# COSMO: Dynamic Uploading Scheduling in mmWave-Based Sensor Networks with Mobile Blockers

YIFEI SUN, Electrical and Electronic Engineering, Southern University of Science and Technology, Shenzhen, China and Computer Science, The University of Hong Kong, Hong Kong, China
BOJIE LV, Electrical and Electronic Engineering, Southern University of Science and Technology, Shenzhen, China
HAISHENG TAN, LINKE Lab, University of Science and Technology of China, Hefei, China
RUI WANG, Electrical and Electronic Engineering, Southern University of Science and Technology, Shenzhen, China
FRANCIS LAU, Computer Science, The University of Hong Kong, Hong Kong, China

Wireless sensor networks (WSNs) leveraging millimeter wave (mmWave) communication for bandwidth-demanding applications is considered in this article. Despite the large bandwidth, the delivery of delay-sensitive information collected by sensors may still face significant latency due to the vulnerability to intermittent link blockage. Hence, the guarantee of low age of information (AoI) in mmWave WSNs is not straightforward. In this article, the wireless sensing and dynamic programming techniques are jointly exploited to relieve the above issue. The former tracks the human blockers and predicts the chance of link blockage; the latter optimizes the transmission of multiple sensors based on the prediction. Particularly, the long-term optimization of sampling, uplink time and power allocation policies in a sensor network can be formulated as an infinite-horizon Markov decision process (MDP) with discounted cost, where the state transition probabilities can be predicted via wireless sensing. A novel low-complexity solution framework, namely COSMO, with a guaranteed performance in the worst case, is proposed. Simulations show that compared with heuristic benchmarks, benefiting from the prediction of the link blockage, COSMO can significantly suppress the average system cost, which consists of both AoI and energy consumption.

Yifei Sun also with The University of Hong Kong.

Part of this work has been accepted in the 29th IEEE International Conference on Parallel and Distributed Systems (ICPADS'23) [30]. This work is supported in part by the 2030 National Key AI Program of China under Grant 2021ZD0110400, in part by National Natural Science Foundation of China under Grant 62171213, and in part by High Level of Special Funds under Grant G03034K004.

Authors' Contact Information: Yifei Sun, Electrical and Electronic Engineering, Southern University of Science and Technology, Shenzhen, Guangdong, China and Computer Science, The University of Hong Kong, Hong Kong, China; e-mail: yfsun2@connect.hku.hk; Bojie Lv (Corresponding author), Electrical and Electronic Engineering, Southern University of Science and Technology, Shenzhen, Guangdong, China; e-mail: lyubj@mail.sustech.edu.cn; Haisheng Tan, LINKE Lab, University of Science and Technology of China, Hefei, Anhui, China; e-mail: hstan@ustc.edu.cn; Rui Wang (Corresponding author), Electrical and Electronic Engineering, Southern University of Science and Technology, Shenzhen, Guangdong, China; e-mail: wang.r@sustech.edu.cn; Francis Lau, Computer Science, The University of Hong Kong, Hong Kong, China; e-mail: fcmlau@cs.hku.hk.



ACM 1550-4859/2024/11-ART119
https://doi.org/10.1145/3696790

CCS Concepts: • **Networks → Wireless local area networks**; **Packet scheduling; • Mathematics of computing → Mathematical optimization**;

Additional Key Words and Phrases: Markov decision process, millimeter wave, age of information, wireless sensor networks, reinforcement learning, scheduling

**ACM Reference Format:**
Yifei Sun, Bojie Lv, Haisheng Tan, Rui Wang, and Francis Lau. 2024. COSMO: Dynamic Uploading Scheduling in mmWave-Based Sensor Networks with Mobile Blockers. *ACM Trans. Sensor Netw.* 20, 6, Article 119 (November 2024), 27 pages. https://doi.org/10.1145/3696790

## 1 Introduction

In **wireless sensor networks (WSNs)**, collecting sensing samples from the sensors (e.g., camera snapshots or broadband radar echos) and maintaining its timeliness at the server are considered as a challenge in many real-time and data-intensive applications. With the ever-increasing dense sensor deployment and throughput demand, communication in the **millimeter wave (mmWave)** band is expected in WSNs due to its wide bandwidth. However, because of the sparse propagation paths, the pencil-shaped beams for overcoming the extremely high path loss, and the weak capability of diffraction, mmWave communications may suffer from the link blockage issue in dynamic environments, which might not be severe in the traditional sub-6GHz communication systems. This will cause severe channel fluctuation, making the delivery of sensing samples outdated [23].

To quantify the freshness of the sensing samples, the concept of **age of information (AoI)** has been adopted as a metric [13], which is defined as the time elapsed since the status of the sensing target is sampled at the sensor. In the data-intensive applications of WSNs, mmWave communication could be adopted for sensing samples' uploading from sensors to the server [25, 36], such that the uploading latency and AoI at the server could be suppressed. However, the link blockage due to mobile blockers in the environment (e.g., human bodies) has become one of the open issues in mmWave communications [6]. Particularly, the blockage of **line-of-sight** (**LoS**) path in mmWave communications would result in significant **signal-to-noise ratio** (**SNR**) degradation, and hence large AoI or even outdated sensing samples at the server. For example, when a worker is patrolling an automatic assembly workshop, where most assembly tasks can be monitored by cameras, he may intermittently block the uplink mmWave transmission of photos or videos from the cameras to the monitoring server, leading to packet loss and retransmission. In fact, with the technique of wireless sensing, the layout of the communication environment can be reconstructed in advance [22, 34], and the locations of the mobile blockers can be tracked in real-time [28], such that potential link degradation due to blockage can be predicted. Intuitively, it would be helpful to suppress the AoI degradation if the scheduler could predict a link blockage of one sensor and raise its priority on sensing sample uploading in advance. To date, however, it is still unknown how to incorporate the above insight into a resource optimization framework, or how to design a joint sampling and uploading policy for mmWave WSNs benefiting from wireless sensing technique.

In this article, we would like to shed some light on the above open problems and propose **COSMO**, **a dynamiC uplOading Scheduling scheme with Mobile blOckers**, to tackle the joint sampling and uploading policy optimization in a mmWave-based sensor network with the awareness of the communication environment. Specifically, the sensors obtain the sensing samples, and deliver them to a server connected to a **base station (BS)** via uplink. There are some walking persons in the communication environment. They are moving randomly and might block some signal paths and degrade the uplink channel. The wireless sensing technique is adopted, such that the static locations of sensors, signal scattering clusters and BS, the real-time positions and random mobility patterns of walking persons are known to the scheduler. Hence, the joint scheduling

of sensors' sampling and uplink transmission in a sequence of frames can be formulated as an infinite-horizon **Markov decision process (MDP)**. The discounted summation of average AoI at the server and average energy consumption in sampling and uplink transmission is considered as the minimization objective. Then, a low-complexity solution framework, namely COSMO, is proposed to address the above MDP. The main contributions of this work are summarized as follows:

- —To the best of our knowledge, it is the first work systemically investigating the AoI-oriented scheduling design with the assistance of mmWave link blockage prediction.
- —In the proposed solution framework, the value function is approximated via analytical expressions. Thus, the computation-intensive value iteration can be eliminated, and the overall computation complexity can be significantly suppressed.
- —A non-trivial analytical lower bound on the performance is derived for COSMO. Notice that it is usually difficult to analyze the performance of AMDP algorithms.

It is shown by numerical simulations that compared with benchmarks without link blockage prediction, COSMO can reduce the average cost up to 52.7%, which demonstrates the necessity of wireless sensing in AoI-oriented mmWave sensor networks.

The remainder of this article is organized as follows: The related works are discussed in Section 2. The system model is introduced in Section 3. In Section 4, the problem of dynamic sampling and uploading scheduling is formulated as an infinite-horizon MDP. In Section 5, a low-complexity suboptimal scheduling framework, namely COSMO, is proposed with a guarantee on the worst-case performance. Finally, the performance of the proposed COSMO scheme is verified by comparing it with the benchmarks in Section 6, and the conclusion is drawn in Section 8.

We use the following notations throughout this article. Bold lowercase $\mathbf{a}$ denotes a column vector, bold uppercase $\mathbf{A}$ denotes a matrix, non-bold letters $a$ and $A$ denote scalar values, and letters $\mathcal{A}$ and $\mathbb{A}$ denote sets. $(a)^+$ denotes $\max(0, a)$. The operator $\otimes$ denotes a Kronecker product. $[\mathbf{A}]_{i,j}$, $\mathbf{A}^{\mathsf{T}}$, and $\mathbf{A}^{\mathsf{H}}$ denote the $(i,j)$-th element, transpose, and conjugate transpose, respectively. $\mathcal{CN}(m, R)$ denotes complex Gaussian distribution with mean $m$ and variance $R$. $\mathcal{U}[a, b]$ denotes a uniform distribution over the interval $[a, b]$. $\mathbb{E}[\cdot]$ denotes an expectation operator. $\mathbb{I}[\cdot]$ denotes an indicator function. Moreover, the main notations used in this article are listed in Table 1

## 2 Related Works

### 2.1 AoI-Oriented Dynamic Programming

There have been a number of research works on AoI-oriented scheduling design in a variety of wireless networks with dynamic programming [9, 16, 26, 35, 39–42]. For example in [35], the minimization of the long-term average AoI for a single sensor with an **energy harvesting** (**EH**) battery was considered, where the scheduling design was formulated as a continuous-time stochastic control problem. In [26], the problem of status update control in an energy-harvesting-enabled source via an mmWave link was formulated as an MDP. In addition to the single-sensor scenario, the works in [41, 42] extended the joint sampling and transmission design to the multi-sensor scenarios. The work in [41] minimized the summation of average AoIs of multiple sensors under the sampling and transmission energy constraints via an MDP formulation. This work was further extended in [42] by considering transmission failures and non-uniform sample size. In [9], the age-aware computation offloading problem in **mobile edge computing** (**MEC**) systems was formulated as a **constrained MDP** (**CMDP**), such that long-term AoI could be minimized with limited energy consumption and wireless bandwidth. In [16], a learning-augmented online algorithm was proposed to minimize AoI and transmission costs for the design of **transmission control protocol** (**TCP**) acknowledgment. However, the algorithms proposed in many of the existing works,

Table 1. Main Notations

| Symbol | Description |
|---|---|
| **Network** | |
| $K/\mathcal{K}$ | Number/Set of sensors |
| $B/\mathcal{B}$ | Number/Set of mobile blockers |
| $N_{\mathrm{R}}/N_{\mathrm{T}}$ | Number of receive/transmit antenna elements |
| $L/\mathbb{L}$ | Number/Set of possible locations for mobile blockers |
| **Channel** | |
| $\mathbf{H}_{t,k}$ | Channel matrix of $k$th sensor in $t$th frame |
| $\mathbb{I}^{\mathrm{o}}_{t,k,i}$ | Blockage indicator of $(k,i)$-th path in $t$th frame |
| $\alpha_{t,k,i}$ | Complex gain of $(k,i)$-th path in $t$th frame |
| $\mathbf{a}_{\mathrm{R}}/\mathbf{a}_{\mathrm{T}}$ | Array response vector of the ULAs at the sensors/the BS |
| $\Phi_{k,i}/\Theta_{k,i}$ | Sine of AoA/AoD of $(k,i)$-th path |
| $\mathbf{w}_{t,k}/\mathbf{f}_{t,k}$ | Analog combiner/precoder of the BS/$k$th sensor |
| **State** | |
| $l^{\mathrm{o}}_{t,b}$ | Location index of the $b$th mobile blocker in $t$th frame |
| $Y_{t,k}$ | Baseband channel power gain of $k$th sensor in $t$th frame |
| $Q_{t,k}$ | Queue length of $k$th sensor in $t$th frame |
| $A^{\mathrm{s}}_{t,k}$ | AoI at $k$th sensor in $t$th frame |
| $A^{\mathrm{d}}_{t,k}$ | AoI for $k$th sensor at the server in $t$th frame |
| **Action** | |
| $s_{t,k}$ | Sampling action for $k$th sensor in $t$th frame |
| $\tau_{t,k}$ | Transmission time allocated to $k$th sensor in $t$th frame |
| $p_{t,k}$ | Transmission power of $k$th sensor in $t$th frame |

such as the Q-learning algorithm used in [9, 26] and the policy iteration algorithm used in [40], are prohibitive in computation complexity for multi-sensor scenarios, as it grows exponentially with respect to the number of sensors. Although the linear approximation of value functions adopted in [41, 42] can reduce the computation complexity, it is difficult to evaluate their performance analytically. As a result, how to design an **approximate MDP (AMDP)** for AoI-oriented scheduling design with a non-trivial performance lower bound and low computational complexity is still unanswered.

### 2.2 Blockage-Aware mmWave Scheduling

There have been a number of works on the transmission scheduling with link blockage awareness in mmWave communication networks [5, 21, 31, 33]. For example, relay-assisted scheduling was proposed in mmWave small cells [33] and mmWave backhaul networks [21] to overcome link blockage. In [5], a blockage-aware proportional fair scheduling was proposed for mmWave communication systems with a four-state blockage model. In these works, the scheduling algorithms are triggered either after blockage happens [21, 33] or significant SNR degradation has been observed [5]. In fact, it is possible to predict the mmWave link degradation via in-band or out-of-band sensing, so that its impacts can be mitigated in advance. For example in [31], the impact of mmWave link degradation due to device rotation and limited **field-of-view** (**FoV**) was mitigated by channel prediction in an AMDP framework, where the average queuing delay was the minimization objective. However, to our best knowledge, there is no existing work on the AoI-oriented scheduling design exploiting sensing and channel prediction.

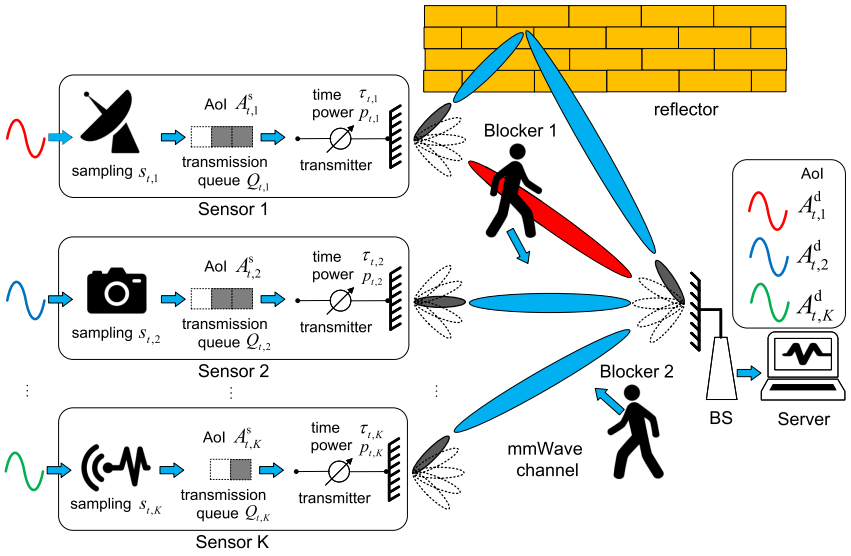


Fig. 1. Network model of the considered mmWave-based WSN.

## 3 System Model

As illustrated in Figure 1, we consider the scheduling of a monitoring system, which delivers the delay-sensitive sensing samples from distributive sensors to a server via high-speed mmWave uplink communications. The system consists of one server connected with the BS and $K$ sensors. Both the BS and sensors are randomly deployed and static. The set of sensors is denoted as $\mathcal{K} \triangleq \{1, 2, \ldots, K\}$. The sensors detect their targets (e.g., taking photos or detecting motions via radar waves), collect sensing samples, and deliver the sensing samples to the server via BS. We shall refer to the action of collecting sensing samples as sampling. The timeliness of sensing samples at the server depends on both policies of sampling and the uplink transmission of samples.

When a sampling action is made at a sensor, the delay in collecting sensing samples at the server comes from the uplink transmission. As shown in Figure 1, both **LoS** and **non-line-of-sight (NLoS)** signal propagation paths (e.g., scattering paths off the walls) may exist between each sensor and the BS. The locations of sensors, BS and scattering clusters are fixed, such that the **angle-of-arrival (AoA)** and **angle-of-departure (AoD)** of these paths are deterministic, which can be sensed in advance via communication signals as in [28]. The analog MIMO architecture with single **radio frequency (RF)** chain and a half-wavelength **uniform linear phased array (ULA)** is adopted at both the BS and the sensors, so that both transmit and receive beams can be aligned to the available paths. It is assumed that the linear phased arrays at the BS and sensors are with $N_{\mathrm{R}}$ and $N_{\mathrm{T}}$ antenna elements, respectively.

Due to the short carrier wavelength, the diffraction of mmWave communication signals across human bodies is usually negligible. Hence, walking persons in the network may block the propagation paths randomly as illustrated by the red path in Figure 1, which may fail the transmission and cause a large delay to the uploading of sensing samples. Note that the issue of link blockage is much more significant in the mmWave band than that in the sub-6GHz band, and the link quality fluctuation would degrade the delay-sensitive communication dramatically. In order to mitigate its impact, a sensing-based predictive scheduling framework is proposed for AoI optimization in mmWave communication systems.

Particularly, it is assumed that there are $B$ walking persons (blockers), denoted by the set $\mathcal{B} \triangleq \{1, 2, \ldots, B\}$. The locations of walking persons can be tracked as in [15, 38]. Therefore, the potential path blockage can be predicted, such that the uplink channel quality of the coming future can be foreseen. As a result, the sampling and uplink transmission can be adjusted to relieve the impact of path blockage. For example, if a LoS path blockage may happen soon, all the uplink transmission resources may be scheduled to the suffering sensor to speed up the sample uploading. In order to facilitate the prediction, the mobility of the $B$ human blockers is modeled as follows: The uplink transmission time is organized by physical-layer frames with a duration $T_{\mathrm{F}}$, where the **channel state information (CSI)** is assumed to be quasi-static in one frame. The two-dimensional space of the network coverage is quantized into a number of indexed grids. Let $\mathbb{L} \triangleq \{1, 2, \ldots, L\}$ be the set of location indexes. The location indexes of the BS and the $k$th sensors ($\forall k$) are denoted as $l^{\mathrm{BS}}$ and $l^{\mathrm{s}}_k$, respectively. The location index of the $b$th mobile human blocker in the $t$th frame is denoted as $l^{\mathrm{o}}_{t,b}$, $\forall t, b$. It is assumed that the mobility of the $b$th mobile blocker follows a time-invariant Markov chain with the following transition probabilities:

$$\Pr\left[l^{\mathrm{o}}_{t+1,b} = \ell' \middle| l^{\mathrm{o}}_{t,b} = \ell\right] = \left[\mathbf{P}^{\mathrm{o}}_b\right]_{\ell,\ell'}, \ \forall b, \ \forall t, \ \forall \ell, \ell' \in \mathbb{L}, \tag{1}$$

where $\mathbf{P}^{\mathrm{o}}_b \in \mathbb{R}^{L\times L}$ denotes the transition matrix of the $b$th blocker's mobility. Given the random motion of the human blockers, the channel model, uplink transmission model and the AoI model are elaborated in the following.

### 3.1 Channel and Uplink Transmission Models

In order to capture the impact of path blockage by human blockers, the geometric channel model [1, 3, 10] is adopted in this article. Specifically, there are at most $M$ NLoS paths and one LoS path from one sensor to the BS. Denote $\mathcal{M} = \{0, 1, 2, \ldots, M\}$ as the index set of propagation paths, where the index of the LoS path is 0 for notation convenience. Moreover, the $i$th path of the $k$th sensor is denoted as the $(k, i)$-th path. Hence, as in [1, 10], the channel matrix $\mathbf{H}_{t,k} \in \mathbb{C}^{N_{\mathrm{R}}\times N_{\mathrm{T}}}$ from the $k$th sensor to the BS in the $t$th frame can be written as[1]

$$\mathbf{H}_{t,k} = \sum_{i\in\mathcal{M}} \mathbb{I}^{\mathrm{o}}_{t,k,i}(\mathcal{L}^{\mathrm{o}}_t)\alpha_{t,k,i}\mathbf{a}_{\mathrm{R}}\left(\Phi_{k,i}\right)\mathbf{a}^{\mathsf{H}}_{\mathrm{T}}\left(\Theta_{k,i}\right), \tag{2}$$

where $\mathcal{L}^{\mathrm{o}}_t \triangleq \{l^{\mathrm{o}}_{t,b}\}_{b\in\mathcal{B}}$ is the aggregation of locations of blockers, and $\mathbb{I}^{\mathrm{o}}_{t,k,i}(\mathcal{L}^{\mathrm{o}}_t) \in \{0, 1\}$ is the blockage indicator. Thus, $\mathbb{I}^{\mathrm{o}}_{t,k,i}(\mathcal{L}^{\mathrm{o}}_t) = 0$ when $(k, i)$-th path is blocked, otherwise $\mathbb{I}^{\mathrm{o}}_{t,k,i}(\mathcal{L}^{\mathrm{o}}_t) = 1$. Moreover, $\Phi_{k,i} \triangleq \sin\phi_{k,i}$, $\Theta_{k,i} \triangleq \sin\theta_{k,i}$, $\phi_{k,i}$ and $\theta_{k,i}$ are the AoA and AoD of the $(k, i)$-th path, respectively. $\alpha_{t,k,i}$ denotes the complex gain of the $(k, i)$-th path in the $t$th frame obeying a complex Gaussian distribution with zero mean and variance $\rho^{-1}_{k,i}$, i.e., $\alpha_{t,k,i} \sim \mathcal{CN}(0, \rho^{-1}_{k,i})$. The path loss $\rho_{k,i}$ depends on the path length and scattering loss. $\mathbf{a}_{\mathrm{R}}(\Phi_{k,i})$ and $\mathbf{a}_{\mathrm{T}}(\Theta_{k,i})$ represent the normalized array response vectors of the ULAs at the sensors and the BS, respectively, which are expressed as in [1, 10] by

$$\mathbf{a}_{\mathrm{R}}(\Phi_{k,i}) \triangleq \frac{1}{\sqrt{N_{\mathrm{R}}}}\left[1, e^{-j\pi\Phi_{k,i}}, \ldots, e^{-j\pi(N_{\mathrm{R}}-1)\Phi_{k,i}}\right]^{\mathsf{T}}, \tag{3}$$

$$\mathbf{a}_{\mathrm{T}}(\Theta_{k,i}) \triangleq \frac{1}{\sqrt{N_{\mathrm{T}}}}\left[1, e^{-j\pi\Theta_{k,i}}, \ldots, e^{-j\pi(N_{\mathrm{T}}-1)\Theta_{k,i}}\right]^{\mathsf{T}}. \tag{4}$$

[1]We assume that the reflectors of mmWave signals in the environment (e.g., walls and big furniture) are of fixed quantity and quasi-static. This is because the dynamics of the reflectors (e.g., change of furniture positions) happen in the timescale of days or months, which is much longer than the timescale of communication scheduling.

All the human blockers are modeled as a disk with radius $r_{\mathrm{B}}$ in the two-dimensional network coverage. Hence, the indicator $\mathbb{I}^{\mathrm{o}}_{t,k,i}$ in (2) can be determined by comparing the blocker radius with the shortest distances between the blockers' centroids and the $i$th path. Particularly, let $\mathcal{P}_{k,i}$ be the set of connected line segments in the $(k,i)$-th path, and $\mathbb{D}(\ell, \mathcal{P}_{k,i})$ be the minimum distance from the location $\ell$ to the line segments in $\mathcal{P}_{k,i}$, then

$$\mathbb{I}^{\mathrm{o}}_{t,k,i}(\mathcal{L}^{\mathrm{o}}_t) = \prod_{b\in\mathcal{B}} \mathbb{I}\left[\mathbb{D}(l^{\mathrm{o}}_{t,b}, \mathcal{P}_{k,i}) \geq r_{\mathrm{B}}\right]. \tag{5}$$

Note that the path loss of the LoS path is usually much smaller than that of the NLoS paths, the uplink transmission suffers from significant degradation when the LoS path is blocked.

*Remark 1 (Non-Stationary Channel Model).* Notice that the channel matrix $\mathbf{H}_{t,k}$ is a superposition of all the propagation paths from the $k$th sensor to the BS. The random gain of each path, $\alpha_{t,k,i}$, is stationary versus time. However, the set of blocked paths depends on the mobility of the blockers. Due to the random mobility of blockers, the blocked paths could be different in different frames. Hence, the distributions of $\mathbf{H}_{t,k}$ in different frames are different, leading to a non-stationary channel model. Moreover, the prediction of future channel distribution relies on the prediction of blockers' future locations.

Let $\mathbf{w}_{t,k} \in \mathbb{C}^{N_{\mathrm{R}}\times 1}$ and $\mathbf{f}_{t,k} \in \mathbb{C}^{N_{\mathrm{T}}\times 1}$ be the analog combiner and precoder of the BS and the $k$th sensor when the $k$th sensor is transmitting in the $t$th frame, respectively. The uplink capacity of the $k$th sensor in the $t$th frame can be expressed as

$$R_{t,k} \triangleq W \log_2\left(1 + p_{t,k} \underbrace{\frac{\left|\mathbf{w}^{\mathrm{H}}_{t,k}\mathbf{H}_{t,k}\mathbf{f}_{t,k}\right|^2}{\|\mathbf{w}_{t,k}\|^2 N_0 W}}_{\text{Baseband gain } Y_{t,k}}\right), \tag{6}$$

where $Y_{t,k}(\mathcal{L}^{\mathrm{o}}_t) \triangleq \frac{|\mathbf{w}^{\mathrm{H}}_{t,k}\mathbf{H}_{t,k}\mathbf{f}_{t,k}|^2}{\|\mathbf{w}_{t,k}\|^2 N_0 W}$ denotes the baseband channel power gain (baseband gain for short), $p_{t,k}$ denotes the transmission power of the $k$th sensor, $N_0$ is the noise power spectral density, and $W$ is the bandwidth. Due to the hardware constraint, the maximum transmission power satisfies

$$p_{t,k} \leq P_{\max},\ \forall t, k \in \mathcal{K}. \tag{7}$$

The analog precoders and combiners are chosen from the following pre-defined codebook composed of a finite number of beam directions:

$$\mathbf{w}_{t,k} \in \mathcal{W} \triangleq \{\mathbf{a}_{\mathrm{R}}(\Phi_q) | q = 1, 2, \ldots, N_{\mathrm{R}}\}, \tag{8}$$

$$\mathbf{f}_{t,k} \in \mathcal{F} \triangleq \{\mathbf{a}_{\mathrm{T}}(\Theta_p) | p = 1, 2, \ldots, N_{\mathrm{T}}\},\ \forall t, k, \tag{9}$$

where $\Phi_q = \frac{2(q-1)}{N_{\mathrm{R}}} - 1$ and $\Theta_p = \frac{2(p-1)}{N_{\mathrm{T}}} - 1$. They might be chosen to maximize the instantaneous **signal-to-noise ratio (SNR)** according to the channel matrix $\mathbf{H}_{t,k}$. However, the overhead of estimating channel matrix with analog MIMO architecture is significant. Instead of channel matrix estimation, the time-varying blockers' locations, constant AoAs, AoDs and path loss of propagation paths can be sensed in advance as in [22, 34]. Hence, the analog precoder and combiner in each frame (say the $t$th frame) are designed to maximize the average SNR as follows:

$$\left(\mathbf{w}_{t,k}, \mathbf{f}_{t,k}\right) = \underset{\substack{\mathbf{a}_{\mathrm{R}}(\Phi_q)\in\mathcal{W} \\ \mathbf{a}_{\mathrm{T}}(\Theta_p)\in\mathcal{F}}}{\arg\max}\ \mathbb{E}_{\mathbf{H}_{t,k}}\left[\left|\mathbf{a}^{\mathrm{H}}_{\mathrm{R}}(\Phi_q)\mathbf{H}_{t,k}\mathbf{a}_{\mathrm{T}}(\Theta_p)\right|^2 \middle| \mathcal{L}^{\mathrm{o}}_t, \Phi_{k,i}, \Theta_{k,i}, \rho_{k,i}, \forall i\right], \tag{10}$$

where the expectation is taken with respect to the random path gains $\{\alpha_{t,k,i}|\forall i\}$. The two-dimensional search over the precoder and combiner codebooks and the method of sample average approximation can be used to find the optimal precoder and combiner pair. However, searching over all $N_{\mathrm{R}}N_{\mathrm{T}}$ precoder and combiner pairs by Monte-Carlo channel simulation is computation-intensive. Therefore, we shall provide a low-complexity solution for precoder and combiner selection in Lemma 1 of Section 5.1.

**Time-division multiple access (TDMA)** is adopted in each frame for multiple access. Let $\tau_{t,k}$ be the transmission time allocated to the $k$th sensor in the $t$th frame, the throughput of the $k$th sensor in the $t$th frame is $\tau_{t,k}R_{t,k}$, where the following constraints should be satisfied:

$$\sum_{k\in\mathcal{K}} \tau_{t,k} = T_{\mathrm{F}},\ \forall t, \tag{11}$$

$$0 \le \tau_{t,k} \le T_{\mathrm{F}},\ \forall t, k \in \mathcal{K}. \tag{12}$$

*Remark 2 (Channel Capacity Prediction).* Given the locations of human blockers in the current frame (say the $t$th frame) $\mathcal{L}_t^{\mathrm{o}}$, the possible locations of these blockers in the future frames (say the $m$th frame, $m > t$) $\mathcal{L}_m^{\mathrm{o}}$ can be predicted according to the mobility transition matrices $\{\mathbf{P}_b^{\mathrm{o}}\}_{b\in\mathcal{B}}$. For each possible value of $\mathcal{L}_m^{\mathrm{o}}$, the analog precoder and combiner for each sensor can be calculated according to (10), and finally, the distributions of channel capacities $\{R_{m,k}\}_{k\in\mathcal{K}}$ can be forecast according to (6). Thus, wireless sensing could enable the capacity prediction of mmWave channel with mobile blockers.

### 3.2 Queuing and AoI Models for Sample Uploading

The procedure of sensing sample collection is as follows: The server delivers a sensing decision to one sensor at the beginning of one frame, a sensing sample representing the latest target status is generated by the sensor and available for uploading since the same frame. Similar to the existing literature, we ignore the downlink signaling delay of sensing decisions and the sampling delay.

The sampling action will induce a constant sampling energy cost denoted by $C^{\mathrm{s}}$. As in [42], it is assumed that the data volume of each sample generated by the $k$th sensor consists of $Q_k^{\max}$ packets, each with $N_{\mathrm{b}}$ information bits. When a new data sample is generated at one sensor, the existing packets from the previous sampling in its uplink queue will be dropped, and new packets will be added. Let $s_{t,k} \in \{0, 1\}$ be the sampling action for the $k$th sensor in the $t$th frame. $s_{t,k} = 0$ indicates no sampling and transmission of the existing sample in the $t$th frame, and $s_{t,k} = 1$ indicates the transmission of the new sample in the $t$th frame. Let $Q_{t,k}$ be the length of the uplink transmission queue (number of uplink packets) of the $k$th sensor at the beginning of the $t$th frame, we have

$$Q_{t+1,k} = \begin{cases} (Q_{t,k} - D_{t,k})^+, & s_{t,k} = 0 \\ Q_k^{\max} - D_{t,k}, & s_{t,k} = 1, \end{cases} \tag{13}$$

where $D_{t,k} = \lfloor R_{t,k}\tau_{t,k}/N_{\mathrm{b}} \rfloor$ denotes the departure packet number of the $k$th sensor in the $t$th frame.

To characterize the freshness of samples, the AoI at the $k$th sensor in the $t$th frame, denoted by $A_{t,k}^{\mathrm{s}}$, is defined as the duration (in terms of frame number) from the generation time of the latest sample to the $t$th frame. Moreover, we introduce an AoI threshold $A_{\max}$ to characterize the

outdated samples.[2] Hence, the AoI dynamics at the $k$th sensor is given by

$$A^{\mathrm{s}}_{t+1,k} = \begin{cases} \min\left\{A^{\mathrm{s}}_{t,k}+1, A_{\max}\right\}, & s_{t,k}=0 \\ 1, & s_{t,k}=1. \end{cases} \tag{14}$$

Moreover, the AoI for the $k$th sensor at the server in the $t$th frame, denoted as $A^{\mathrm{d}}_{t,k}$, is defined as the duration (in terms of frame number) from the generation time of the latest received sample to the $t$th frame. Hence, $A^{\mathrm{d}}_{t,k}$ increases in each frame, unless a new sensing sample from the $k$th sensor is completely received or the AoI threshold $A_{\max}$ is reached. Thus,

$$A^{\mathrm{d}}_{t+1,k} = \begin{cases} \min\left\{A^{\mathrm{s}}_{t,k}+1, A_{\max}\right\}, & D_{t,k} \geq Q_{t,k} \\ \min\left\{A^{\mathrm{d}}_{t,k}+1, A_{\max}\right\}, & \text{otherwise}. \end{cases} \tag{15}$$

## 4 Problem Formulation

This article aims at achieving a good tradeoff performance between the average AoI at the server and the average energy consumption. The former depends on the sampling action, uplink time and power allocations of all the frames. Clearly, due to the random motion of the human blockers $\{\mathcal{L}^{\mathrm{o}}_t|\forall t\}$ and random path gains $\{\alpha_{t,k,i}|\forall t,k,i\}$, it is impossible for the BS to determine the above actions of all the frames in a deterministic manner. Instead, we shall formulate their optimization as an infinite-horizon MDP. Particularly, the system state, scheduling policy, and system cost are first defined below.

$$\begin{aligned}
&\Pr\left[\mathcal{S}_{t+1}\middle|\mathcal{S}_t,\Omega(\mathcal{S}_t)\right] \\
&= \prod_{b\in\mathcal{B}} \Pr\left[l^{\mathrm{o}}_{t+1,b}\middle|l^{\mathrm{o}}_{t,b}\right] \times \Pr\left[\mathcal{Y}_{t+1}\middle|\mathcal{L}^{\mathrm{o}}_{t+1}\right] \\
&\quad\times \prod_{k\in\mathcal{K}} \left\{(1-s_{t,k})\mathbb{I}\left[A^{\mathrm{s}}_{t+1,k}=\min\{A^{\mathrm{s}}_{t,k}+1, A_{\max}\}\right] + s_{t,k}\mathbb{I}\left[A_{t+1,k}=1\right]\right\} \\
&\quad\times \prod_{k\in\mathcal{K}} \left\{(1-s_{t,k})\mathbb{I}\left[Q_{t+1,k}=(Q_{t,k}-D_{t,k})^+\middle|Y_{t,k},\tau_{t,k},p_{t,k}\right]\right. \\
&\qquad \left. + s_{t,k}\mathbb{I}\left[Q_{t+1,k}=Q^{\max}_k - D_{t,k}\middle|Y_{t,k},\tau_{t,k},p_{t,k}\right]\right\} \\
&\quad\times \prod_{k\in\mathcal{K}} \left\{\mathbb{I}\left[(Q_{t,k}-D_{t,k})^+=0\right]\mathbb{I}\left[A^{\mathrm{d}}_{t+1,k}=\min\{A^{\mathrm{s}}_{t,k}+1, A_{\max}\}\right]\right. \\
&\qquad \left. + \mathbb{I}\left[(Q_{t,k}-D_{t,k})^+\neq 0\right]\mathbb{I}\left[A^{\mathrm{d}}_{t+1,k}=\min\{A^{\mathrm{d}}_{t,k}+1, A_{\max}\}\right]\right\}
\end{aligned} \tag{16}$$

*Definition 1 (System State).* At the beginning of the $t$th frame, the global system state is uniquely specified by a tuple $\mathcal{S}_t \triangleq (\mathcal{L}^{\mathrm{o}}_t, \mathcal{Y}_t, \mathcal{Q}_t, \mathcal{A}^{\mathrm{s}}_t, \mathcal{A}^{\mathrm{d}}_t)$, where $\mathcal{L}^{\mathrm{o}}_t$ is the set of location indices of all the mobile blockers, $\mathcal{Y}_t \triangleq \{Y_{t,k}\}_{k\in\mathcal{K}}$ is the set of baseband gains of all sensors, $\mathcal{Q}_t \triangleq \{Q_{t,k}\}_{k\in\mathcal{K}}$ is the set of uplink queue lengths, $\mathcal{A}^{\mathrm{s}}_t \triangleq \{A^{\mathrm{s}}_{t,k}\}_{k\in\mathcal{K}}$ is the set of AoIs at the sensors, and $\mathcal{A}^{\mathrm{d}}_t \triangleq \{A^{\mathrm{d}}_{t,k}\}_{k\in\mathcal{K}}$ is the set of AoIs at the server. Moreover, the local system state of the $k$th sensor in the $t$th frame is defined by $\mathcal{S}_{t,k} \triangleq (\mathcal{L}^{\mathrm{o}}_t, Y_{t,k}, Q_{t,k}, A^{\mathrm{s}}_{t,k}, A^{\mathrm{d}}_{t,k})$.

*Definition 2 (Action and Policy).* The local scheduling action of the $k$th sensor, including the sampling decision, the uplink transmission time and power, is defined as $\mathbf{a}_{t,k} \triangleq (s_{t,k}, \tau_{t,k}, p_{t,k})$.

[2]We assume that the AoI is upper bounded by $A_{\max}$ as in [41, 42]. This is because, for time-critical applications, it is meaningless for the server to receive significantly outdated status information from the sensors. Thus, the AoIs exceeding a threshold are equally bad for the server. Moreover, for tractability, $A_{\max}$ is finite but can be arbitrarily large.

The global scheduling action is defined as the aggregation of the local actions of all the sensors; thus, $\mathbf{a}_t \triangleq \{\mathbf{a}_{t,k}\}_{k\in\mathcal{K}}$. Hence, the scheduling policy, denoted as $\Omega$, is a mapping from the system state $\mathcal{S}_t$ to the scheduling action, i.e., $\Omega(\mathcal{S}_t) = \mathbf{a}_t$.

Hence, given the scheduling policy $\Omega$, the state transition probability can be written as in (16). Moreover, in the $t$th frame, the per-frame cost is defined as the weighted sum of AoIs at the server for all sensors, energy consumption of sampling and uplink transmission, and outdated AoI penalties at the server for all sensors. That is,

$$g(\mathcal{S}_t, \Omega(\mathcal{S}_t)) = \sum_{k\in\mathcal{K}} \left[A^{\mathrm{d}}_{t,k} + w_{\mathrm{P}}(s_{t,k}C^{\mathrm{s}} + \tau_{t,k}p_{t,k}) + w_{\mathrm{Q}}\mathbb{I}\left[A^{\mathrm{d}}_{t,k} = A_{\max}\right]\right], \tag{17}$$

where $w_{\mathrm{P}}$ and $w_{\mathrm{Q}}$ denote the weights for energy consumption and AoI outdatedness penalty, respectively. Note that the first and third terms of (17) imply a nonlinear cost function of AoI [27]. Hence, the overall average cost from the 1-st frame is defined as

$$\overline{G}(\mathcal{S}_1, \Omega) \triangleq \lim_{T\to\infty} \left[\mathbb{E}_{\boldsymbol{\mathcal{Y}}, \mathcal{L}^{\mathrm{o}}} \sum_{t=1}^{T} \gamma^{t-1} g_t(\mathcal{S}_t, \Omega(\mathcal{S}_t)) \middle| \mathcal{S}_1\right], \tag{18}$$

where $\gamma \in (0,1)$ is the discount factor, and the expectation is taken on $\boldsymbol{\mathcal{Y}} \triangleq \{\mathcal{Y}_1, \mathcal{Y}_2, \ldots, \mathcal{Y}_T\}$ and $\mathcal{L}^{\mathrm{o}} \triangleq \{\mathcal{L}^{\mathrm{o}}_1, \mathcal{L}^{\mathrm{o}}_2, \ldots, \mathcal{L}^{\mathrm{o}}_T\}$. As a result, the joint sampling and uploading optimization can be formulated as the following infinite-horizon MDP with discounted cost.

$$\mathrm{P1}: \ \Omega^{\star} = \arg\min_{\Omega} \ \overline{G}(\mathcal{S}_1, \Omega) \tag{19a}$$

$$\text{s.t.} \quad \text{Constraints in (7), (11), (12).} \tag{19b}$$

The Bellman's equations of the above MDP are

$$W(\mathcal{S}) = \min_{\Omega(\mathcal{S})} \left[g\big(\mathcal{S}, \Omega(\mathcal{S})\big) + \gamma \sum_{\mathcal{S}'} W(\mathcal{S}') \Pr\left[\mathcal{S}' \middle| \mathcal{S}, \Omega(\mathcal{S})\right]\right], \ \forall \mathcal{S}, \tag{20}$$

where $W(\cdot)$ is the value function of the optimal scheduling policy (i.e., the optimal value function), and $\mathcal{S}'$ is the system state in the next frame given system state $\mathcal{S}$ and scheduling action $\Omega(\mathcal{S})$. Moreover, the policy minimizing the **right-hand-side** (**RHS**) of the above Bellman's equations is optimal [4].

Note that the baseband gain in the system state is continuously and independently distributed in all frames, which can be eliminated from the value function to reduce the complexity. We first define the local and global *abstract state* [14] with the baseband gain eliminated, i.e., $\widetilde{\mathcal{S}}_{t,k} \triangleq (\mathcal{L}^{\mathrm{o}}_t, Q_{t,k}, A^{\mathrm{s}}_{t,k}, A^{\mathrm{d}}_{t,k})$ and $\widetilde{\mathcal{S}}_t \triangleq (\mathcal{L}^{\mathrm{o}}_t, \boldsymbol{Q}_t, \mathcal{A}^{\mathrm{s}}_t, \mathcal{A}^{\mathrm{d}}_t)$. By taking expectation on both sides of (20), the Bellman's equations with respect to the abstract state can be simplified as

$$W(\widetilde{\mathcal{S}}) = \mathbb{E}_{\mathcal{Y}} \min_{\Omega(\mathcal{S})} \left[g\big(\mathcal{S}, \Omega(\mathcal{S})\big) + \gamma \sum_{\widetilde{\mathcal{S}}'} W(\widetilde{\mathcal{S}}') \Pr\left[\widetilde{\mathcal{S}}' \middle| \mathcal{S}, \Omega(\mathcal{S})\right]\right], \ \forall \widetilde{\mathcal{S}}, \tag{21}$$

where $\widetilde{\mathcal{S}}'$ is the abstract state in the next frame. The optimal value function with respect to the abstract state can be represented as follows:

$$W(\widetilde{\mathcal{S}}) = \mathbb{E}_{\mathcal{Y}} W(\mathcal{S}) = \min_{\Omega} \lim_{T\to\infty} \sum_{t=1}^{T} \mathbb{E}^{\Omega}_{\mathcal{Y},\mathcal{L}^{\circ}} \gamma^{t-1} \Big[ g\big(\mathcal{S}_t, \Omega(\mathcal{S}_t)\big) \Big| \widetilde{\mathcal{S}}_1 = \widetilde{\mathcal{S}} \Big], \ \forall \widetilde{\mathcal{S}}. \tag{22}$$

With the value function of the optimal policy and the current system state $\mathcal{S}_t$, the optimal action in the $t$th frame is given by

$$\Omega^{\star}(\mathcal{S}_t) = \arg\min_{\Omega(\mathcal{S}_t)} \left[ g_t\big(\mathcal{S}_t, \Omega(\mathcal{S}_t)\big) + \gamma \sum_{\widetilde{\mathcal{S}}_{t+1}} W(\widetilde{\mathcal{S}}_{t+1}) \Pr\left[\widetilde{\mathcal{S}}_{t+1} \middle| \mathcal{S}_t, \Omega(\mathcal{S}_t)\right] \right], \ \forall t, \forall \mathcal{S}_t. \tag{23}$$

Although conventional approaches such as **value iteration algorithm** (**VIA**) can be used to find the optimal scheduling policy [4], they suffer from the *curse of dimensionality*: due to the huge system state space, the evaluation of $\{W(\widetilde{\mathcal{S}}_t) | \forall \widetilde{\mathcal{S}}_t\}$ is prohibitive. In the following section, we shall propose a low-complexity scheduling scheme to address this issue.

## 5 Low-Complexity Scheduling

In this section, COSMO, a low-complexity scheduling scheme with non-trivial analytical performance bound is proposed. COSMO includes two stages. Particularly,

- —in Section 5.1, an offline procedure for value function approximation is introduced.[3] We design a heuristic scheduling policy as the reference policy, whose value function (average discounted cost) can be analytically expressed;
- —in Section 5.2, the online stage for per-frame scheduling is introduced. By approximating the optimal value function via the above-derived value function, the scheduling action of each frame can be obtained by solving the optimization problem in (23). As a result, the stochastic policy optimization in Problem P1 is decomposed into a deterministic optimization problem for each frame. Then we propose an **alternative optimization (AO)** algorithm to derive the suboptimal solution of the deterministic optimization problems;
- —in Section 5.3, we analyze the performance bound and time complexity of COSMO.

### 5.1 Decoupled Reference Policy

A heuristic reference policy is proposed to provide an expression of an achievable average discounted cost in this part. In fact, given the policy, the system evolves as a Markov chain, whose cost can be derived via the transition matrix. However, the dimension of the transition matrix grows exponentially with respect to the number of sensors, which makes the above approach infeasible. In order to address this issue, in the reference policy, the transmission time allocation of the $K$ sensors is fixed. Hence, the state transition of the whole network can be decoupled into $K$ time-invariant Markov chains with significantly smaller state space for the $K$ sensors respectively. Specifically, We adopt the following decoupled policy as the reference policy[4]:

Policy 1 (Decoupled Reference Policy $\Pi$). *The reference policy, denoted as* $\Pi \triangleq (s^{\Pi}_{t,k}, \tau^{\Pi}_{t,k}, p^{\Pi}_{t,k})$, *is elaborated below.*

[3]This offline procedure is required to be conducted only once before a large brunch of scheduling frames as long as the reflectors are static. However, the reflector dynamics can be easily detected by periodic environmental reconstruction via wireless sensing techniques [28], and then the parameters of the channel model shall be renewed for the update of approximate value functions.

[4]In fact, the reference policy is not required to be specified as the particular policy in this work and can be any policy as long as its value function can be analytically expressed. In this work, for elaboration convenience, we specified the reference policy with fixed transmission time and power allocation.

—*The sampling decision is made when the uplink queue is empty, i.e.,* $s^{\Pi}_{t,k} = \mathbb{I}[Q_{t,k} = 0]$.

—*The transmission time allocation is proportional to the data volume of corresponding sampling, i.e.,*

$$\tau^{\Pi}_{t,k} = T_{\mathrm{F}} \frac{Q^{\max}_k}{\sum_{k' \in \mathcal{K}} Q^{\max}_{k'}}. \tag{24}$$

—*The transmission power is constant, i.e.,* $p^{\Pi}_{t,k} = P^{\Pi}$.

With the reference policy $\Pi$, the average discounted cost of each sensor can be calculated according to the transition matrices of their local abstract states, respectively. Note that the space of the local abstract state of the $k$th sensor, denoted as $\widetilde{\mathcal{S}}_{t,k} = (\mathcal{L}^{\mathrm{o}}_t, Q_{t,k}, A^{\mathrm{s}}_{t,k}, A^{\mathrm{d}}_{t,k})$, is with a cardinality of $L^B(Q^{\max}_k + 1)A^2_{\max}$. We can use the vector $\mathbf{s}_{t,k} \in \mathbb{R}^{L^B(Q^{\max}_k+1)A^2_{\max}\times 1}$ to represent the distribution of local abstract state of $k$th sensor in the $t$th frame. The $\kappa\big(\varsigma(\mathcal{L}^{\mathrm{o}}_t), \epsilon(Q_{t,k}, A^{\mathrm{s}}_{t,k}, A^{\mathrm{d}}_{t,k})\big)$-th entry of $\mathbf{s}_{t,k}$ denotes the probability of the local abstract state $\widetilde{\mathcal{S}}_{t,k} = (\mathcal{L}^{\mathrm{o}}_t, Q_{t,k}, A^{\mathrm{s}}_{t,k}, A^{\mathrm{d}}_{t,k})$ in the $t$th frame, where

$$\kappa(\varsigma, \epsilon) \triangleq (\varsigma - 1)\left(Q^{\max}_k + 1\right) A^2_{\max} + \epsilon, \tag{25}$$

$$\varsigma(\mathcal{L}^{\mathrm{o}}_t) \triangleq \sum_{b \in \mathcal{B}} l^{\mathrm{o}}_{t,b} L^{B-b}, \tag{26}$$

and

$$\epsilon\left(Q_{t,k}, A^{\mathrm{s}}_{t,k}, A^{\mathrm{d}}_{t,k}\right) \triangleq Q_{t,k}A^2_{\max} + \left(A^{\mathrm{s}}_{t,k} - 1\right) A_{\max} + A^{\mathrm{d}}_{t,k}. \tag{27}$$

As a result, the time-invariant transition probability matrix of the local abstract state of the $k$th sensor $\mathbf{P}_k \in \mathbb{R}^{L^B(Q^{\max}_k+1)A^2_{\max}\times L^B(Q^{\max}_k+1)A^2_{\max}}$ satisfies

$$\mathbf{s}_{t+1,k} = \mathbf{P}^{\mathsf{T}}_k \mathbf{s}_{t,k} = \left(\mathbf{P}^{\mathsf{T}}_k\right)^t \mathbf{s}_{1,k}. \tag{28}$$

In order to derive the expression of $\mathbf{P}_k$, we first introduce the following lemma on the distribution of departure packet number.

Lemma 1. *With sufficiently large* $N_{\mathrm{R}}$ *and* $N_{\mathrm{T}}$*, given the reference policy* $\Pi$ *and the blocker's location* $\mathcal{L}^{\mathrm{o}}_t$*, the optimal solution of (10) can be determined by* $(\mathbf{w}_{t,k}, \mathbf{f}_{t,k}) = (\mathbf{a}_{\mathrm{R}}(\Phi_{k,i^\star}), \mathbf{a}_{\mathrm{T}}(\Theta_{k,i^\star}))$*, and the* ***probability mass function (PMF)*** *of departure packet number can be written by (29), where* $i^\star = i^\star(k, \mathcal{L}^{\mathrm{o}}_t) = \arg\max_i \mathbb{I}^{\mathrm{o}}_{t,k,i}(\mathcal{L}^{\mathrm{o}}_t)\rho^{-1}_{k,i}$.

Proof. Please refer to Appendix A. □

With Lemma 1, the complicated precoder and combiner optimization in (10) can be solved by selecting the best one among the $(M + 1)$ propagation paths. The distribution of packet departure can be determined accordingly. As a result, we have the following lemma on the transition matrix $\mathbf{P}_k$.

Lemma 2. *With sufficiently large* $N_{\mathrm{R}}$ *and* $N_{\mathrm{T}}$*, given the reference policy* $\Pi$*, the transition probability matrix of local abstract state of the* $k$*th sensor is given by (30), where* $\mathbf{M}^{(\ell)}_k \in \mathbb{R}^{(Q^{\max}_k+1)A^2_{\max}\times(Q^{\max}_k+1)A^2_{\max}}$ *is given by Table 2,* $\mathbf{P}^{\mathrm{o}} \triangleq \mathbf{P}^{\mathrm{o}}_1 \otimes \mathbf{P}^{\mathrm{o}}_2 \otimes \cdots \otimes \mathbf{P}^{\mathrm{o}}_B \in \mathbb{R}^{L^B\times L^B}$*, and* $\mathbf{P}^{\mathrm{o}}_b$ *is defined in (1).*

Proof. Please refer to Appendix B. □

Table 2. Non-Zero Entries of Matrix $\mathbf{M}_k^{(\ell)}$

| $Q_{t,k}$ | $A^{\mathrm{s}}_{t,k}$ | $A^{\mathrm{d}}_{t,k}$ | $Q_{t+1,k}$ | $A^{\mathrm{s}}_{t+1,k}$ | $A^{\mathrm{d}}_{t+1,k}$ | $\left[\mathbf{M}_k^{(\ell)}\right]_{\epsilon(Q_{t,k},A^{\mathrm{s}}_{t,k},A^{\mathrm{d}}_{t,k}),\epsilon(Q_{t+1,k},A^{\mathrm{s}}_{t+1,k},A^{\mathrm{d}}_{t+1,k})}$ |
|---|---|---|---|---|---|---|
| 0 | $1,\ldots,A_{\max}$ | $1,\ldots,A_{\max}$ | 0 | 1 | 1 | $\Pr\left[D^{\Pi}_{t,k}\geq Q_k^{\max}\middle\vert\varsigma(\mathcal{L}^{\mathrm{o}}_t)=\ell\right]$ |
| 0 | $1,\ldots,A_{\max}$ | $1,\ldots,A_{\max}$ | $1,\ldots,Q_k^{\max}$ | 1 | $\min\{A^{\mathrm{d}}_{t,k}+1,A_{\max}\}$ | $\Pr\left[D^{\Pi}_{t,k}=Q_k^{\max}-Q_{t+1,k}\middle\vert\varsigma(\mathcal{L}^{\mathrm{o}}_t)=\ell\right]$ |
| $1,\ldots,Q_k^{\max}$ | $1,\ldots,A_{\max}$ | $1,\ldots,A_{\max}$ | 0 | $\min\{A^{\mathrm{s}}_{t,k}+1,A_{\max}\}$ | $\min\{A^{\mathrm{d}}_{t,k}+1,A_{\max}\}$ | $\Pr\left[D^{\Pi}_{t,k}\geq Q_{t,k}\middle\vert\varsigma(\mathcal{L}^{\mathrm{o}}_t)=\ell\right]$ |
| $1,\ldots,Q_k^{\max}$ | $1,\ldots,A_{\max}$ | $1,\ldots,A_{\max}$ | $1,\ldots,Q_{t,k}$ | $\min\{A^{\mathrm{s}}_{t,k}+1,A_{\max}\}$ | $\min\{A^{\mathrm{d}}_{t,k}+1,A_{\max}\}$ | $\Pr\left[D^{\Pi}_{t,k}=Q_{t,k}-Q_{t+1,k}\middle\vert\varsigma(\mathcal{L}^{\mathrm{o}}_t)=\ell\right]$ |

$$\Pr\left[D^{\Pi}_{t,k}=d\middle|\mathcal{L}^{\mathrm{o}}_t\right]=\exp\left(-\frac{\rho_{k,i^\star}N_0W}{P^{\Pi}}\left(2^{\frac{dN_{\mathrm{b}}}{W\tau^{\Pi}_{t,k}}}-1\right)\right)-\exp\left(-\frac{\rho_{k,i^\star}N_0W}{P^{\Pi}}\left(2^{\frac{(d+1)N_{\mathrm{b}}}{W\tau^{\Pi}_{t,k}}}-1\right)\right),\tag{29}$$

$$\mathbf{P}_k=\begin{pmatrix}[\mathbf{P}^{\mathrm{o}}]_{1,1}\mathbf{M}_k^{(1)} & [\mathbf{P}^{\mathrm{o}}]_{1,2}\mathbf{M}_k^{(1)} & \cdots & [\mathbf{P}^{\mathrm{o}}]_{1,L^B}\mathbf{M}_k^{(1)}\\ [\mathbf{P}^{\mathrm{o}}]_{2,1}\mathbf{M}_k^{(2)} & [\mathbf{P}^{\mathrm{o}}]_{2,2}\mathbf{M}_k^{(2)} & \cdots & [\mathbf{P}^{\mathrm{o}}]_{2,L^B}\mathbf{M}_k^{(2)}\\ \vdots & \vdots & \ddots & \vdots\\ [\mathbf{P}^{\mathrm{o}}]_{L^B,1}\mathbf{M}_k^{(L^B)} & [\mathbf{P}^{\mathrm{o}}]_{L^B,2}\mathbf{M}_k^{(L^B)} & \cdots & [\mathbf{P}^{\mathrm{o}}]_{L^B,L^B}\mathbf{M}_k^{(L^B)}\end{pmatrix},\tag{30}$$

Finally, the value function of the reference policy (referred to as the approximate value function), representing the average discounted cost with the reference policy, is given by the following theorem.

Theorem 1 (Value Function of Reference Policy Π). *With the reference policy* Π*, the approximate value function is given by*

$$W^{\Pi}(\widetilde{\mathcal{S}}_t)=\sum_{k\in\mathcal{K}}\underbrace{\left(\mathbf{e}^{L^B(Q_k^{\max}+1)A^2_{\max}}_{\kappa\left(\varsigma(\mathcal{L}^{\mathrm{o}}_t),\epsilon(Q_{t,k},A^{\mathrm{s}}_{t,k},A^{\mathrm{d}}_{t,k})\right)}\right)^{\mathsf{T}}[\mathbf{I}-\gamma\mathbf{P}_k]^{-1}\mathbf{g}_k}_{\triangleq W_k^{\Pi}(\widetilde{\mathcal{S}}_{t,k})}+\frac{1}{1-\gamma}w_{\mathrm{P}}T_{\mathrm{F}}P^{\Pi},\tag{31}$$

*where* $\mathbf{e}^N_n$ *denotes an* $N\times 1$ *column vector whose* $n$*th element is 1 and 0 otherwise,* $\mathbf{g}_k\in\mathbb{R}^{L^B(Q_k^{\max}+1)A^2_{\max}\times 1}$ *is the cost vector for all abstract local state at the* $k$*th sensor, and* $W_k^{\Pi}(\widetilde{\mathcal{S}}_{t,k}$ *is the approximate local value function. Specifically,*

$$[\mathbf{g}_k]_{\kappa\left(\varsigma(\mathcal{L}^{\mathrm{o}}_t),\epsilon(Q_{t,k},A^{\mathrm{s}}_{t,k},A^{\mathrm{d}}_{t,k})\right)}\triangleq A^{\mathrm{d}}_{t,k}+\mathbb{I}[Q_{t,k}=0]w_{\mathrm{P}}C^{\mathrm{s}}+\mathbb{I}[A^{\mathrm{d}}_{t,k}=A_{\max}]w_{\mathrm{Q}}.\tag{32}$$

Proof. Please refer to Appendix C. □

As a result, the approximate value function can be evaluated in a distributed and offline manner. The procedure to calculate the approximate local value function $W_k^{\Pi}(\widetilde{\mathcal{S}}_{t,k})$ $(\forall\widetilde{\mathcal{S}}_{t,k})$ at each sensor (say the $k$th sensor) is summarized in Algorithm 1.

### 5.2 Scheduling with Approximate Value Function

Substituting the optimal value function $W(\widetilde{\mathcal{S}}_{t+1})$ of the problem in (23) with the approximate value function $W^{\Pi}(\widetilde{\mathcal{S}}_{t+1})$, the proposed scheduling policy in one frame (say the $t$th frame) given the global system state $\mathcal{S}_t$ can be obtained from the following optimization problem. In other words, the complicated policy optimization in Problem P1 is decomposed into the following deterministic

**ALGORITHM 1:** Evaluation of approximate local value function at the $k$-th sensor

**Output:**
$W_k^{\Pi}(\widetilde{\mathcal{S}}_{t,k})$ $(\forall \widetilde{\mathcal{S}}_{t,k})$: Approximate local value function

1 Compute $i^{\star}(k, \mathcal{L}_t^{\mathrm{o}})$ via Lemma 1.
2 Compute $\mathbf{P}_k$ via Lemma 2.
3 Compute $W_k^{\Pi}(\widetilde{\mathcal{S}}_{t,k})$ $(\forall \widetilde{\mathcal{S}}_{t,k})$ via its definition in Theorem 1.
4 **return** $W_k^{\Pi}(\widetilde{\mathcal{S}}_{t,k})$ $(\forall \widetilde{\mathcal{S}}_{t,k})$

optimization problem at the beginning of each frame.

$$\mathrm{P2}: \left\{(s_{t,k}^{\star}, \tau_{t,k}^{\star}, p_{t,k}^{\star})\right\}_{k\in\mathcal{K}} = \underset{\left\{(s_{t,k},\tau_{t,k},p_{t,k})\right\}_{k\in\mathcal{K}}}{\arg\min} \sum_{k\in\mathcal{K}} w_{\mathrm{P}}(\tau_{t,k}p_{t,k} + s_{t,k}C^{\mathrm{s}}) + \gamma \sum_{\widetilde{\mathcal{S}}_{t+1}} W^{\Pi}(\widetilde{\mathcal{S}}_{t+1}) \Pr\left[\widetilde{\mathcal{S}}_{t+1} \middle| \mathcal{S}_t, (s_{t,k}, \tau_{t,k}, p_{t,k})_{k\in\mathcal{K}}\right], \tag{33}$$

$$\text{s.t.} \qquad \text{Constraints in (7), (11), (12).} \tag{34}$$

*Remark 3 (Predictive Scheduling).* Note that the LoS path is usually much better than the NLoS paths. If the LoS path of one sensor (say the $k$th sensor) might be blocked soon, the penalty of AoI outdatedness will be easily triggered. Thus, its local value function $W_k^{\Pi}(\widetilde{\mathcal{S}}_{t,k})$ could be large unless more transmission resources are scheduled to this sensor such that its AoI is maintained at a low level. As a result, the prediction of future path blockage is considered in the scheduling of the current frame.

Problem P2 is a mixed continuous and discrete optimization problem with coupled variables $\{s_{t,k}\}_{k\in\mathcal{K}}$, $\{\tau_{t,k}\}_{k\in\mathcal{K}}$ and $\{p_{t,k}\}_{k\in\mathcal{K}}$. An AO algorithm [32] is proposed in Algorithm 2 to obtain the suboptimal solution in each frame (say the $t$th frame). Let $d_{t,k}$ be the number of transmission packets from the $k$th sensor in the $t$th frame, the optimizations of $\{d_{t,k}\}_{k\in\mathcal{K}}$ and $\{d_{t,k}^{(n)}\}_{k\in\mathcal{K}}$ are equivalent. Hence, we optimize $\{s_{t,k}\}_{k\in\mathcal{K}}$, $\{\tau_{t,k}\}_{k\in\mathcal{K}}$ and $\{d_{t,k}\}_{k\in\mathcal{K}}$ alternatively in Algorithm 2.

Particularly, let $\{s_{t,k}^{(n)}\}_{k\in\mathcal{K}}$, $\{\tau_{t,k}^{(n)}\}_{k\in\mathcal{K}}$ and $\{d_{t,k}^{(n)}\}_{k\in\mathcal{K}}$ be the correspondingly optimized variables in the $n$th iteration, respectively. Initializing them with the reference policy $\Pi$ (Line 1 of Algorithm 2), i.e., $(s_{t,k}^{(0)}, \tau_{t,k}^{(0)}, d_{t,k}^{(0)}) = (s_{t,k}^{\Pi}, \tau_{t,k}^{\Pi}, d_{t,k}^{\Pi})$, the entire procedure of solving P2 consists of a number of iterations. In the $n$th iteration, the sub-problems P2.1$(n,k)$, P2.2$(n,k)$, $\forall k$, and P2.3$(n)$ in (37), (38) and (39) are solved respectively (Lines 5, 7, 8), where

$$p_{t,k}\left(d_{t,k}\right) = \frac{1}{Y_{t,k}}\left[2^{\wedge}\left(\frac{d_{t,k}N_{\mathrm{b}}}{\tau_{t,k}^{(n-1)}W}\right) - 1\right], \tag{35}$$

$$p_{t,k}(\tau_{t,k}) = \frac{1}{Y_{t,k}}\left[2^{\wedge}\left(\frac{d_{t,k}^{(n)}N_{\mathrm{b}}}{\tau_{t,k}W}\right) - 1\right]. \tag{36}$$

Notice that the originally coupled sampling decision $s_{t,k}^{(n)}$ and transmission packet number allocation $d_{t,k}^{(n)}$ of all sensors are decoupled given the transmission time allocation $\{\tau_{t,k}^{(n)}\}_{k\in\mathcal{K}}$, thus the

**ALGORITHM 2:** Online scheduling via AO algorithm.

**Input:**

$\mathcal{S}_t$: System state in the $t$th frame

$\{W_k^{\Pi}(\widetilde{\mathcal{S}}_{t,k})\}_{\forall k, \widetilde{\mathcal{S}}_{t,k}}$: Approximate local value functions via Algorithm 1.

**Output:**

$\Psi^{\infty}(\mathcal{S}_t)$: Converged action in the $t$th frame.

$n \leftarrow 0$, $s_{t,k}^{(0)} \leftarrow s_{t,k}^{\Pi}$, $\tau_{t,k}^{(0)} \leftarrow \tau_{t,k}^{\Pi}$, $d_{t,k}^{(0)} \leftarrow d_{t,k}^{\Pi}$.

**while** *not converge* **do**

  $n \leftarrow n+1$

  **for** $k \in \mathcal{K}$ **do in parallel**

    *Solve* $s_{t,k}^{(n)}$ *in* P2.1$(n,k)$ *via comparison.*

  **for** $k \in \mathcal{K}$ **do in parallel**

    *Solve* $d_{t,k}^{(n)}$ *in* P2.2$(n,k)$ *via one-dimensional search.*

  *Solve* $\{\tau_{t,k}^{(n)}\}_{k\in\mathcal{K}}$ *in* P2.3$(n)$ *via Lemma 3.*

  $p_{t,k}^{(n)} = \frac{1}{Y_{t,k}}\left[2^{\wedge}\left(\frac{d_{t,k}^{(n)} N_{\mathrm{b}}}{\tau_{t,k}^{(n)} W}\right) - 1\right], \forall k$

**return** $\Psi^{\infty}(\mathcal{S}_t) \leftarrow \{(s_{t,k}^{\infty}, \tau_{t,k}^{\infty}, p_{t,k}^{\infty})\}_{k\in\mathcal{K}}$

complexity is significantly reduced.

$$\mathrm{P2.1}(n,k): \; s_{t,k}^{(n)} = \underset{s_{t,k}\in\{0,1\}}{\arg\min} \; w_{\mathrm{P}} s_{t,k} C^{\mathrm{s}} + \gamma \sum_{\widetilde{\mathcal{S}}_{t+1,k}} W_k^{\Pi}(\widetilde{\mathcal{S}}_{t+1,k}) \Pr\left[\widetilde{\mathcal{S}}_{t+1,k} \middle| \mathcal{S}_{t,k}, (s_{t,k}, \tau_{t,k}^{(n-1)}, p_{t,k}^{(n-1)})\right], \tag{37}$$

$$\begin{aligned}\mathrm{P2.2}(n,k): \; d_{t,k}^{(n)} = \underset{d_{t,k}}{\arg\min} \; & w_{\mathrm{P}} \tau_{t,k}^{(n-1)} p_{t,k}(d_{t,k}) \\ & + \gamma \sum_{\widetilde{\mathcal{S}}_{t+1,k}} W_k^{\Pi}(\widetilde{\mathcal{S}}_{t+1,k}) \Pr\left[\widetilde{\mathcal{S}}_{t+1,k} \middle| \mathcal{S}_{t,k}, \left(s_{t,k}^{(n)}, \tau_{t,k}^{(n-1)}, p_{t,k}(d_{t,k})\right)\right] && \text{(38a)} \\ \text{s.t.} \quad & p_{t,k}(d_{t,k}) \le P_{\max} && \text{(38b)}\end{aligned}$$

$$\begin{aligned}\mathrm{P2.3}(n): \quad \left\{\tau_{t,k}^{(n)}\right\}_{k\in\mathcal{K}} = \; & \underset{\{\tau_{t,k}\}_{k\in\mathcal{K}}}{\arg\min} \sum_{k\in\mathcal{K}} \tau_{t,k} p_{t,k}(\tau_{t,k}) && \text{(39a)} \\ \text{s.t.} \quad & \sum_{k\in\mathcal{K}} \tau_{t,k} = T_{\mathrm{F}}, && \text{(39b)} \\ & 0 \le \tau_{t,k} \le T_{\mathrm{F}}, \quad \forall k \in \mathcal{K}, && \text{(39c)} \\ & p_{t,k}(\tau_{t,k}) \le P_{\max}, \; \forall k \in \mathcal{K} && \text{(39d)}\end{aligned}$$

The optimal solution of P2.1$(n,k)$, $\forall k$ can be derived by evaluating the binary local sampling action for $s_{t,k} = 0$ and $s_{t,k} = 1$ and choosing the one with the smaller value of the objective function. Similarly, the optimal solution of P2.2$(n,k)$, $\forall k$ can be solved by one-dimensional search over $d_{t,k} \in \{0, 1, \ldots, Q_{t,k}\}$.

Moreover, P2.3$(n)$ is a convex optimization problem and the optimal solution can be derived by the following lemma.

Lemma 3. *The optimal solution of* P2.3($n$) *is given by*

$$\tau_{t,k}^{(n)} = \frac{d_{t,k}N_b}{W} \max\left\{\frac{\ln 2}{1+\mathbb{W}_0\left(\frac{Y_{t,k}\nu^\star-1}{e}\right)}, \frac{1}{\log_2(1+P_{\max}Y_{t,k})}\right\}, \tag{40}$$

*where* $\mathbb{W}_0(\cdot)$ *denotes the principal branch of the Lambert W function, and* $\nu^\star$ *denotes the optimal Lagrangian multiplier for the equality constraint (11) which can be solved by*

$$\sum_{k\in\mathcal{K}} \frac{d_{t,k}N_{\mathrm{b}}}{W} \max\left\{\frac{\ln 2}{1+\mathbb{W}_0\left(\frac{Y_{t,k}\nu^\star-1}{e}\right)}, \frac{1}{\log_2(1+P_{\max}Y_{t,k})}\right\} - T_{\mathrm{F}} = 0. \tag{41}$$

Proof. Please refer to Appendix D. □

Because the **left-hand-side** (**LHS**) of (41) is non-increasing, the bisection method can be used to solve $\nu^\star$. Finally, given $d_{t,k}^{(n)}$ and $\tau_{t,k}^{(n)}$, the transmission power allocation in the $n$th iteration, denoted by $p_{t,k}^{(n)}$, can be derived analytically.

Since the iterative optimization of the sub-problems P2.1($n,k$), P2.2($n,k$), $\forall k$, and P2.3($n$) would lead to a non-increasing value of the objective function in P2, the above iteration will always converge (Line 10). In summary, the overall procedure of COSMO is described in Algorithm 3.

**ALGORITHM 3:** COSMO

1 **Stage 1: Offline Evaluation of Approximate Value Functions**
2 **At each sensor (say the $k$th sensor):**
3 Each sensor evaluates its approximate local value function $W_k^{\Pi}(\widetilde{\mathcal{S}}_{t,k})$ $(\forall\widetilde{\mathcal{S}}_{t,k})$ via Algorithm 1 and reports the results to the server.
4 **Stage 2: Centralized Online Scheduling**
5 **At the beginning of each frame (say the $t$th frame):**
6 The server collects the real-time locations of the mobile blockers $\mathcal{L}_t^{\mathrm{o}}$ via wireless sensing and baseband gains $\mathcal{Y}_t$ via channel estimation.
7 The server calculates the proposed policy in the $t$th frame $\Psi^\infty(\mathcal{S}_t)$ via Algorithm 2 and broadcasts the policy to the sensors.
8 The sensors conduct status sampling and uplink transmission according to $\Psi^\infty(\mathcal{S}_t)$.

### 5.3 Performance Bound and Complexity Analysis

The analytical performance bound and complexity analysis are provided in this part. Denote

$$\Psi^{(n)}: \widetilde{\mathcal{S}}_t \to \left\{s_{t,k}^{(n)}, \tau_{t,k}^{(n)}, p_{t,k}^{(n)}\right\}_{k\in\mathcal{K}},$$

as the scheduling policy obtained after $n$ iterations of Algorithm 2,

$$\widetilde{W}^{\Psi^{(n)}}(\widetilde{\mathcal{S}}_t),$$

as the corresponding value functions, and

$$\Psi^\infty: \widetilde{\mathcal{S}}_t \to \left\{s_{t,k}^{\infty}, \tau_{t,k}^{\infty}, p_{t,k}^{\infty}\right\}_{k\in\mathcal{K}},$$

as the scheduling policy after convergence. The performance of $\Psi^\infty$ and $\Psi^{(n)}$ can be bounded as elaborated in the following lemma.

Lemma 4 (Performance Bound). *The average discounted cost of policies $\Psi^{(n)}$ and $\Psi^{\infty}$ can be bounded by*

$$W(\widetilde{\mathcal{S}}_t) \leq W^{\Psi^{\infty}}(\widetilde{\mathcal{S}}_t) \leq \cdots \leq \widetilde{W}^{\Psi^{(n)}}(\widetilde{\mathcal{S}}_t) \leq \cdots \leq \widetilde{W}^{\Psi^{(1)}}(\widetilde{\mathcal{S}}_t) \leq W^{\Pi}(\widetilde{\mathcal{S}}_t). \tag{42}$$

Proof. Since $W(\widetilde{\mathcal{S}}_t)$ is the optimal value function, it is the lower bound of the value function of an arbitrary policy. The proof of $W^{\Psi^{\infty}}(\widetilde{\mathcal{S}}_t) \leq \cdots \leq \widetilde{W}^{\Psi^{(n)}}(\widetilde{\mathcal{S}}_t) \leq \cdots \leq \widetilde{W}^{\Psi^{(1)}}(\widetilde{\mathcal{S}}_t)$ and $\widetilde{W}^{\Psi^{(1)}}(\widetilde{\mathcal{S}}_t) \leq W^{\Pi}(\widetilde{\mathcal{S}}_t)$ resembles the proof of the policy improvement property in Chapter II of [4]. □

As shown in Algorithm 3, the procedure of COSMO consists of two stages: the evaluation of the approximate local value functions (Stage 1) and the online scheduling (Stage 2). On one hand, the computational complexity of Stage 1 is $O[L^{4B}A_{\max}^{8}\sum_{k\in\mathcal{K}}(Q_k^{\max})^4]$, which grows linearly with the number of sensors $K$. Note that the computation of Stage 1 can be distributed to the sensors, the computational complexity per sensor in Stage 1 will not scale with $K$. On the other hand, the computational complexity of per-frame optimization in Stage 2 (Algorithm 2) is $O[N_{\text{iter}}^{\text{AO}}L^{B}A_{\max}^{2}\sum_{k\in\mathcal{K}}(Q_k^{\max})^2]$, where $N_{\text{iter}}^{\text{AO}}$ denotes the number of iterations in Algorithm 2. Based on the fact that the convergence of AO in Algorithm 2 only requires a few iterations (as demonstrated by our numerical simulations in the next section), thus $N_{\text{iter}}^{\text{AO}}$ is small, the most computation-intensive part of COSMO lies in Stage 1.

As a comparison, VIA is requested in the conventional solution of MDP. For VIA with $\tau_{\text{D}}$ and $p_{\text{D}}$ levels of transmission time and power respectively, the computation complexity is $O\left(N_{\text{iter}}^{\text{VIA}}|\widetilde{\mathbb{S}}|^2|\mathbb{A}|\right)$, where $N_{\text{iter}}^{\text{VIA}}$ denotes the number of iterations in the VIA, $|\widetilde{\mathbb{S}}| \triangleq L^{B}A_{\max}^{2K}\prod_{k\in\mathcal{K}}Q_k^{\max}$ and $|\mathbb{A}| \triangleq 2^{K}\tau_{\text{D}}^{K}p_{\text{D}}^{K}$ denote the cardinalities of the abstract state space and action space, respectively. It can be observed that the computation complexity of VIA grows exponentially with the number of sensors $K$, which makes the computation of the optimal value function prohibitive. Thus, benefiting from the analytical approximation of the optimal value functions in COSMO, the computation complexity is essentially reduced.

## 6 Simulations and Discussions

In this section, the performance of COSMO is demonstrated via simulations, where a number of benchmarks are used in the comparison. We summarize the key findings of our simulations as follows:

- —COSMO can converge after only a few iterations and reduce the average per-frame cost by 19.5%–52.7% compared with the benchmarks.
- —COSMO can proactively keep the AoIs of sensors with future channel degradation at a low level in example traces, which intuitively verifies the benefits of exploiting mobile blocker detection in AoI-oriented scheduling design.
- —COSMO shows better performance compared to the benchmarks with robustness under a variety of configurations.

### 6.1 Simulation Setup

As illustrated in Figure 2, we consider a 20m × 20m square room with walls serving as the reflectors of the NLoS paths ($M = 4$), where the BS is deployed at the center block of the room and the locations of sensors are uniformly distributed near the walls. The mobility of the human blockers with radius $r_{\text{B}} = 0.3$m follows a modified random walk. The probability of staying in the same grid in the next frame is 0.90, while the probability of moving to one of the feasible neighboring blocks is based on the probability distribution of human locations in a room which can be calculated by

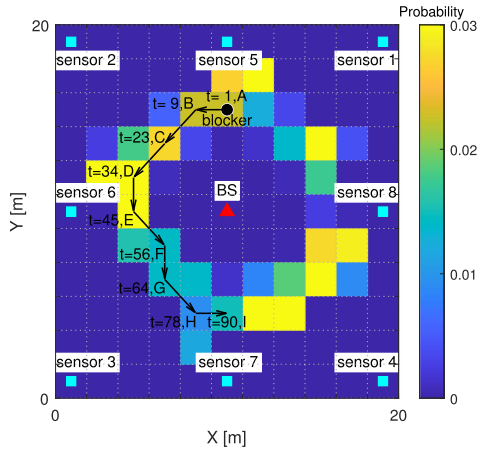


Fig. 2. Simulation scenario and an illustrative trajectory of the blocker in scenario of single blocker.

Table 3. Parameter Configuration of the Simulation

| Parameter | Symbol | Value |
|---|---|---|
| Carrier frequency | $f_c$ | 60.48 GHz |
| Number of sensors | $K$ | 8 |
| Number of blockers | $B$ | 1 |
| Number of receive/transmit antenna elements | $N_R$, $N_T$ | 128,64 [3] |
| Path loss | $\rho_{k,i}$ | LoS: $32.5 + 20\log(f_c) + 20\log(R)$ [18]<br>NLoS: $32.5 + 20\log(f_c) + 20\log(R) + 15$ [34] |
| Number of possible locations of the blocker | $L$ | 48 |
| Bandwidth | $W$ | 400 MHz [37] |
| Frame duration | $T_F$ | 10 ms [29] |
| Packet size | $N_b$ | 200 KB |
| Noise power spectral density | $N_0$ | -174 dBm/Hz [37] |
| Data volume | $Q_k^{\max}$ | $\mathcal{U}(3, 4)$ |
| Threshold for outdated AoI | $A_{\max}$ | 10 |
| Maximum transmission power | $P_{\max}$ | 100 mW [2] |
| Discount factor | $\gamma$ | 0.98 |
| Weights for energy consumption | $w_P$ | 10000 |
| Weights for outdated AoI penalty | $w_Q$ | 15 |
| Sampling energy consumption | $C^s$ | $2 \times 10^{-4}$ J |
| Transmission power of reference policy | $P^\Pi$ | 20 mW |

wireless localization techniques such as [28]. To generate a practical location probability distribution and trajectories of human blockers, we use the OMNI1 dataset [19] which includes 1,600 human trajectories through a lab captured by an omni-directional camera. The location probability distribution is illustrated in Figure 2 where the colors indicate the probability of each grid and an illustrative blocker trajectory generated from it is shown by the arrows. Other parameters are summarized in Table 3.

We evaluate COSMO with different numbers of AO iterations in solving P2, namely, $\Psi^{(1)}$, $\Psi^{(2)}$, $\Psi^{(3)}$, $\Psi^{\infty}$, and compare them with the following three benchmark policies, which are referred to as

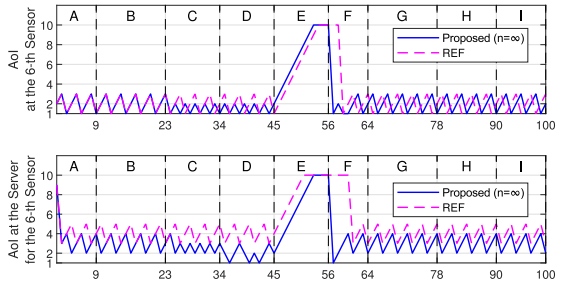


Fig. 3. An AoI trial of the 6-th sensor.

REF, DBP and LAF, respectively. All the benchmarks adopt the same sampling policy and transmission power allocation as the reference policy, i.e., $s_{t,k} = \mathbb{I}[Q_{t,k} = 0]$ and $p_{t,k} = P^{\Pi}$, while the transmission time allocation is elaborated as follows:

Benchmark 1 (Reference Policy, REF). *The transmission time allocation of each sensor is proportional to the corresponding data volume of the sample at each sensor, i.e.,* $\tau_{t,k} = T_{\mathrm{F}} Q_k^{\max} / \sum_{k' \in \mathcal{K}} Q_{k'}^{\max}$.

Benchmark 2 (Dynamic Backpressure [8], DBP). *The sensor with the largest product of buffer length and uplink capacity, i.e.,* $\arg\max_k Q_{t,k} R_{t,k}$, *is scheduled for transmission sequentially until the transmission time of the frame is used up.*

Benchmark 3 (Largest-AoI First, LAF). *The sensor with the largest AoI at the server, i.e.,* $\arg\max_k A_{t,k}^{\mathrm{d}}$, *is scheduled for transmission sequentially until the transmission time of the frame is used up.*

### 6.2 Simulation Results

The insights on blockage-prediction-based scheduling of COSMO can be obtained in Figure 3, which shows a trial of the AoIs at the 6-th sensor and the server. The corresponding trajectory is illustrated in Figure 2, where the time indexes indicate the frames when the blocker moves to the grids. For example, the blocker remains in position A since the 1-st frame and moves left to position B in the 9-th frame. In this trajectory, the LoS path between the BS and the 6-th sensor is blocked when the blocker is located at the position E from the 45-th frame to the 55-th frame. In this period, only the NLoS paths are available for sample uploading at the 6-th sensor. The REF policy cannot detect such channel degradation, hence the samples are uploaded at a constant frequency. With COSMO, when the blocker is at positions B, C and D, the samples from the 6-th sensor are uploaded with higher frequencies (3, 2, and 2 frames per uploading), resulting in an average AoIs of 3, 2.5, and 2 frames, respectively. Thus, as the blocker moves closer to the LoS path between the sensor and the server, the COSMO will give a higher priority of sample uploading to the sensor, such that the penalty of AoI outage at the server can be suppressed. This can be observed from Figure 3.

Figure 4(a) displays the **cumulative distribution functions** (**CDFs**) of the per-frame cost of COSMO with the different numbers of AO iterations ($n = 1, 2, 3, \infty$) as well as the three benchmarks, where the number of sensors is $K = 8$. The average per-frame costs of COSMO with the number of AO iterations $n = 1, 2, 3, \infty$ are 171.3, 139.2, 133.7 and 133.2, respectively, while those for the three benchmarks, i.e., REF, DBP and LAF, are 282.0, 165.4 and 255.3, respectively. The cost reductions of the converged COSMO ($\Psi^{\infty}$) over the benchmarks are 52.7%, 19.5%, and 47.8%, respectively, indicating the average gain performance of COSMO. Notice that COSMO can

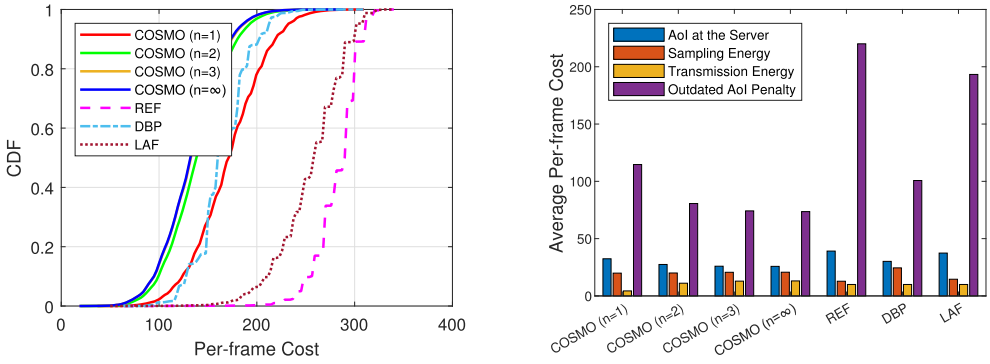


Fig. 4. (a) CDF of the per-frame cost. (b) Average performance of the AoI at the server, sampling energy, transmission energy, and outdated AoI penalty.

significantly reduce the average per-frame cost of the REF scheme by 39.2% with only one-step AO iteration, and it converges in only a few AO iterations. This indicates the efficiency of the proposed online scheduling via AO algorithm.

Figure 4(b) sketches the bar chart of the four components of per-frame cost. REF has the largest penalty on AoI outdatedness, because its scheduling is independent of the system state. DBP has the highest sampling energy consumption indicating its highest sampling frequency. This is because it tends to select the sensor with high throughput. Although LAF always prioritizes the sensors with the largest AoIs at the server in per-frame scheduling, it fails to reduce the average penalty on AoI outdatedness in the long run, due to its obliviousness to channel variation. On the other hand, COSMO achieves the best balance between AoI and energy consumption.

### 6.3 Sensitivity Analysis

We further investigate the impact of the number of sensors $K$, the number of blockers $B$, sampling energy consumption $C^{\mathrm{s}}$, as well as the weights for energy consumption and AoI outdatedness penalty $w_{\mathrm{P}}$ and $w_{\mathrm{Q}}$.

**Impact of the Number of Blockers.** The impact of the number of blockers $B$ on average per-frame cost is shown in Figure 5. As the number of blockers increases, the average per-frame cost of the policies increases because the LoS path of each sensor will suffer a higher probability of blockage. COSMO overperforms the benchmarks at both the circumstances with or without mobile blockers which shows its robustness to both static and dynamic environments.

**Impact of the Number of Sensors.** The average per-frame cost versus the number of sensors $K$ is studied in Figure 6. The average per-frame cost of COSMO is always lower than the benchmarks, which verifies the better performance of COSMO and its scalability to the number of sensors.

**Impact of Sampling Energy Consumption.** The average per-frame cost versus the sampling energy consumption $C^{\mathrm{s}}$ is shown in Figure 7. Note that $C^{\mathrm{s}} = 0$ corresponds to the scenario that either sampling at the sensor consumes negligible energy or the sensors are sampling in real-time and the server needs to determine whether to refresh the buffer with the latest status information. COSMO has the lowest average per-frame cost which implies that it is adaptive to various types of sensors.

**Impact of Weights in Cost Function.** Because the demands for energy consumption and information freshness/outdatedness may vary across different monitoring applications. We also investigate the impact of the two weights for energy consumption and AoI outdatedness penalty

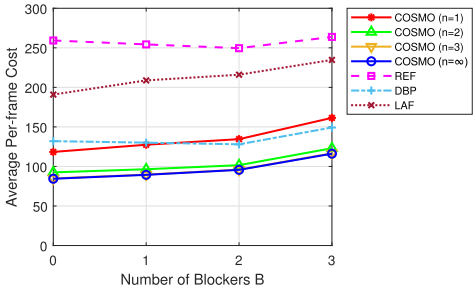


Fig. 5. Average per-frame cost versus the number of blockers.

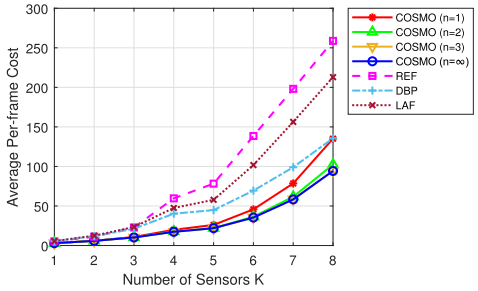


Fig. 6. Average per-frame cost versus the number of sensors.

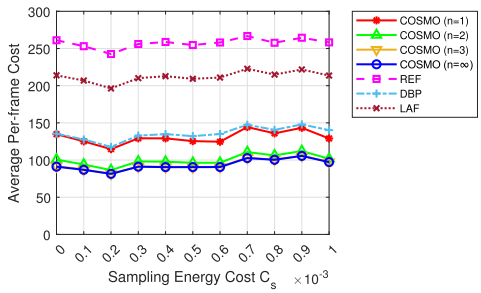


Fig. 7. Average per-frame cost versus sampling energy cost.

in the cost function, i.e., $w_P$ and $w_Q$. As illustrated in Figure 8, COSMO can always achieve the lowest average per-frame cost with different demands on energy consumption and information freshness.

**Performance Scalability.** In the above sensitivity analysis, the numbers of blockers or sensors are increased to 3 and 8, respectively. It is expected that when both numbers are even larger, the performance gain of COSMO over various benchmarks would become more significant. In fact, the gain of COSMO comes from the prediction of mmWave link blockage. Intuitively, when a link blockage is likely to happen at a sensor in the coming future, it is likely to make a sampling decision and upload the sample as fast as possible, such that the risk of AoI outdatedness at the server during the link blockage period is suppressed. With more sensors and blockers, there is generally a larger chance that COSMO could benefit from such prediction. This is the reason that an increasing trend of performance gain can be observed in Figure 6.

## 7 Discussions

As an initial attempt to exploit environment sensing and human motion tracking in the AoI-oriented scheduling design, our current design exhibits some limitations that will be addressed in the future. First, the performance of COSMO depends closely on the reference policy. In fact, any policy can be taken as the reference policy as long as its value function can be analytically expressed. Hence, it is interesting to investigate the design of reference policy with better performance. For example, it is possible to optimize the transmission and sampling parameters of the current reference policy as a new reference policy.

Second, we consider static sensors in this article. In fact, both sensors and blockers might move in practice. For example, sensors might be installed on **user equipment** (**UE**), vehicles or robotic platforms [7, 12, 17, 20] in indoor or outdoor scenarios with walking people. In these new scenarios,

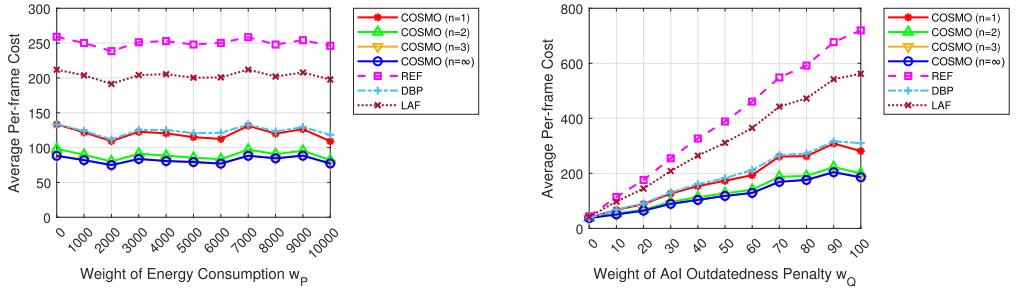


Fig. 8. Average per-frame cost versus the weights for (a) energy consumption and (b) AoI outdatedness penalty.

since sensors are moving, their locations as well as the locations of mobile blockers should be involved in the system state. Since both the blockers and other sensors might block the uplink transmission of one sensor, the motions of all blockers and sensors are coupled in the transmission scheduling, leading to huge system state space. As a result, a low-complexity algorithm design becomes necessary.

Finally, COSMO tracks the motion of $B$ blockers in this article. When $B$ increases, the computational complexity of COSMO grows exponentially. This is because the locations of all blockers are included in the local system state of each sensor, and the state space grows exponentially with $B$. Hence, the link blockage model should be revised in order to keep the complexity of COSMO at a low level. In fact, when $B$ is large, it is not practical to track the trajectories of all the blockers individually. One promising method to model the link blockage is to establish: (1) the relation between the link blockage probability and the number of mobile blockers in close proximity to the link, and (2) a stochastic process describing the dynamics of blockers' number in each region of the network.

## 8 Conclusion

In this article, we formulate the dynamic scheduling of sampling and uploading in an mmWave-based WSN with random human blockage as an infinite-horizon MDP with discounted cost, where the weighted sum of average AoI and system energy consumption is the minimization objective. Since the system state space grows exponentially with respect to the number of sensors, an approximate MDP solution framework for COSMO is proposed to address the curse of dimensionality. In the proposed framework of COSMO, the optimal value function is approximated by an analytical expression derived from a reference policy. The numerical evaluation of the value function in conventional approximate MDP solutions can be eliminated. Finally, the policy iteration based on the analytical value function is significantly more efficient than the lookup-table-based value function in conventional solution frameworks. Thus, the solution complexity is significantly reduced. Moreover, the impact of random human blockage on future costs is predicted in the approximate value function and mitigated in the scheduling of the current frame.

## Appendices

## A Proof of Lemma 1

According to [24], with sufficient large $N_R$ and $N_T$, $f_R(\Phi) \to 1$ and $f_T(\Theta) \to 1$ if $\Phi = 0$ and $\Theta = 0$, otherwise $f_R(\Phi) \to 0$ and $f_T(\Theta) \to 0$, where $f_R(\Phi) = |\mathbf{a}_R^H(0)\mathbf{a}_R(\Phi)|$ and $f_T(\Theta) = |\mathbf{a}_T^H(0)\mathbf{a}_T(\Theta)|$.

Hence, substituting (2) into (10), we can derive

$$\left(\mathbf{w}_{t,k}, \mathbf{f}_{t,k}\right) = \underset{\substack{\mathbf{a}_{\mathrm{R}}(\Phi_q)\in\mathcal{W} \\ \mathbf{a}_{\mathrm{T}}(\Theta_p)\in\mathcal{F}}}{\arg\max}\ \mathbb{E}_{(\alpha_{t,k,i})_{i\in\mathcal{M}}}\left|\sum_{i\in\mathcal{M}} \mathbb{I}^{\mathrm{o}}_{t,k,i}(\mathcal{L}^{\mathrm{o}}_t)\alpha_{t,k,i} \times f_{\mathrm{R}}(\Phi_{k,i}-\Phi_q) f_{\mathrm{T}}(\Theta_p-\Theta_{k,i})\right|^2 \tag{43}$$

$$= \underset{\substack{\mathbf{a}_{\mathrm{R}}(\Phi_q)\in\mathcal{W} \\ \mathbf{a}_{\mathrm{T}}(\Theta_p)\in\mathcal{F}}}{\arg\max}\ \mathbb{I}^{\mathrm{o}}_{t,k,i}(\mathcal{L}^{\mathrm{o}}_t)\rho^{-1}_{k,i}. \tag{44}$$

Then the analog precoder and combiner selection in (44) can be degraded into the multipath selection among the $(M+1)$ propagation paths. Let $i^\star = \arg\max_i \mathbb{I}^{\mathrm{o}}_{t,k,i}(\mathcal{L}^{\mathrm{o}}_t)\rho^{-1}_{k,i}$, (44) can be written by $\left(\mathbf{w}_{t,k}, \mathbf{f}_{t,k}\right) = \left(\mathbf{a}_{\mathrm{R}}(\Phi_{k,i^\star}), \mathbf{a}_{\mathrm{T}}(\Theta_{k,i^\star})\right)$. Then the baseband gain can be represented by $Y_{t,k}(\mathcal{L}^{\mathrm{o}}_t) = \left|\alpha_{t,k,i^\star}\right|^2/(N_0 W)$.

Given $\alpha_{t,k,i^\star} \sim \mathcal{CN}(0, \rho^{-1}_{k,i^\star})$, $|\alpha_{t,k,i^\star}|^2 = \Re^2[\alpha_{t,k,i^\star}] + \Im^2[\alpha_{t,k,i^\star}]$ follows an exponential distribution, i.e., $|\alpha_{t,k,i^\star}|^2 \sim \mathrm{Exp}(\rho_{k,i^\star})$. Then $Y_{t,k}(\mathcal{L}^{\mathrm{o}}_t) \sim \mathrm{Exp}(\rho_{k,i^\star} N_0 W)$ and thus the CDF of $Y_{t,k}(\mathcal{L}^{\mathrm{o}}_t)$ can be written by

$$\Pr[Y_{t,k}(\mathcal{L}^{\mathrm{o}}_t) \le x] = 1 - \exp(-\rho_{k,i^\star} N_0 W x). \tag{45}$$

Therefore, the PMF of departure packet number is given by

$$\Pr\left[D^{\Pi}_{t,k} = d \middle| \mathcal{L}^{\mathrm{o}}_t\right] = \Pr\left[D^{\Pi}_{t,k} \le d+1 \middle| \mathcal{L}^{\mathrm{o}}_t\right] - \Pr\left[D^{\Pi}_{t,k} \le d \middle| \mathcal{L}^{\mathrm{o}}_t\right] \tag{46}$$

$$= \Pr\left[Y_{t,k}(\mathcal{L}^{\mathrm{o}}_t) \le \frac{2^{\wedge}\left(\frac{(d+1)N_{\mathrm{b}}}{W\tau^{\Pi}_{t,k}}\right)-1}{P^{\Pi}}\right] - \Pr\left[Y_{t,k}(\mathcal{L}^{\mathrm{o}}_t) \le \frac{2^{\wedge}\left(\frac{dN_{\mathrm{b}}}{W\tau^{\Pi}_{t,k}}\right)-1}{P^{\Pi}}\right]. \tag{47}$$

## B Proof of Lemma 2

In fact, $\mathbf{M}^{(\ell)}_k$ represents the transition probability matrix of local abstract state with $\mathcal{L}^{\mathrm{o}}_t$ eliminated, i.e., $(Q_{t,k}, A^{\mathrm{s}}_{t,k}, A^{\mathrm{d}}_{t,k})$, conditioned on $\varsigma(\mathcal{L}^{\mathrm{o}}_t) = \ell$. Then the derivation of (30) is straightforward. We have the following discussion on all possible cases for $\mathbf{M}^{(\ell)}_k$ defined in Table 2.

- **Case 1** $(Q_{t,k} = 0, 1 \le A^{\mathrm{s}}_{t,k} \le A_{\max}, 1 \le A^{\mathrm{d}}_{t,k} \le A_{\max}, Q_{t+1,k} = 0, A^{\mathrm{s}}_{t+1,k} = 1, A^{\mathrm{d}}_{t+1,k} = 1)$: This means that sampling and transmission of all packets of the new sample are accomplished with the probability of $Q^{\max}_k$ departure packets.
- **Case 2** $(Q_{t,k} = 0, 1 \le A^{\mathrm{s}}_{t,k} \le A_{\max}, 1 \le A^{\mathrm{d}}_{t,k} \le A_{\max}, 1 \le Q_{t+1,k} \le Q^{\max}_k, A^{\mathrm{s}}_{t+1,k} = 1, A^{\mathrm{d}}_{t+1,k} = \min\{A^{\mathrm{d}}_{t,k}+1, A_{\max}\})$: This means that sampling and transmission of $(Q^{\max}_k - Q_{t+1,k})$ packets of the new sample are accomplished with the probability of $(Q^{\max}_k - Q_{t+1,k})$ departure packets.
- **Case 3** $(1 \le Q_{t,k} \le Q^{\max}_k, 1 \le A^{\mathrm{s}}_{t,k} \le A_{\max}, 1 \le A^{\mathrm{d}}_{t,k} \le A_{\max}, Q_{t+1,k} = 0, A^{\mathrm{s}}_{t+1,k} = \min\{A^{\mathrm{s}}_{t,k}+1, A_{\max}\}, A^{\mathrm{d}}_{t+1,k} = \min\{A^{\mathrm{s}}_{t,k}+1, A_{\max}\})$: This means that transmission of all remaining $Q_{t+1,k}$ packets of the current sample is accomplished with the probability of $Q_{t+1,k}$ departure packets.
- **Case 4** $(1 \le Q_{t,k} \le Q^{\max}_k, 1 \le A^{\mathrm{s}}_{t,k} \le A_{\max}, 1 \le A^{\mathrm{d}}_{t,k} \le A_{\max}, 1 \le Q_{t+1,k} \le Q_{t,k}, A^{\mathrm{s}}_{t+1,k} = \min\{A^{\mathrm{s}}_{t,k}+1, A_{\max}\}, A^{\mathrm{d}}_{t+1,k} = \min\{A^{\mathrm{d}}_{t,k}+1, A_{\max}\})$: This means that transmission of $(Q_{t,k} - Q_{t+1,k})$ packets of the current sample is accomplished with the probability of $(Q_{t,k} - Q_{t+1,k})$ departure packets.

## C Proof of Theorem 1

The approximate value function is given by

$$W^{\Pi}(\widetilde{\mathcal{S}}_t) = \lim_{T\to\infty} \sum_{t=1}^{T} \gamma^{t-1} \left[ \sum_{k\in\mathcal{K}} \left( \mathbf{e}_{\kappa\left(\varsigma(\mathcal{L}_t^{\mathrm{o}}),\epsilon(Q_{t,k},A_{t,k}^{\mathrm{s}},A_{t,k}^{\mathrm{d}})\right)}^{L^B(Q_k^{\max}+1)A_{\max}^2} \right)^{\mathsf{T}} \mathbf{P}_k^{t-1}\mathbf{g}_k + w_{\mathrm{P}} T_{\mathrm{F}} P^{\Pi} \right]. \tag{48}$$

Since the reference policy adopts constant transmission time and power allocations, the per-frame cost for transmission power consumption, and thus its discounted cumulative sum corresponding to the second term in (31), results in a constant. In fact, $\mathbf{e}_{\kappa(\varsigma(\mathcal{L}_t^{\mathrm{o}}),\epsilon(Q_{t,k},A_{t,k}^{\mathrm{s}},A_{t,k}^{\mathrm{d}}))}^{L^B(Q_k^{\max}+1)A_{\max}^2}$ represents the probability vector for a deterministic local abstract state $\widetilde{\mathcal{S}}_{t,k} = (\mathcal{L}_t^{\mathrm{o}}, Q_{t,k}, A_{t,k}^{\mathrm{s}}, A_{t,k}^{\mathrm{d}})$, $\mathbf{P}_k$ represents the transition probability matrix, and the $\kappa(\varsigma(\mathcal{L}_t^{\mathrm{o}}),\epsilon(Q_{t,k},A_{t,k}^{\mathrm{s}},A_{t,k}^{\mathrm{d}}))$-th entry of $\mathbf{g}_k$ represents the per-frame cost for sampling energy cost, and AoIs and outdated AoI penalties at the server. Since the reference policy samples only when the queue becomes empty, the per-frame cost for sampling is counted only for $Q_{t,k} = 0$. Therefore, $(\mathbf{e}_{\kappa(\varsigma(\mathcal{L}_t^{\mathrm{o}}),\epsilon(Q_{t,k},A_{t,k}^{\mathrm{s}},A_{t,k}^{\mathrm{d}}))}^{L^B(Q_k^{\max}+1)A_{\max}^2})^{\mathsf{T}}\mathbf{P}_k^{t-1}\mathbf{g}_k$ represents the expected per-frame cost in the $t$-th frame. The derivation from (48) to (31) resembles the proof in Appendix B(3) of [11].

## D Proof of Lemma 3

Substituting $p_{t,k}(\tau_{t,k})$ by (36), P2.3($n$) can be rewritten by

$$\mathrm{P2.3}(n): \ \left\{\tau_{t,k}^{(n)}\right\}_{k\in\mathcal{K}} = \underset{\{\tau_{t,k}\}_{k\in\mathcal{K}}}{\arg\min} \sum_{k\in\mathcal{K}} \frac{\tau_{t,k}}{Y_{t,k}} \left[ 2^{\wedge}\left( \frac{d_{t,k}^{(n)} N_{\mathrm{b}}}{\tau_{t,k} W} \right) - 1 \right] \tag{49a}$$

$$\text{s.t.} \quad \tau_{t,k}^{(n)} \geq \frac{d_{t,k}^{(n)} N_b}{W \log_2(1 + P_{\max} Y_{t,k})}, \ \forall k \in \mathcal{K} \tag{49b}$$

$$\sum_{k\in\mathcal{K}} \tau_{t,k} = T_{\mathrm{F}}. \tag{49c}$$

By introducing Lagrange multipliers $\mu_k^{\star}$ for inequality constraints (49b) and $\nu^{\star}$ for equality constraint (49c), the Lagrangian is given by

$$\begin{aligned} &L\left(\{\tau_{t,k}\}_{k\in\mathcal{K}}, \{\mu_k\}_{k\in\mathcal{K}}, \nu\right) \\ &= \sum_{k\in\mathcal{K}} \frac{\tau_{t,k}}{Y_{t,k}} \left[ 2^{\wedge}\left( \frac{d_{t,k}^{(n)} N_{\mathrm{b}}}{\tau_{t,k} W} \right) - 1 \right] - \sum_{k\in\mathcal{K}} \mu_k \left( \tau_{t,k} - \frac{d_{t,k}^{(n)} N_b}{W \log_2(1 + P_{\max} Y_{t,k})} \right) + \nu \left( \sum_{k\in\mathcal{K}} \tau_{t,k} - T_{\mathrm{F}} \right). \end{aligned} \tag{50}$$

Hence, the ***Karush-Kuhn-Tucker*** (**KKT**) conditions can be written as follows:

$$\tau_{t,k}^{(n)} \geq \frac{d_{t,k}^{(n)} N_{\mathrm{b}}}{W \log_2(1 + P_{\max} Y_{t,k})}, \qquad \forall k \in \mathcal{K}, \tag{51a}$$

$$\sum_{k\in\mathcal{K}} \tau_{t,k}^{(n)} = T_{\mathrm{F}}, \tag{51b}$$

$$\mu_k^{\star} \geq 0, \qquad \forall k \in \mathcal{K}, \tag{51c}$$

$$\mu_k^{\star} \tau_{t,k}^{(n)} = 0, \qquad \forall k \in \mathcal{K}, \tag{51d}$$

$$\frac{\partial L}{\partial \tau_{t,k}^{(n)}} = \frac{1}{Y_{t,k}} \left[ \left( 1 - \frac{d_{t,k}^{(n)} N_{\mathrm{b}} \ln 2}{W \tau_{t,k}^{(n)}} \right) 2^{\wedge}\left( \frac{d_{t,k}^{(n)} N_{\mathrm{b}}}{W \tau_{t,k}^{(n)}} \right) - 1 \right] - \mu_k^{\star} + \nu^{\star} = 0, \quad \forall k \in \mathcal{K}. \tag{51e}$$

Since P2.3($n$) is a convex problem, by solving the above equations, proof of lemma 3 is straightforward.

## References


[1] Ahmed Alkhateeb, Omar El Ayach, Geert Leus, and Robert W. Heath. 2014. Channel estimation and hybrid precoding for millimeter wave cellular systems. *IEEE Journal of Selected Topics in Signal Processing* 8, 5 (2014), 831–846. DOI: https://doi.org/10.1109/JSTSP.2014.2334278

[2] Mario Alonzo, Stefano Buzzi, Alessio Zappone, and Ciro D'Elia. 2019. Energy-efficient power control in cell-free and user-centric massive mimo at millimeter wave. *IEEE Transactions on Green Communications and Networking* 3, 3 (2019), 651–663. DOI: https://doi.org/10.1109/TGCN.2019.2908228

[3] Omar El Ayach, Sridhar Rajagopal, Shadi Abu-Surra, Zhouyue Pi, and Robert W. Heath. 2014. Spatially sparse precoding in millimeter wave MIMO systems. *IEEE Transactions on Wireless Communications* 13, 3 (2014), 1499–1513. DOI: https://doi.org/10.1109/TWC.2014.011714.130846

[4] Dimitri Bertsekas. 2012. *Dynamic Programming and Optimal Control (4 ed.).* Vol. 2. Athena scientific, Belmont, MA, USA.

[5] Fadhil Firyaguna, Andrea Bonfante, Jacek Kibiłda, and Nicola Marchetti. 2020. Performance evaluation of scheduling in 5G-mmWave networks under human blockage. arXiv:2007.13112. Retrieved from https://arxiv.org/abs/2007.13112

[6] Fadhil Firyaguna, Jacek Kibilda, Carlo Galiotto, and Nicola Marchetti. 2021. Performance analysis of indoor mmwave networks with ceiling-mounted access points. *IEEE Transactions on Mobile Computing* 20, 5 (2021), 1940–1950. DOI: https://doi.org/10.1109/TMC.2020.2972282

[7] Margarita Gapeyenko, Andrey Samuylov, Mikhail Gerasimenko, Dmitri Moltchanov, Sarabjot Singh, Mustafa Riza Akdeniz, Ehsan Aryafar, Nageen Himayat, Sergey Andreev, and Yevgeni Koucheryavy. 2017. On the temporal effects of mobile blockers in urban millimeter-wave cellular scenarios. *IEEE Transactions on Vehicular Technology* 66, 11 (2017), 10124–10138. DOI: https://doi.org/10.1109/TVT.2017.2754543

[8] Leonidas Georgiadis, Michael J. Neely, and Leandros Tassiulas. 2006. Resource allocation and cross-layer control in wireless networks. *Foundations and Trends® in Networking* 1, 1 (2006), 1–144.

[9] Xingqiu He, Sheng Wang, Xiong Wang, Shizhong Xu, and Jing Ren. 2022. Age-based scheduling for monitoring and control applications in mobile edge computing systems. In *Proceedings of the IEEE INFOCOM 2022 - IEEE Conference on Computer Communications*. 1009–1018. DOI: https://doi.org/10.1109/INFOCOM48880.2022.9796654

[10] Robert W. Heath, Nuria González-Prelcic, Sundeep Rangan, Wonil Roh, and Akbar M. Sayeed. 2016. An overview of signal processing techniques for millimeter wave MIMO systems. *IEEE Journal of Selected Topics in Signal Processing* 10, 3 (2016), 436–453. DOI: https://doi.org/10.1109/JSTSP.2016.2523924

[11] Shanfeng Huang, Bojie Lv, Rui Wang, and Kaibin Huang. 2020. Scheduling for mobile edge computing with random user Arrivals–An approximate MDP and reinforcement learning approach. *IEEE Transactions on Vehicular Technology* 69, 7 (2020), 7735–7750. DOI: https://doi.org/10.1109/TVT.2020.2990482

[12] Ish Kumar Jain, Rajeev Kumar, and Shivendra S. Panwar. 2019. The impact of mobile blockers on millimeter wave cellular systems. *IEEE Journal on Selected Areas in Communications* 37, 4 (2019), 854–868. DOI: https://doi.org/10.1109/JSAC.2019.2898756

[13] Sanjit Kaul, Marco Gruteser, Vinuth Rai, and John Kenney. 2011. Minimizing age of information in vehicular networks. In *Proceedings of the 2011 8th Annual IEEE Communications Society Conference on Sensor, Mesh and Ad Hoc Communications and Networks*. 350–358. DOI: https://doi.org/10.1109/SAHCN.2011.5984917

[14] Lihong Li, Thomas J. Walsh, and Michael L. Littman. 2006. Towards a unified theory of state abstraction for MDPs. In *Proceedings of the AI&M.*

[15] Fan Liu, Yuanhao Cui, Christos Masouros, Jie Xu, Tony Xiao Han, Yonina C. Eldar, and Stefano Buzzi. 2022. Integrated sensing and communications: Toward dual-functional wireless networks for 6G and beyond. *IEEE Journal on Selected Areas in Communications* 40, 6 (2022), 1728–1767. DOI: https://doi.org/10.1109/JSAC.2022.3156632

[16] Zhongdong Liu, Keyuan Zhang, Bin Li, Yin Sun, Y. Thomas Hou, and Bo Ji. 2024. Learning-augmented Online Minimization of Age of Information and Transmission Costs. arXiv:2403.02573. Retrieved from https://arxiv.org/abs/2403.02573

[17] Bojie Lv and Rui Wang. 2024. Delay-aware two-time-scale scheduling for mmWave systems with mobility and environment knowledge. *IEEE Transactions on Communications* 72, 7 (2024), 4026–4040. DOI: https://doi.org/10.1109/TCOMM.2024.3366390

[18] A. Maltsev, V. Erceg, E. Perahia, C. Hansen, R. Maslennikov, A. Lomayev, A. Sevastyanov, A. Khoryaev, G. Morozov, M. Jacob, S. Priebe, T. Kürner, S. Kato, H. Sawada, K. Sato, and H. Harada. 2010. *Channel Models for 60 GHz WLAN Systems, doc.: IEEE 802.11-09/0334r8.* Institute of Electrical, Piscataway, NJ, USA.

[19] Brendan Morris and Mohan Trivedi. 2009. Learning trajectory patterns by clustering: Experimental studies and comparative evaluation. In *Proceedings of the 2009 IEEE Conference on Computer Vision and Pattern Recognition*. 312–319. DOI: https://doi.org/10.1109/CVPR.2009.5206559

[20] Linh Nguyen, Sarath Kodagoda, Ravindra Ranasinghe, and Gamini Dissanayake. 2021. Mobile robotic sensors for environmental monitoring using Gaussian Markov random field. *Robotica* 39, 5 (2021), 862–884.

[21] Yong Niu, Weiguang Ding, Hao Wu, Yong Li, Xinlei Chen, Bo Ai, and Zhangdui Zhong. 2019. Relay-assisted and QoS aware scheduling to overcome blockage in mmWave backhaul networks. *IEEE Transactions on Vehicular Technology* 68, 2 (2019), 1733–1744. DOI: https://doi.org/10.1109/TVT.2018.2890308

[22] Joan Palacios, Paolo Casari, and Joerg Widmer. 2017. JADE: Zero-knowledge device localization and environment mapping for millimeter wave systems. In *Proceedings of the IEEE INFOCOM 2017 - IEEE Conference on Computer Communications*. 1–9. DOI: https://doi.org/10.1109/INFOCOM.2017.8057183

[23] Theodore S. Rappaport, Shu Sun, Rimma Mayzus, Hang Zhao, Yaniv Azar, Kevin Wang, George N. Wong, Jocelyn K. Schulz, Mathew Samimi, and Felix Gutierrez. 2013. Millimeter wave mobile communications for 5G cellular: It will work!*IEEE Access* 1 (2013), 335–349. DOI: https://doi.org/10.1109/ACCESS.2013.2260813

[24] A. M. Sayeed. 2002. Deconstructing multiantenna fading channels. *IEEE Transactions on Signal Processing* 50, 10 (2002), 2563–2579. DOI: https://doi.org/10.1109/TSP.2002.803324

[25] Kinza Shafique, Bilal A. Khawaja, Farah Sabir, Sameer Qazi, and Muhammad Mustaqim. 2020. Internet of things (IoT) for next-generation smart systems: A review of current challenges, future trends and prospects for emerging 5G-IoT scenarios. *IEEE Access* 8 (2020), 23022–23040. DOI: https://doi.org/10.1109/ACCESS.2020.2970118

[26] Marzieh Sheikhi and Vesal Hakami. 2021. AoI-aware status update control for an energy harvesting source over an uplink mmWave channel. In *Proceedings of the 2021 7th International Conference on Signal Processing and Intelligent Systems*. 01–06. DOI: https://doi.org/10.1109/ICSPIS54653.2021.9729335

[27] Yin Sun and Benjamin Cyr. 2019. Sampling for data freshness optimization: Non-linear age functions. *Journal of Communications and Networks* 21, 3 (2019), 204–219. DOI: https://doi.org/10.1109/JCN.2019.000035

[28] Yifei Sun, Jie Li, Tong Zhang, Rui Wang, Xiaohui Peng, Xiao Han, and Haisheng Tan. 2022. An indoor environment sensing and localization system via mmWave phased array. *Journal of Communications and Information Networks* 7, 4 (2022), 383–393. DOI: https://doi.org/10.23919/JCIN.2022.10005216

[29] Yifei Sun, Bojie Lv, Rui Wang, Haisheng Tan, and Francis C. M. Lau. 2024. Predictive delay-aware scheduling with receiver rotation detection and mmWave channel learning. *IEEE Transactions on Wireless Communications* (2024), 1–1. DOI: https://doi.org/10.1109/TWC.2024.3400377

[30] Yifei Sun, Bojie Lv, Rui Wang, Haisheng Tan, and Francis C. M. Lau. 2023. Dynamic uploading scheduling in mmWave-based sensor networks via mobile blocker detection. In *Proceedings of the 2023 IEEE 29th International Conference on Parallel and Distributed Systems*. 2057–2066. DOI: https://doi.org/10.1109/ICPADS60453.2023.00280

[31] Yifei Sun, Bojie Lv, Rui Wang, Haisheng Tan, and Francis C. M. Lau. 2023. Predictive resource allocation in mmWave systems with rotation detection. In *Proceedings of the ICC 2023 - IEEE International Conference on Communications*. 2753–2759. DOI: https://doi.org/10.1109/ICC45041.2023.10278584

[32] Dawei Wang, Menghan Wu, Chinmay Chakraborty, Lingtong Min, Yixin He, and Manisha Guduri. 2024. Covert communications in air-ground integrated urban sensing networks enhanced by federated learning. *IEEE Sensors Journal* 24, 5 (2024), 5636–5643. DOI: https://doi.org/10.1109/JSEN.2023.3322784

[33] Yibing Wang, Yong Niu, Hao Wu, Bo Ai, Zhangdui Zhong, Dapeng Oliver Wu, and Tutun Juhana. 2019. Relay assisted concurrent scheduling to overcome blockage in full-duplex millimeter wave small cells. *IEEE Access* 7 (2019), 105755–105767. DOI: https://doi.org/10.1109/ACCESS.2019.2931876

[34] Teng Wei, Anfu Zhou, and Xinyu Zhang. 2017. Facilitating robust 60 GHz network deployment by sensing ambient reflectors. In *Proceedings of the 14th USENIX Symposium on Networked Systems Design and Implementation*. 213–226.

[35] Xianwen Wu, Jing Yang, and Jingxian Wu. 2018. Optimal status update for age of information minimization with an energy harvesting source. *IEEE Transactions on Green Communications and Networking* 2, 1 (2018), 193–204. DOI: https://doi.org/10.1109/TGCN.2017.2778501

[36] Liu Yang, Keping Yu, Simon X. Yang, Chinmay Chakraborty, Yinzhi Lu, and Tan Guo. 2022. An intelligent trust cloud management method for secure clustering in 5G enabled internet of medical things. *IEEE Transactions on Industrial Informatics* 18, 12 (2022), 8864–8875. DOI: https://doi.org/10.1109/TII.2021.3128954

[37] Candy Yiu and Suresh Singh. 2009. Empirical capacity of mmWave WLANS. *IEEE Journal on Selected Areas in Communications* 27, 8 (2009), 1479–1487. DOI: https://doi.org/10.1109/JSAC.2009.091017

[38] Ronghui Zhang, Xiaojun Jing, Sheng Wu, Chunxiao Jiang, Junsheng Mu, and F. Richard Yu. 2021. Device-free wireless sensing for human detection: The deep learning perspective. *IEEE Internet of Things Journal* 8, 4 (2021), 2517–2539. DOI: https://doi.org/10.1109/JIOT.2020.3024234

[39] Xi Zheng, Sheng Zhou, Zhiyuan Jiang, and Zhisheng Niu. 2019. Closed-form analysis of non-linear age of information in status updates with an energy harvesting transmitter. *IEEE Transactions on Wireless Communications* 18, 8 (2019), 4129–4142. DOI: https://doi.org/10.1109/TWC.2019.2921372

[40] Bo Zhou and Walid Saad. 2018. Optimal sampling and updating for minimizing age of information in the internet of things. In *Proceedings of the 2018 IEEE Global Communications Conference*. 1–6. DOI: https://doi.org/10.1109/GLOCOM.2018.8647281

[41] Bo Zhou and Walid Saad. 2019. Joint status sampling and updating for minimizing age of information in the internet of things. *IEEE Transactions on Communications* 67, 11 (2019), 7468–7482. DOI: https://doi.org/10.1109/TCOMM.2019.2931538

[42] Bo Zhou and Walid Saad. 2020. Minimum age of information in the internet of things with non-uniform status packet sizes. *IEEE Transactions on Wireless Communications* 19, 3 (2020), 1933–1947. DOI: https://doi.org/10.1109/TWC.2019.2959777

Received 14 February 2024; revised 22 July 2024; accepted 17 September 2024