal{S}_t)\big)+W_t^{\Pi}\big(\mathcal{Q}_t^{\mathrm{D}}(\mathcal{S}_t, \Omega_t)\big) \big\},$$
$$\text{s.t.} \quad 0 \leq P_t \leq P_{\max}.$$

Since both the current system cost and approximate value function in the $t$-th frame can be decomposed by each UE, P2 can be decomposed into $K$ sub-problems. The $k$-th sub-problem is given by

**P2(k) :** $P_{t,k}^{\psi} =\arg\min_{P_{t,k}} \big\{ w_p P_{t,k} + W_t^{\Pi}\big(\mathcal{Q}_t^{\mathrm{D}}(\mathcal{S}_t, \pi_{t,k})\big) \big\},$
$$\text{s.t.} \quad 0 \leq P_{t,k} \leq P_{\max},$$

where $\pi_{t,k} \triangleq (k, P_{t,k})$ denotes the policy that the $k$-th UE is selected. Moreover, denote the minimized objective of P2(k) as $G_{t,k}^{\psi}$, then the optimal solution of P2 can be derived by $d_t^{\Psi} =\arg\min_{k\in\mathcal{K}} G_{t,k}^{\psi}$ and $P_t^{\Psi} = P_{t,d_t^{\Psi}}^{\psi}$.

The transmission power allocation for P2(k), which is a discrete optimization problem, can be achieved via the following one-dimensional search.

**Lemma 4** (Local Power Optimization). *The optimized transmission power for the $k$-th UE (solution of P2(k)) is*

$$P_{t,k}^{\psi} = \arg\min_{P_{t,k}\in\mathscr{P}_{t,k}} \big\{ w_{\mathrm{P}}P_{t,k} + \Delta\mathbf{z}_{t,k}^{\mathsf{T}}(P_{t,k})\mathbf{V}_{t,k} \big\}, \quad (19)$$

*where $\mathscr{P}_{t,k} \triangleq \Big\{0, \frac{2^{\frac{B}{WN_{\mathrm{F}}}}-1}{Y_{t,k}}, \frac{2^{\frac{2B}{WN_{\mathrm{F}}}}-1}{Y_{t,k}}, \ldots, \min\big(\frac{2^{\frac{Q_{\max}B}{WN_{\mathrm{F}}}}-1}{Y_{t,k}}, P_{\max}\big)\Big\}$ is the feasible power set, $\Delta\mathbf{z}_{t,k}(P_{t,k}) = \mathbf{1}_{Q_{t,k}^{\mathrm{D}}(P_{t,k})+1} - \mathbf{1}_{Q_{t,k}+1}, Q_{t,k}^{\mathrm{D}}(P_{t,k}) = \big(Q_{t,k}-\lfloor\frac{WR_{t,k}(P_{t,k})T_{\mathrm{F}}}{B}\rfloor\big)^+, \mathbf{V}_{t,k} = \sum_{\tau=t+1}^{T}\mathbf{Y}_{\tau,k}^{\mathsf{T}}\mathbf{c}^{(1)}+\sum_{\tau=t+2}^{T+1}\mathbf{Y}_{\tau,k}^{\mathsf{T}}\mathbf{c}^{(2)}, and*

$$\mathbf{Y}_{\tau,k} = \begin{cases} \mathbf{P}_k^{\mathsf{T}} & \tau = t+1, \\ \mathbf{X}_k^{\mathsf{T}}(t+1,\tau)\mathbf{Y}_{t+1,k} & \tau = t+2, \ldots, T. \end{cases} \quad (20)$$

*Proof.* The feasible power set $\mathscr{P}_{t,k}$ is the minimum required transmission power to transmit an integer number of packets, thus it will not affect the optimality. Then P2(k) can be solved by one-dimensional search in $\mathscr{P}_{t,k}$. $\Delta\mathbf{z}_{t,k}^{\mathsf{T}}(P_{t,k})\mathbf{V}_{t,k} = W_{t,k}^{\Pi}(Q_{t,k}^{\mathrm{D}}(P_{t,k}))-W_{t,k}^{\Pi}(Q_{t,k})$, where $W_{t,k}^{\Pi}(Q_{t,k})$ is a constant in P2(k). Thus, the proof is straightforward. $\square$

$$[\mathbf{P}_k]_{i,j} = \begin{cases} \Pr[A_{t,k} = j - i] & 1 \le i \le Q_{\max} \text{ and } i \le j \le Q_{\max}, \\ \Pr[A_{t,k} \ge Q_{\max} + 1 - i] & 1 \le i \le Q_{\max} \text{ and } j = Q_{\max} + 1, \\ 1 & i = Q_{\max} + 1 \text{ and } j = Q_{\max} + 1, \\ 0 & \text{otherwise} \end{cases} \tag{17}$$

$$[\mathbf{M}_{t,k}]_{i,j} = \begin{cases} \Pr[D_{t,k}^\Pi \ge i - 1] \Pr[A_{t,k} = 0] & 1 \le i \le Q_{\max} + 1 \text{ and } j = 1, \\ \Pr[A_{t,k} - \min(D_{t,k}^\Pi, i-1) = j - i] & 1 \le i \le Q_{\max} + 1 \text{ and } 2 \le j \le Q_{\max}, \\ \Pr[A_{t,k} - \min(D_{t,k}^\Pi, i-1) \ge Q_{\max} + 1 - i] & 1 \le i \le Q_{\max} + 1 \text{ and } j = Q_{\max} + 1 \end{cases} \tag{18}$$

As a summary, the system can be scheduled in a semi-distributed manner as elaborated below.

- **Step 1:** At the beginning of one scheduling period, each UE (say the $k$-th UE) calculates $\{\mathbf{V}_{t,k} | \forall t\}$ locally based on SCSI $\mathcal{I}_t^{\text{SCSI}}$ and the angular velocity $\omega_k$.
- **Step 2:** At the beginning of the $t$-th frame ($\forall t$), each UE (say the $k$-th UE) calculates the optimal power $P_{t,k}^\psi$ by assuming it is selected, and reports $G_{t,k}^\psi$ and $P_{t,k}^\psi$ to the BS via uplink signaling channel.
- **Step 3:** After receiving the reports from all UEs, the BS determines the improved scheduling policy $\Psi_t$.

## V. SIMULATIONS AND DISCUSSION

In this part, the performance of the proposed algorithm is demonstrated via numerical simulations with existing benchmark algorithms for comparison. There are eight UEs in the system where four UEs are static with zero angular velocities (indexed with $1 \sim 4$) while the other four UEs are rotating with angular velocity 2 rad/s (indexed with $5 \sim 8$). We compare the proposed algorithms with the following three benchmarks, which are referred to BM1, BM2 and BM3. For fair comparison, the fixed transmission power of the base policy and benchmarks are the same, i.e., $P_{\text{BSL}} = P_{\text{BM}}$.

**BM 1** (Dynamic Backpressure). *The transmission power to each selected UE is fixed to $P_{\text{BM}}$. The UE selection is based on the backpressure algorithm [21] according to the real-time data rate and queue length, i.e., $d_t = \arg\max_{k \in \mathcal{K}} R_{t,k}(P_{\text{BM}}) Q_{t,k}$.*

**BM 2** (Largest-Rate First). *The transmission power to each selected UE is fixed to $P_{\text{BM}}$. In each frame, the UE with the largest data rate is selected, i.e., $d_t = \arg\max_{k \in \mathcal{K}} R_{t,k}(P_{\text{BM}})$.*

**BM 3** (Longest-Queue First). *The transmission power to each selected UE is fixed to $P_{\text{BM}}$. In each frame, the UE with the longest queue is selected, i.e., $d_t = \arg\max_{k \in \mathcal{K}} Q_{t,k}$.*

The instantaneous SNR and queue length of a static UE (indexed by $k = 1$) and a rotating UE (indexed by $k = 5$) in one realization of scheduling period are illustrated in Fig. 2. The 5-th UE cannot find appropriate combiner due to UE rotation and limited FoV in the middle of the scheduling period, leading to weak SNRs. It can be observed that the proposed scheme can predict the low SNR period of rotating UEs and schedule more transmission opportunities to them before the low SNR period, so that the packet drop rate can be significantly reduced. As a comparison, the benchmarks suffer from high packet
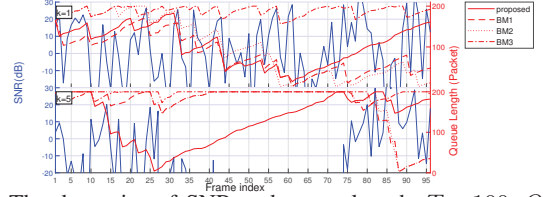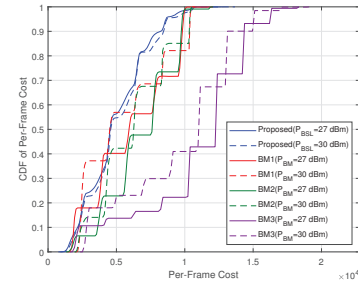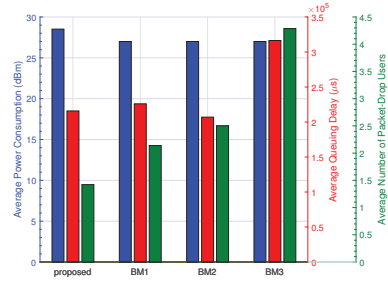


Fig. 2. The dynamics of SNR and queue length. $T = 100$, $Q_{\max} = 200$, $Q_{1,k} \sim \mathcal{U}(100, 200)$, $\lambda_k \sim \mathcal{U}(2, 6)$, $W = 400$ MHz, $N_F = 10$ ms, $B = 30000$ bytes, $N_R = 16$, $N_T = 32$, $\theta_{\max} = -\theta_{\min} = 90°$, $\phi_{\max} = -\phi_{\min} = 30°$, $P_{\max} = 30$ dBm, $N_k^{\text{cl}} = 3$, $N_{k,i}^{\text{ray}} = 20$, $\sigma_{\alpha;k,i}^2 \sim \mathcal{U}(4 \times 10^{-15}, 4 \times 10^{-14})$, $\sigma_{\phi;k,i} = \sigma_{\theta;k,i} = 5°$, $N_0 = -174$ dBm/Hz, $w_P = 1500$, $w_Q = 2000$, and $P_{\text{BSL}} = P_{\text{BM}} = 27$ dBm.



(a)



(b)

Fig. 3. (a) CDF of the per-frame cost. (b) Average transmission power, queuing delay and the number of packet-drop UEs.

drop rate during the low SNR period. This demonstrates the performance gain of the sensing-based channel prediction and the proposed predictive scheduling framework.

While Fig. 2 shows the system performance in a snapshot, Fig. 3(a) shows CDFs of per-frame cost of the proposed scheme as well as the benchmarks. It can be observed that the proposed algorithm has significantly better CDF curves than the benchmarks. More insights can be obtained from Fig. 3(b), where average transmission power, queuing delay, and the number of packet drop UEs are illustrated. BM1 has the medium cost of delay and packet drop penalty. This is because it makes UE selection according to both the queue length and

data rate. BM2 has the least cost of delay because it attempts to decrease the queue length as much as possible in every single frame. BM3 takes only the QSI into account but neglects the channel state information (CSI), which results in the worst performance. The proposed scheme manages to achieve the minimum packet drop rate, while keeping the average queuing delay in a low level. This demonstrates the benefits of channel prediction of the proposed scheme in suppressing the packet drop rate with non-stationary mmWave channel statistics.

## VI. CONCLUSION

In this paper, we consider the downlink transmission scheduling in an mmWave cell. Each UE is rotating with a predictable angular velocity for a number of frames, where the angular velocity of rotation can be measured by embedded motion sensors and reported to the BS. We first propose an SCSI-based beam alignment scheme. Then, we formulate the joint optimization of the downlink UE selection and power allocation as a finite-horizon MDP. To address the *curse of dimensionality*, we finally propose a novel approximate MDP approach via one-step policy improvement over a base policy. Simulations show that the proposed MDP solution framework can effectively exploit the motion sensors to predict the future performance, resulting in an efficient scheduling algorithm.

## APPENDIX A
## PROOF OF LEMMA 1

Note that $Y_{t,k} = |\mathbf{w}_{t,k}^{\mathsf{H}} \mathbf{H}_{t,k} \mathbf{f}_{t,k}|^2 = \Re^2(\mathbf{w}_{t,k}^{\mathsf{H}} \mathbf{H}_{t,k} \mathbf{f}_{t,k}) + \Im^2(\mathbf{w}_{t,k}^{\mathsf{H}} \mathbf{H}_{t,k} \mathbf{f}_{t,k})$. According to [22], when $N_{\mathrm{R}}$ and $N_{\mathrm{T}}$ are both sufficiently large, we have $\left|\mathbf{a}_{\mathrm{R}}^{\mathsf{H}}(\phi_{q_t^\dagger}) \mathbf{a}_{\mathrm{R}}(\phi_{t,k,i,\ell})\right|^2 \to \{0,1\}$, $\left|\mathbf{a}_{\mathrm{T}}^{\mathsf{H}}(\theta_{p_t^\dagger}) \mathbf{a}_{\mathrm{T}}(\theta_{t,k,i,\ell})\right|^2 \to \{0,1\}$. Denote the set $\mathcal{I}_{t,k,i} \triangleq \left\{\ell \big| |\mathbf{a}_{\mathrm{R}}^{\mathsf{H}}(\phi_{q_t^\dagger}) \mathbf{a}_{\mathrm{R}}(\phi_{t,k,i,\ell})|^2 \to 1, |\mathbf{a}_{\mathrm{T}}^{\mathsf{H}}(\theta_{p_t^\dagger}) \mathbf{a}_{\mathrm{T}}(\theta_{t,k,i,\ell})|^2 \to 1, \Lambda_{\mathrm{R}}(\phi_{t,k,i,\ell}) \Lambda_{\mathrm{T}}(\theta_{t,k,i,\ell}) = 1\right\}$, then $\Phi_{t,k,i} \triangleq |\mathcal{I}_{t,k,i}|$ will follow the binomial distribution. Conditioned on $\{\Phi_{t,k,i}|\forall i\}$, the real and imaginary parts of $\mathbf{w}_{t,k}^{\mathsf{H}} \mathbf{H}_{t,k} \mathbf{f}_{t,k}$ will follow normal distributions, e.g., $\Re\left(\mathbf{w}_{t,k}^{\mathsf{H}} \mathbf{H}_{t,k} \mathbf{f}_{t,k}\right)\big|_{\{\Phi_{t,k,i}|\forall i\}} \sim \mathcal{N}\left(0, \frac{1}{2} \sum_{i=1}^{N_k^{\mathrm{cl}}} \Phi_{t,k,i} \sigma_{\alpha;k,i}^2\right)$. Remind that if $\chi_2^2$ is a random variable following chi-squared distribution with degrees of freedom 2, then the CDF of $\chi_2^2$ is $F_{\chi_2^2}(x) = 1 - \exp(-x/2)$, $x > 0$. Hence, Lemma 1 is straightforward.

## APPENDIX B
## PROOF OF LEMMA 3

$\mathbf{u}_{t,k}$ and $\mathbf{s}_{t,\tau,k}$ represent the post-decision and pre-decision probability vector for the $k$-th queue respectively. $[\mathbf{c}^{(1)}]_i$ and $[\mathbf{c}^{(2)}]_i$ represent the per-frame queuing and packet-drop cost for the $k$-th UE in the $\tau$-th frame for cases $t+1 \le \tau \le T$ and $\tau = T+1$, respectively. $\mathbf{M}_{t,k}$ and $\mathbf{P}_k$ are transition probability matrices for the $k$-th queue considering both packet departure and arrival and only the packet arrivals, respectively. In (16), $\mathbf{s}_{t,\tau,k}^{\mathsf{T}} \mathbf{c}^{(1)}$ and $\mathbf{s}_{t,T+1,k}^{\mathsf{T}} \mathbf{c}^{(2)}$ counts for the average queuing and packet-drop cost in the $\tau$-th frame for cases $t+1 \le \tau \le T$ and $\tau = T+1$, respectively. Hence, Lemma 3 is straightforward.

## REFERENCES

[1] J. G. Andrews, S. Buzzi, W. Choi, S. V. Hanly, A. Lozano, A. C. K. Soong, and J. C. Zhang, "What will 5G be?" *IEEE J. Sel. Areas Commun.*, vol. 32, no. 6, pp. 1065–1082, 2014.

[2] J. Bao, D. Sun, and H. Li, "Motion sensor aided beam tracking in mobile devices of millimeter-wave communications," in *2018 IEEE International Conference on Communications (ICC)*, 2018, pp. 1–7.

[3] J. Cui, Y. Liu, Z. Ding, P. Fan, and A. Nallanathan, "Optimal user scheduling and power allocation for millimeter wave NOMA systems," *IEEE Trans. Wireless Commun.*, vol. 17, no. 3, pp. 1502–1517, 2018.

[4] W. Zhang, Y. Wei, S. Wu, W. Meng, and W. Xiang, "Joint beam and resource allocation in 5G mmWave small cell systems," *IEEE Trans. Veh. Technol.*, vol. 68, no. 10, pp. 10 272–10 277, 2019.

[5] W. Hao, M. Zeng, G. Sun, O. Muta, O. A. Dobre, S. Yang, and H. Gacanin, "Codebook-based max-min energy-efficient resource allocation for uplink mmWave MIMO-NOMA systems," *IEEE Trans. Commun.*, vol. 67, no. 12, pp. 8303–8314, 2019.

[6] B. Soleimani and M. Sabbaghian, "Cluster-based resource allocation and user association in mmWave femtocell networks," *IEEE Trans. Commun.*, vol. 68, no. 3, pp. 1746–1759, 2020.

[7] H. Zhang, S. Huang, C. Jiang, K. Long, V. C. M. Leung, and H. V. Poor, "Energy efficient user association and power allocation in millimeter-wave-based ultra dense networks with energy harvesting base stations," *IEEE J. Sel. Areas Commun.*, vol. 35, no. 9, pp. 1936–1947, 2017.

[8] M. Hussain, M. Scalabrin, M. Rossi, and N. Michelusi, "Mobility and blockage-aware communications in millimeter-wave vehicular networks," *IEEE Trans. Veh. Technol.*, vol. 69, no. 11, pp. 13 072–13 086, 2020.

[9] F. Liu, W. Yuan, C. Masouros, and J. Yuan, "Radar-assisted predictive beamforming for vehicular links: Communication served by sensing," *IEEE Trans. Wireless Commun.*, vol. 19, no. 11, pp. 7704–7719, 2020.

[10] W. Yuan, F. Liu, C. Masouros, J. Yuan, D. W. K. Ng, and N. González-Prelcic, "Bayesian predictive beamforming for vehicular networks: A low-overhead joint radar-communication approach," *IEEE Trans. Wireless Commun.*, vol. 20, no. 3, pp. 1442–1456, 2021.

[11] D. Zhang, A. Li, M. Shirvanimoghaddam, Y. Li, and B. Vucetic, "Exploring AoA/AoD dynamics in beam alignment of mobile millimeter wave MIMO systems," *IEEE Trans. Veh. Technol.*, vol. 68, no. 6, pp. 6172–6176, 2019.

[12] Y. Lin, C. Shen, and Z. Zhong, "Sensor-aided predictive beam tracking for mmWave phased array antennas," in *2019 IEEE Globecom Workshops (GC Wkshps)*, 2019, pp. 1–5.

[13] S. M. LaValle, A. Yershova, M. Katsev, and M. Antonov, "Head tracking for the Oculus Rift," in *2014 IEEE International Conference on Robotics and Automation (ICRA)*, 2014, pp. 187–194.

[14] W. Zhang, W. Zhang, and J. Wu, "UAV beam alignment for highly mobile millimeter wave communications," *IEEE Trans. Veh. Technol.*, vol. 69, no. 8, pp. 8577–8585, 2020.

[15] F. Liu, Y. Cui, C. Masouros, J. Xu, T. X. Han, Y. C. Eldar, and S. Buzzi, "Integrated sensing and communications: Toward dual-functional wireless networks for 6G and beyond," *IEEE J. Sel. Areas Commun.*, vol. 40, no. 6, pp. 1728–1767, 2022.

[16] I. Mavromatis, A. Tassi, R. J. Piechocki, and A. Nix, "Beam alignment for millimetre wave links with motion prediction of autonomous vehicles," in *Antennas, Propagation RF Technology for Transport and Autonomous Platforms 2017*, 2017, pp. 1–8.

[17] A. Forenza, D. J. Love, and R. W. Heath, "Simplified spatial correlation models for clustered MIMO channels with different array configurations," *IEEE Trans. Veh. Technol.*, vol. 56, no. 4, pp. 1924–1934, 2007.

[18] Y.-P. Lin, "Hybrid MIMO-OFDM beamforming for wideband mmWave channels without instantaneous feedback," *IEEE Trans. Signal Process.*, vol. 66, no. 19, pp. 5142–5151, 2018.

[19] B. Lv, L. Huang, and R. Wang, "Joint downlink scheduling for file placement and delivery in cache-assisted wireless networks with finite file lifetime," *IEEE Trans. Commun.*, vol. 67, no. 6, pp. 4177–4192, 2019.

[20] W. B. Powell, *Approximate Dynamic Programming: Solving the curses of dimensionality*. John Wiley & Sons, 2007, vol. 703.

[21] L. Georgiadis, M. J. Neely, and L. Tassiulas, *Resource allocation and cross-layer control in wireless networks*. Now Publishers Inc, 2006.

[22] A. Sayeed, "Deconstructing multiantenna fading channels," *IEEE Trans. Signal Process.*, vol. 50, no. 10, pp. 2563–2579, 2002.