Joint Sparse Observation and Coding Design for Multiple Phenomena Monitoring

Chengcheng Han, Li Chen, Nan Zhao, *Senior Member, IEEE,* Yunfei Chen, *Senior Member, IEEE,* and F. Richard Yu, *Fellow, IEEE*

***Abstract*—Energy-efficient designs play an important role in the Internet of Things (IoT) that monitors multiple phenomena, due to the limited power supply and complicated observation. In this paper, taking into account the power consumptions of observation, coding, and communication, we propose a joint sparse observation and coding scheme for energy-efficient monitoring of multiple phenomena using IoT. Through the analysis of outage performance, we find that the sparse observation and coding scheme can achieve the performance of the full observation scheme in which all nodes observe all phenomena with lower power consumption due to the dynamic and selective observation and coding. With the derived achievable rates and network power consumption, we study the trade-off between achievable rates and network power consumption that is determined by both the observation matrix and the coding matrix. For given rate constraints, we propose an optimization problem to minimize the network power consumption by jointly designing the observation and coding matrices. To solve this NP-hard problem efficiently, we propose a low-complexity algorithm with the convex-concave procedure. Moreover, to improve performance in high noise environment, we adopt collaboration among nodes to suppress observation noises and equalize bad observations by utilizing observation diversity. Finally, simulation results illustrate the superior performance of the proposed schemes.**

***Index Terms*—Energy efficiency, multiple phenomena monitoring, observation diversity, sparse observation and coding.**

I. Introduction

The Internet of Things (IoT) with a huge number of battery-powered nodes has been widely utilized to monitor diverse phenomena [1]–[5]. Due to the limited lifetime of battery, it is extremely important to investigate energy-efficient schemes for IoT.

To improve the energy efficiency of IoT, node selection has been widely used in the context of phenomenon observation, where a subset of nodes is selected to transmit information to the fusion center (FC) according to a certain metric. With the help of auxiliary Boolean variables, the problem of node selection for sensing was studied in [6], where the Boolean variables can determine whether or not its corresponding node is selected. With the limited energy budget, the authors of [7] extended one-step node selection schemes into a multi-step node selection scheme to schedule the transmission of observation from nodes to the FC to minimize estimation error. For a better trade-off between information accuracy and energy consumption, the optimal node selection scheme was provided by optimizing the energy efficiency of nodes for a given information utility [8], [9]. Considering the main energy consumption of IoT, the work of [10] studied the optimal node selection for data transmission with energy conservation. Utilizing energy harvesting as the power supply, the node selection approach was adopted in [11] for heterogeneous sensor networks to achieve energy-efficient spatial field reconstruction. To improve energy efficiency, grouping the selected nodes into multiple clusters is an effective method [12]–[15], which can maintain a longer life of distributed nodes by reducing the power loss in propagation and retransmission.

Introducing collaboration into multi-cluster IoT is widely regarded as a beneficial approach to improve energy efficiency. A universal framework of multi-cluster IoT with collaboration was studied in [16], where the nodes within the same cluster performed collaboration for their observations. It was shown that transmitting the collaborative observations over the best available channels and assigning the power levels matched to channel quality can use the minimum energy consumption to achieve the minimum estimation error. This work ignored collaboration costs. Considering the collaboration cost in practice, the collaborative IoT to estimate phenomenon was further studied, where the nodes were selected to update their observations by collaborating with their adjacent nodes [17]. The optimal selection and collaboration strategy was designed to minimize the expected mean square error (MSE) subject to the given power constraints. With the error-free and low-cost collaboration links, [18] provided a power allocation scheme of distributed IoT by optimizing the collaboration coefficients of nodes. The reason that collaboration can improve the energy efficiency of IoT is that it can utilize the spatial diversity of observation and transmission.

The energy efficiency of IoT can be further improved by sparsely optimizing the selection and collaboration of nodes. The energy-optimal collaboration scheme of IoT was studied in [19], where only a subset of nodes was permitted to collaborate for their observations. This optimal collaboration strategy indicates that a partially collaborating network can yield performance close to that of a fully collaborating network

This research was supported by National Natural Science Foundation of China (Grant No. 62071445), and the Fundamental Research Funds for the Central Universities (Grant No. WK3500000007 and Grant No. YD3500002001) (*Corresponding author: Li Chen*).

L. Chen and C. Han are with CAS Key Laboratory of Wireless-Optical Communications, University of Science and Technology of China, Hefei 230052, China (e-mail: chenli87@ustc.edu.cn;hancheng@mail.ustc.edu.cn).

N. Zhao is with the School of Info. and Commun. Eng., Dalian University of Technology, Dalian, China (email: zhaonan@dlut.edu.cn).

Y. Chen is with the School of Engineering, University of Warwick, Coventry CV4 7AL, U.K. (e-mail: Yunfei.Chen@warwick.ac.uk).

F.R. Yu is with the Department of Systems and Computer Engineering, Carleton University, Ottawa, ON, K1S 5B6, Canada (email: richard.yu@carleton.ca).

yet consumes much less energy. In the study of [20], a sparsity-aware node selection problem was formulated where the number of selected nodes was minimized subject to a certain estimation quality. The authors of [21] further explored the above sparsity-aware node selection problem and proposed a relaxed sparsity-aware node selection approach with low-complexity. With the help of promoting sparsity, the design of node selection scheme in IoT was further studied in [22]–[24], which reduced energy consumption. Ignoring the cost of selection and collaboration, the authors in [25] formulated the design of node selection and collaboration schemes as the problem of recovering a sparse matrix. Incorporating the cost of selection and collaboration, the authors in [26] minimized the total energy consumption subject to a given quality-of-service (QoS) requirement in a sparsely optimized node selection and collaboration scheme.

Most of the aforementioned works only reduce the power consumption of communication, while the study in [27] shows that the power consumption of observation is roughly comparable with that of communication in energy-consuming IoT. The power consumption of coding in IoT is also counted into the network power consumption [28]. In fact, in the IoT monitoring multiple phenomena, observation and coding consume more power than communication due to the complicated observation and coding.

The above discussion implies that the power consumptions of observation, coding, and communication should be fully considered to achieve an energy-efficient IoT that monitors multiple phenomena. However, this brings the following challenges. First, the full observation scheme requiring all nodes to observe all phenomena may no longer be optimal for IoT due to the huge power consumption in observation and coding. Also, the power consumptions of observation, coding, and communication jointly determine the performance of IoT, which leads to a more complicated trade-off between observation performance and power consumption. To cope with the first challenge, in this paper, we propose an energy-efficient scheme for IoT to monitor multiple phenomena, which utilizes the sparse observation and coding method. In this method, nodes can dynamically choose any part of the phenomena for observation and coding. To tackle the second challenge, for given observation performance requirements, we propose an optimization problem to minimize the network power consumption by jointly designing the observed relation of nodes to phenomena and the coding matrix of observations. Furthermore, to improve the observation performance in high noise environment, we study the collaboration among nodes and propose a sparse observation and collaborative coding scheme, which can utilize the observation diversity to suppress the observation noise and equalize the bad observation. The main contributions of this paper are summarized as follows.

1) **Sparse observation and coding scheme**. We propose a new sparse observation and coding scheme for IoT to achieve the energy-efficient observation of multiple phenomena, where the nodes can dynamically choose any part of the phenomena for observation and coding. Through the analysis of outage performance, we find that the sparse observation and coding scheme can achieve the performance of the full observation scheme in which all nodes observe all phenomena with lower power consumption due to the dynamic and selective observation and coding. With the derived achievable rates and network power consumption, we find that there is a trade-off between achievable performance and power consumption that is determined by both the observation matrix and the coding matrix. Specifically, activating more nodes to observe and encode more phenomena can improve the achievable rates of phenomena but leads to an increase in network power consumption.
2) **Joint optimization of observation and coding matrices**. To achieve the optimal performance trade-off, for given rate constraints, we formulate a joint optimization problem of observation and coding matrices to minimize the network power consumption. This optimization problem is intractable due to the coupled observation and coding matrices, the discontinuous and non-differentiable l_0-norm, and the non-convex objective function and constraints. To solve it effectively, we first decouple the coding matrix from the observation matrix and reformulate an equivalent sparse coding optimization problem. Then, we approximate the discontinuous and non-differentiable l_0-norm with the concave smooth function and obtain a continuous non-convex optimization problem. Furthermore, we relax the rate constraints and equivalently introduce the auxiliary variables to transform the non-convex optimization problem into a solvable difference of convex (DC) programming problem. Finally, an algorithm with the convex-concave procedure (CCP) is proposed to provide an effective solution to the original problem.
3) **Collaboration among nodes and optimization**. To further improve the performance, we study the collaboration among nodes and propose a sparse observation and collaborative coding scheme, where the nodes observing the same phenomenon are required to share their observations. The collaboration utilizes the observation diversity of nodes to suppress the observation noise and equalize the bad observation, which improves the SINRs of observed phenomena. The improvement of SINRs is at the expense of increased power consumption and shortened transmission duration, which restricts the achievable rates of observed phenomena. Consequently, collaboration has a complicated effect on the trade-off between achievable rates and network power consumption. To achieve the optimal trade-off, we formulate an NP-hard problem of minimizing the network power consumption with given rate constraints. To solve it effectively, a CCP based algorithm with low-complexity is proposed, which requires to solve $\mathcal{O}\left(\log\left(\epsilon_1^{-1}\right)(\theta_0-\epsilon)\beta^{-1}\right)$ convex problems.

The rest of this paper is organized as follows. In Section II, the system model of IoT with sparse observation and coding scheme is proposed. Section III Section provides the outage performance analysis and derives the achievable rate. IV studies the optimization problem to minimize the network power consumption for given rate constraints. Section V introduces collaboration into the sparse observation and coding scheme

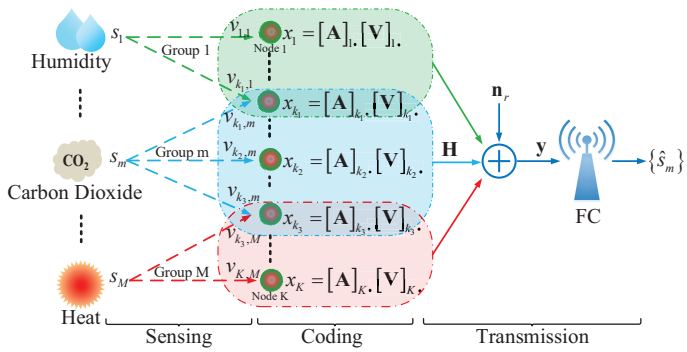


Figure 1: The sparse observation and coding scheme in IoT.

and optimizes the performance. In Section VI, simulation results are presented to illustrate the superior performance of the proposed schemes, followed by the conclusion in Section VII.

Notation: $\mathbb{E}\{\cdot\}$ and $\Re\{\cdot\}$ denote the expectation and real part of variables, respectively. The Hadamard product between vector $\mathbf{a}$ and $\mathbf{b}$ is denoted by $\mathbf{a}\circ\mathbf{b}$. Given a matrix $\mathbf{A}$, $\mathbf{A}^H$, $\mathbf{A}^T$, and $\mathrm{rank}(\mathbf{A})$ denote the conjugate transposition, the transposition, and the rank of $\mathbf{A}$, respectively. The matrix with K rows and M columns is denoted by $\mathbf{A}_{K\times M}$, and its k-th row vector and m-th column vector are denoted by $[\mathbf{A}]_{k\cdot}$ and $[\mathbf{A}]_{\cdot m}$, respectively. The cardinality of set $\mathcal{A}$ is denoted by $|\mathcal{A}|$. For a square matrix $\mathbf{A}$, $\mathbf{A}\succeq\mathbf{0}$ means that $\mathbf{A}$ is the positive semi-definite. The $a\sim\mathcal{CN}\left(\mu,\sigma_o^2\right)$ indicates that a is a circularly symmetric complex Gaussian random variable with mean μ and variance σ_o^2.

II. System Model

In this section, we first present the network model, followed by the power consumption model.

A. Network Model

Consider an IoT system based on sparse observation and coding scheme to monitor multiple phenomena $\mathbf{s}=[s_1,\cdots,s_M]^T$, as illustrated in Fig. 1. This IoT system consists of K nodes with single-antenna and one FC with N antennas. During each observation, nodes can dynamically choose any part of the phenomena $\mathbf{s}$ for observation. According to the observed relation between nodes and phenomena, K nodes naturally form M overlapped groups, denoted by $\mathcal{Q}=\{\mathcal{Q}_1,\cdots,\mathcal{Q}_M\}$ where the nodes belonging to group $\mathcal{Q}_m$ need to observe phenomenon s_m. The nodes belonging to multiple groups are required to linearly encode the observation of different phenomena and concurrently transmit the coded observation to the FC for recovery.

The IoT based on sparse observation and coding scheme consists of three phases: sensing, coding, and transmission. Specifically, in the sensing phase, all nodes independently select part of the phenomena for observation. The observation is disturbed by the additive white Gaussian noise (AWGN). The observation of phenomenon s_m at node k is given as

$$v_{k,m}=c_{k,m}\left(s_m+n_{k,m}\right),k\in\mathcal{K},m\in\mathcal{M}, \tag{1}$$

where $c_{k,m}\in\{0,1\}$ represents the observed relation between node k and phenomenon s_m, $c_{k,m}=1$ indicates that node k observes s_m and $c_{k,m}=0$ otherwise, $n_{k,m}$ is the observation noise of s_m at node k and satisfies $n_{k,m}\sim\mathcal{CN}\left(0,\sigma_o^2\right)$, $\mathcal{K}=\{1,\cdots,K\}$, $\mathcal{M}=\{1,\cdots,M\}$. The observation of phenomena at all nodes can be given by

$$\mathbf{V}=\mathbf{C}\circ\left(\mathbf{S}^T+\mathbf{N}\right), \tag{2}$$

where $\mathbf{C}_{K\times M}$ represents the observed relation of K nodes to M phenomena and is termed as the observation matrix, $\mathbf{S}_{M\times K}=[\mathbf{s},\cdots,\mathbf{s}]$, and $\mathbf{N}_{K\times M}$ represents the independent and identically distributed (i.i.d) AWGN in the observation of M phenomena at K nodes.

In the coding phase, the observations of different phenomena are linearly encoded at nodes. Each node encodes the observation of different phenomena and produces a coded scalar observation by multiplying a coding vector with length M. The coded observation of node k is given by

$$x_k=[\mathbf{A}]_{k\cdot}[\mathbf{V}]_{k\cdot}^T=[\mathbf{A}]_{k\cdot}\left([\mathbf{C}]_{k\cdot}^T\circ\left(\mathbf{s}+[\mathbf{N}]_{k\cdot}^T\right)\right),k\in\mathcal{K}, \tag{3}$$

where $\mathbf{A}_{K\times M}$ represents the coding matrix whose k-th row vector $[\mathbf{A}]_{k\cdot}$ is the coding vector of node k. The coded observation of all nodes is given as

$$\begin{aligned}\mathbf{x}=&\begin{bmatrix}[\mathbf{A}]_{1\cdot}\left([\mathbf{C}]_{1\cdot}^T\circ\left(\mathbf{s}+[\mathbf{N}]_{1\cdot}^T\right)\right)\\ \vdots\\ [\mathbf{A}]_{K\cdot}\left([\mathbf{C}]_{K\cdot}^T\circ\left(\mathbf{s}+[\mathbf{N}]_{K\cdot}^T\right)\right)\end{bmatrix}\\ =&\begin{bmatrix}[\mathbf{A}]_{1\cdot}\circ[\mathbf{C}]_{1\cdot}\\ \vdots\\ [\mathbf{A}]_{K\cdot}\circ[\mathbf{C}]_{K\cdot}\end{bmatrix}\mathbf{s}+\begin{bmatrix}\left([\mathbf{A}]_{1\cdot}\circ[\mathbf{C}]_{1\cdot}\right)[\mathbf{N}]_{1\cdot}^T\\ \vdots\\ \left([\mathbf{A}]_{K\cdot}\circ[\mathbf{C}]_{K\cdot}\right)[\mathbf{N}]_{K\cdot}^T\end{bmatrix}\\ =&\left(\mathbf{A}\circ\mathbf{C}\right)\mathbf{s}+\mathbf{n}_o,\end{aligned} \tag{4}$$

where $\mathbf{n}_o$ represents the coded observed noise and satisfies $\mathbf{n}_o\sim\mathcal{CN}\left(\mathbf{0},\mathbf{R}_1\right)$, $\mathbf{R}_1$ is a diagonal matrix whose k-th diagonal element is $\|[\mathbf{A}\circ\mathbf{C}]_{k\cdot}\|_2^2\sigma_o^2$.

In the transmission phase, all nodes concurrently transmit their coded observations to the FC for recovery. The received observation at the FC is given by

$$\mathbf{y}=\mathbf{H}\mathbf{x}+\mathbf{n}_r=\sum_{m=1}^{M}\mathbf{H}\left[\mathbf{A}\circ\mathbf{C}\right]_{\cdot m}s_m+\mathbf{H}\mathbf{n}_o+\mathbf{n}_r, \tag{5}$$

where $\mathbf{H}_{N\times K}$ represents the channels from K nodes to the FC that are the i.i.d Rayleigh block fading channels, $\mathbf{n}_r$ represents the received noise that is the i.i.d AWGN, i.e., $\mathbf{n}_r\sim\mathcal{CN}\left(\mathbf{0},\sigma^2\mathbf{I}\right)$. We assume that the channel state information (CSI) is known by the FC. With the known CSI, the FC can recover phenomena $\mathbf{s}$ by processing the received observation $\mathbf{y}$ through the linear minimum mean-square error (MMSE) receiver. Using the linear MMSE receiver, the estimate of phenomenon s_m $(m\in\mathcal{M})$ can be calculated as

$$\hat{s}_m=\mathbf{d}_m^H\left(\sum_{n\neq m}^{M}\mathbf{d}_n\mathbf{d}_n^H+\mathbf{H}\mathbf{R}_1\mathbf{H}^H+\sigma^2\mathbf{I}\right)^{-1}\mathbf{y}, \tag{6}$$

where $\mathbf{d}_m=\mathbf{H}\left[\mathbf{A}\circ\mathbf{C}\right]_{\cdot m},m\in\mathcal{M}$.

B. Power Consumption Model

The network power consumption of IoT is composed of the sensing power consumption, the coding power consumption, and the transmission power consumption. These three power consumptions are determined by the working mode of nodes, i.e., active mode and sleep mode [29]. For active nodes, the power is mainly consumed to observe phenomena, encode observations, transmit coded observations, and keep basic circuits operating. While the sleep nodes only consume a few power to keep necessary circuits operating. Using the empirical linear model in [30], the power consumption of node k can be given as

$$P_k = \begin{cases} P_k^o + P_k^a + P_k^t, & \text{active mode}, \\ P_k^{sl}, & \text{sleep mode}, \end{cases} \tag{7}$$

where P_k^o is the total power consumption of sensing and coding, P_k^a and P_k^{sl} are constant circuit power consumption when node k works in active and sleep mode respectively, and P_k^t is transmission power consumption. The power consumption of sensing and coding is proportional to the amount of data generated [29]. The total power consumption of sensing and coding at node k can be given by

$$P_k^o = \sum_{m=1}^{M} c_{k,m} p_m R_{g,m}, k \in \mathcal{K}, \tag{8}$$

where p_m is the total power consumption to generate and encode 1-bit observation of s_m. According to (3), the transmission power consumption of node k ($k \in \mathcal{K}$) is given by

$$\begin{aligned} P_k^t &= \mathbb{E}\left\{x_k x_k^H\right\} \\ &= \mathbb{E}\left\{[\mathbf{A}\circ\mathbf{C}]_{k\cdot}\left(\mathbf{s}+[\mathbf{N}]_{k\cdot}^T\right)\left(\mathbf{s}+[\mathbf{N}]_{k\cdot}^T\right)^H[\mathbf{A}\circ\mathbf{C}]_{k\cdot}^H\right\} \\ &= \left(1+\sigma_o^2\right)\|[\mathbf{A}\circ\mathbf{C}]_{k\cdot}\|_2^2. \end{aligned} \tag{9}$$

The network power consumption is given as

$$\begin{aligned} P = \sum_{k=1}^{K} P_k = &\sum_{m=1}^{M} \|[\mathbf{C}]_{\cdot m}\|_0 u_m + \sum_{k=1}^{K} \left\|\|[\mathbf{C}]_{k\cdot}\|_2^2\right\|_0 P_k^g \\ &+ \sum_{k=1}^{K} P_k^{sl} + \left(1+\sigma_o^2\right)\|\mathbf{A}\circ\mathbf{C}\|_2^2, \end{aligned} \tag{10}$$

where $u_m = p_m R_{g,m}$, $P_k^g = P_k^a - P_k^{sl}$.

III. Performance Analysis

In this section, we derive the outage probability and the achievable rate for the sparse observation and coding scheme.

A. Outage Probability

The signal-to-interference-plus-noise ratio (SINR) of phenomenon s_m after the linear MMSE receiver is given by

$$\text{SINR}_m = \mathbf{d}_m^H \left(\sum_{n\neq m}^{M} \mathbf{d}_n \mathbf{d}_n^H + \mathbf{H}\mathbf{R}_1\mathbf{H}^H + \sigma^2\mathbf{I}\right)^{-1} \mathbf{d}_m. \tag{11}$$

Definition 1. (Outage probability.) For the received SINR γ and its threshold γ_{T}, the outage probability can be defined by

$$P_{\text{o}} = \Pr\left(\gamma < \gamma_{\text{T}}\right), \tag{12}$$

which represents the probability that the QoS requirement of the observed phenomenon is unsatisfied.

In order to derive the outage probability, we need the probability density function (PDF) of the received SINR, which is derived in the following Lemma 1.

Lemma 1. (The PDF of SINR_m.) The PDF of SINR_m can be given by

$$\begin{aligned} f(\gamma) = &\frac{1}{\beta_m}\frac{\left(\frac{\gamma}{\beta_m}\right)^{N-1}\mathrm{e}^{-\frac{\rho\gamma}{\beta_m}}}{(N-1)!\left(1+\frac{\rho^N\gamma}{\beta_m}\right)^M}\left[\sum_{k=0}^{M-1}\binom{M-1}{k}\frac{N!}{(N-k)!}\rho^{-k}\right. \\ &\left.+\frac{\gamma}{\beta_m}\sum_{k=0}^{M-1}\binom{M-1}{k}\frac{(N-1)!}{(N-k-1)!}\rho^{-k}\right], \end{aligned} \tag{13}$$

where $\beta_m = \|[\mathbf{A}\circ\mathbf{C}]_{\cdot m}\|_2^2$, $\rho = \|\mathbf{A}\circ\mathbf{C}\|_2^2\sigma_o^2 + \sigma^2$, $N!$ is the factorial of the non-negative integer N, $\binom{M}{k}$ is the number of k-combinations from a given set with M elements.

Proof: Please refer to Appendix A. ■

Proposition 1. (Outage probability.) The outage probability can be given by

$$\begin{aligned} P_{\text{o},m} = &\sum_{k=0}^{M-1}\binom{M-1}{k}\frac{N\gamma_0^{k-N}}{(N-k)}\mathcal{I}\left(N, N-M+1, \gamma_0^{-1}, \frac{\gamma_{\text{T}}}{\beta_m}\right) \\ &+\sum_{k=0}^{M-1}\binom{M-1}{k}\frac{\gamma_0^{k-N}}{(N-k-1)!}\mathcal{I}\left(N+1, N-M+2, \gamma_0^{-1}, \frac{\gamma_{\text{T}}}{\beta_m}\right), \end{aligned} \tag{14}$$

where $\gamma_0 = \rho^{-1}$, the function $\mathcal{I}(a,b,z,\gamma)$ can be given by

$$\begin{aligned} \mathcal{I}(a,b,z,\gamma) = &\frac{\mathrm{e}^z}{\Gamma(a)}\sum_{i=0}^{a-1}\binom{a-1}{i}\Big[\Gamma(b-i-1,z) - \\ &\Gamma\big(b-i-1,(1+\gamma)z\big)\Big](-1)^i z^{i-b+1}, \end{aligned}$$

$\Gamma(\cdot)$ is the gamma function [31, eq. (8.310.1)], and $\Gamma(\cdot,\cdot)$ is the incomplete gamma function [31, eq. (8.350.2)].

Proof: Please refer to Appendix B. ■

Definition 2. (Diversity order.) The diversity order for the outage probability is defined as

$$d = -\lim_{\gamma\to\infty}\frac{\log(P_{\text{o}}(\gamma))}{\log(\gamma)}. \tag{15}$$

Proposition 2. (Diversity order.) The diversity order can be given by

$$d = N - M + 1. \tag{16}$$

Proof: Please refer to Appendix C. ■

Remark 1. (Diversity order compare.) From the diversity order in (16) and [32, Lemma 6], we can find that the sparse observation and coding scheme can achieve the same diversity order as the traditional scheme [33]. It implies that the sparse observation and coding scheme can achieve a performance close to that of the traditional scheme with lower power consumption.

B. Achievable Rate

To demonstrate the achievable performance of the proposed scheme, the achievable rates of the observed phenomena are derived. The achievable rate of observed phenomenon s_m can be given by

$$R_m = \log_2\left(1 + \text{SINR}_m\right), m \in \mathcal{M}. \tag{17}$$

To guarantee the success of observation, all the generated data of phenomena must be successfully transmitted to the FC for recovery. It requires that the achievable rate $R_m, m \in \mathcal{M}$ must exceed the generating rate of observed phenomena $R_{g,m}, m \in \mathcal{M}$, i.e., $R_m \geq R_{g,m}, m \in \mathcal{M}$, which is termed as the QoS requirement of phenomena. This is equivalent to $\text{SINR}_m \geq \gamma_{g,m}, m \in \mathcal{M}$, where $\gamma_{g,m} = 2^{R_{g,m}} - 1, m \in \mathcal{M}$.

From (10) and (17), we find that both the derived achievable rates and the network power consumption are closely related to the observation matrix $\mathbf{C}$ and the coding matrix $\mathbf{A}$. Specifically, activating more nodes to observe and encode more phenomena can improve achievable rates but leads to an increased network power consumption. It implies that there is a trade-off between achievable rates and network power consumption on the joint design of $\mathbf{C}$ and $\mathbf{A}$. Therefore, to satisfy the given QoS requirements of phenomena, we jointly optimize $\mathbf{C}$ and $\mathbf{A}$ to minimize the network power consumption.

IV. Power Consumption Minimization

In this section, we study the optimization problem to minimize the network power consumption by jointly optimizing the observation matrix and the coding matrix for given QoS requirements.

A. Problem Formulation

Without loss of generality, we ignore the constant term of network power consumption for brevity. The minimization problem of network power consumption subject to the given QoS constraints can be formulated as

$$\begin{aligned}
\mathcal{P}_0\colon \min_{\mathbf{C},\mathbf{A}} \quad & \sum_{m=1}^{M} \|[\mathbf{C}]_{\cdot m}\|_0 u_m + \sum_{k=1}^{K} \left\| \|\left[\mathbf{C}\right]_{k\cdot}\|_2^2 \right\|_0 P_k^g \\
& + \left(1+\sigma_o^2\right) \|\mathbf{A} \circ \mathbf{C}\|_2^2 \\
\text{s.t.} \quad & C_1 : \text{SINR}_m \geq \gamma_{g,m}, m \in \mathcal{M}, \\
& C_2 : c_{k,m} \in \{0,1\}, k \in \mathcal{K}, m \in \mathcal{M}, \\
& C_3 : (1-c_{k,m}) a_{k,m} = 0, k \in \mathcal{K}, m \in \mathcal{M}.
\end{aligned} \tag{18}$$

Constraint C_1 is the QoS requirements of different phenomena. Constraint C_2 represents the observed relation of nodes to phenomena during each observation. Constraint C_3 reveals the coupling relation between the observation matrix and the coding matrix. From (18), we find that problem $\mathcal{P}_0$ is intractable due to the following challenges.

1) **Coupled observation and coding matrices**. The observation matrix and coding matrix are coupled as nodes are required to choose part of the phenomena for sequential observation and coding.
2) **Discontinuous and non-differentiable l_0-norm**. Due to the dynamical and selective observed relation between nodes and phenomena, the objective function naturally contains the discontinuous and non-differentiable l_0-norm.
3) **Non-convex objective function and constraints**. The objective function is still non-convex even if the l_0-norm is replaced by a certain continuous function. Due to the inverse matrix in SINRs, the QoS constraints are non-convex. The other constraints are also non-convex.

Considering these challenges, problem $\mathcal{P}_0$ is an NP-hard problem. Only an exhaustive search can find the global optimum of $\mathcal{P}_0$, which requires the solution to 2^{KM} quadratically constrained quadratic programming (QCQP) subproblems with semi-definite relaxation (SDR). Due to the heavy computation of solving so many subproblems, the exhaustive search is impractical even for small-scale optimization problems.

B. Problem Solving

To solve the intractable problem $\mathcal{P}_0$ efficiently, we adopt the following approaches to cope with the aforementioned challenges. First, we decouple the coding matrix $\mathbf{A}$ from the observation matrix $\mathbf{C}$ by using equivalent variable substitutions to rewrite $\mathcal{P}_0$ as a sparse coding optimization problem without loss of optimality. Then, the discontinuous and non-differentiable l_0-norm in the objective function is approximated by the concave smooth function. Moreover, the auxiliary variables and the lower bound of SINR_m are introduced to transform the sparse coding optimization problem into a solvable DC programming problem. Finally, a CCP based algorithm with low-complexity is proposed to provide an effective solution to $\mathcal{P}_0$.

The following Lemma 2 decouples the coding matrix $\mathbf{A}$ from the observation matrix $\mathbf{C}$ by utilizing the equivalent relation between observation and coding of different phenomena at nodes.

Lemma 2. (Decoupling of observation and coding matrices.) There is an equivalent substitution between the observation matrix $\mathbf{C}$ and the coding matrix $\mathbf{A}$

$$c_{k,m} = \|a_{k,m}\|_0 k \in \mathcal{K}, m \in \mathcal{M}, \tag{19}$$

which decouples $\mathbf{A}$ from $\mathbf{C}$ without changing optimality.

Proof: Lemma 2 can be readily proved by contradiction. Assume that there exists a $c_{k,m} \neq 0$ to obtain P_k^{o*}, which is the minimum power consumption of sensing and coding at node k. If $a_{k,m} = 0$ there is $a_{k,m} v_{k,m} = 0$, which means that the value of $c_{k,m}$ is no longer important. Thus, we can set $c_{k,m} = 0$ to obtain the smaller power consumption than P_k^{o*}, i.e., $P_k^o < P_k^{o*}$. It contradicts the above assumption. Therefore, to achieve the optimal energy efficiency, there must be $c_{k,m} = \|a_{k,m}\|_0$. ■

Note that the coding matrix $\mathbf{A}$ has the same sparsity as the observation matrix $\mathbf{C}$ when $\mathbf{A}$ is decoupled from $\mathbf{C}$ by Lemma 2. With the sparse coding matrix $\mathbf{A}$, we equivalently rewrite

problem $\mathcal{P}_0$ as the following sparse coding optimization problem

$$\begin{aligned}\mathcal{P}_{\text{SC}}\text{: }\min_{\mathbf{A}} \quad & \sum_{m=1}^{M}\|[\mathbf{A}]_{\cdot m}\|_0 u_m+\sum_{k=1}^{K}\left\|\|[\mathbf{A}]_{k\cdot}\|_2^2\right\|_0 P_k^g \\ & +\left(1+\sigma_o^2\right)\|\mathbf{A}\|_2^2 \\ \text{s.t.} \quad & C_1: \text{SINR}_m \geq \gamma_{g,m},\ m\in\mathcal{M}.\end{aligned} \tag{20}$$

From (20), we find that problem $\mathcal{P}_{\text{SC}}$ is still challenging as the objective function contains the discontinuous and non-differentiable l_0-norm and the QoS constraints are non-convex.

To deal with the discontinuous and non-differentiable l_0-norm in the objective function, we utilize the continuous concave smooth logarithmic function in the following Lemma 3 to approximate l_0-norm.

Lemma 3. (Approximation of l_0-norm [34, Eq. (21)].) The continuous concave smooth logarithmic function is commonly regarded as an effective approximation of l_0-norm, whose effectiveness has been proved in [34]–[36]. The continuous concave smooth logarithmic function is given by

$$f_\theta(x)=\frac{\log(\frac{x}{\theta}+1)}{\log(\frac{1}{\theta}+1)}, \theta>0, \tag{21}$$

where θ is termed as the smoothness factor that can control the smoothness of approximation and the approximation error of l_0-norm. Specifically, a larger θ leads to a smoother function but a worse approximation of l_0-norm and vice versa.

With the smoothed approximation function of l_0-norm in Lemma 3, the sparse coding optimization problem $\mathcal{P}_{\text{SC}}$ can be rewritten as the following continuous optimization problem

$$\begin{aligned}\mathcal{P}_1\text{: }\min_{\mathbf{A}} \ & \sum_{m=1}^{M}\sum_{k=1}^{K} f_\theta\left(|a_{k,m}|^2\right)u_m+\sum_{k=1}^{K} f_\theta\left(\|[\mathbf{A}]_{k\cdot}\|_2^2\right)P_k^g \\ & +\left(1+\sigma_o^2\right)\sum_{m=1}^{M}\|[\mathbf{A}]_{\cdot m}\|_2^2 \\ \text{s.t.} \ & C_1: \text{SINR}_m \geq \gamma_{g,m}, m\in\mathcal{M}.\end{aligned} \tag{22}$$

Problem $\mathcal{P}_1$ is still intractable as the objective function and QoS constraints are non-convex.

To solve the non-convex problem $\mathcal{P}_1$, we introduce the auxiliary variables and the lower bound of SINR_m to transform $\mathcal{P}_1$ into a solvable DC programming problem[1]. The objective function of $\mathcal{P}_1$ can be equivalently represented as the normal DC form by introducing the auxiliary variable matrix $\mathbf{T}_{K\times M}$, which is shown in the following Lemma 4.

Lemma 4. (Equivalent DC form of objective function.) We replace $\mathbf{A}$ with $\mathbf{T}$ by $t_{k,m}\geq|a_{k,m}|^2, k\in\mathcal{K}, m\in\mathcal{M}$, where $t_{k,m}$ is the element of $\mathbf{T}$ at the row k and column m, i.e., $\mathbf{T}=[\mathbf{t}_1,\cdots,\mathbf{t}_M]$, $\mathbf{t}_m=[t_{1,m},\cdots,t_{K,m}]^T$. With this auxiliary variable substitution, the objective function of $\mathcal{P}_1$ can be equivalently rewritten as

$$\begin{aligned}&\left(1+\sigma_o^2\right)\sum_{m=1}^{M}\sum_{k=1}^{K} t_{k,m}- \\ &\left(-\sum_{m=1}^{M}\sum_{k=1}^{K} f_\theta\left(t_{k,m}\right)u_m-\sum_{k=1}^{K} f_\theta\left(\sum_{m=1}^{M} t_{k,m}\right)P_k^g\right).\end{aligned} \tag{23}$$

Note that the objective function in (23) follows the form of DC programming as the first and second terms are both convex functions.

Proof: Lemma 4 can be readily proved by contradiction. From (21), we find that the approximation function $f_\theta(x)$ is the strictly monotonically increasing function of x. Thus, the objective function in (23) is the strictly monotonically increasing function of $t_{k,m}$. If $t_{k,m}^*>|a_{k,m}^*|^2$, there must be a $t_{k,m}$ satisfying $|a_{k,m}^*|^2\leq t_{k,m}<t_{k,m}^*$ to further decrease the objective function, which is a contradiction. Therefore, the optimal solution must satisfy $|a_{k,m}^*|^2=t_{k,m}^*$, and introducing auxiliary variables $\mathbf{T}$ does not change the optimality. ■

To transform the QoS constraints C_1 into the normal DC form, we relax SINR_m by utilizing its lower bound γ_m, which is given in the following Lemma 5.

Lemma 5. (Lower bound of SINR_m.) The lower bound of SINR_m can be given by γ_m, i.e., $\text{SINR}_m\geq\gamma_m$, where

$$\gamma_m=\frac{[\mathbf{A}]_{\cdot m}^H\mathbf{H}^H\mathbf{H}[\mathbf{A}]_{\cdot m}}{\sum_{n\neq m}^{M}[\mathbf{A}]_{\cdot n}^H\mathbf{H}^H\mathbf{H}[\mathbf{A}]_{\cdot n}+\sum_{k=1}^{K}\|[\mathbf{A}]_{k\cdot}\|_2^2\mathbf{h}_k^H\mathbf{h}_k\sigma_o^2+\sigma^2}. \tag{24}$$

Proof: The proof of Lemma 5 is given at Proof D. ■

Substituting γ_m into QoS constraints C_1, we obtain the relaxed QoS constraints $\gamma_m\geq\gamma_{g,m}, m\in\mathcal{M}$. Then we equivalently rewrite the relaxed QoS constraints as

$$\begin{aligned}&\sum_{n\neq m}^{M}[\mathbf{A}]_{\cdot n}^H\left(\mathbf{H}^H\mathbf{H}+\sigma_o^2\mathbf{R}_2\right)[\mathbf{A}]_{\cdot n}\gamma_{g,m}+\gamma_{g,m}\sigma^2 \\ &-[\mathbf{A}]_{\cdot m}^H\left(\mathbf{H}^H\mathbf{H}-\gamma_{g,m}\sigma_o^2\mathbf{R}_2\right)[\mathbf{A}]_{\cdot m}\leq 0, m\in\mathcal{M},\end{aligned} \tag{25}$$

where $\mathbf{R}_2\in\mathbb{R}^{K\times K}$ is a diagonal matrix and its k-th diagonal element is $[\mathbf{H}]_{\cdot k}^H[\mathbf{H}]_{\cdot k}$. Note that the QoS constraints in (25) follow the form of DC programming as the sum of the first two terms and the third term are both convex functions.

Proposition 3. (DC programming problem.) The problem $\mathcal{P}_1$ can be relaxed as a DC programming problem

$$\begin{aligned}\mathcal{P}_2\text{:}\min_{\mathbf{A},\mathbf{T}} \ & \sum_{m=1}^{M}\sum_{k=1}^{K} f_\theta(t_{k,m})u_m+\sum_{k=1}^{K} f_\theta\left(\sum_{m=1}^{M} t_{k,m}\right)P_k^g \\ & +\left(1+\sigma_o^2\right)\sum_{m=1}^{M}\sum_{k=1}^{K} t_{k,m} \\ \text{s.t.} \ & C_4: |a_{k,m}|^2-t_{k,m}\leq 0, k\in\mathcal{K}, m\in\mathcal{M}, \\ & C_5: \sum_{n\neq m}^{M}[\mathbf{A}]_{\cdot n}^H\left(\mathbf{H}^H\mathbf{H}+\sigma_o^2\mathbf{R}_2\right)[\mathbf{A}]_{\cdot n}\gamma_{g,m}+\gamma_{g,m}\sigma^2 \\ & -[\mathbf{A}]_{\cdot m}^H\left(\mathbf{H}^H\mathbf{H}-\gamma_{g,m}\sigma_o^2\mathbf{R}_2\right)[\mathbf{A}]_{\cdot m}\leq 0, m\in\mathcal{M}.\end{aligned} \tag{26}$$

[1]The DC programming problem is the optimization problem whose objective function and constraints can be represented as the difference of two convex functions.

Proof: With Lemma 4 and 5, the objective function and QoS constraints of $\mathcal{P}_1$ can be transformed into the objective function and constraint C_5 of problem $\mathcal{P}_2$, which is a normal DC form. Moreover, constraint C_4 is convex. Thus, $\mathcal{P}_2$ is a standard DC programming problem. ■

The local optimal solution to DC programming problems can be readily provided by CCP methods [37]. Hence, we propose a two-level loop CCP based algorithm to provide a local optimum to $\mathcal{P}_2$, which contains an outer and inner loop.

In the outer loop, we update the smoothness factor θ by $\theta = \beta\theta, 0 < \beta < 1$ to approximate l_0-norm with smooth function $f_\theta(x)$. Specifically, the performance of the smoothed l_0-norm approximation $f_\theta(x)$ depends on the smoothness factor θ. When x is large, θ should be large so that the approximation algorithm can explore the entire variable space; when x is small, θ should be small so that $f_\theta(x)$ has behavior close to l_0-norm. The update rule $\theta = \beta\theta$ has been proved to be effective by [38].

In the inner loop, with the fixed smoothed l_0-norm, we use CCP to provide an effective solution, which requires several iterations. In each iteration, we transform the objective function and constraints of the DC programming problem into convex forms. Specifically, we replace the concave part in DC functions with their first order Taylor expansions, which leads to a convex subproblem. In a certain iteration of CCP, the subproblem takes the following form in (27), as shown at the bottom of this page, which is a convex problem. Then we solve a sequence of convex subproblems successively during iterations and find a local optimum to be the starting point of the next outer loop. After multiple iterations, a local optimum is provided to problem $\mathcal{P}_2$.

Moreover, the CCP based algorithm requires a feasible point as the initial starting point. The feasible point is obtained by solving the full observation case where all nodes observe all phenomena[2], which is a special case of $\mathcal{P}_0$ and is termed as $\mathcal{P}_{\text{INI}}$. Problem $\mathcal{P}_{\text{INI}}$ contains non-convex QoS constraints but can be readily solved through the SDR method [39]. With the SDR method, problem $\mathcal{P}_{\text{INI}}$ is given as

$$\begin{aligned} \mathcal{P}_{\text{INI}}\text{:} \min_{\{\mathbf{A}_m\}} \quad & \left(1+\sigma_o^2\right)\sum_{m=1}^{M}\operatorname{tr}\left(\mathbf{A}_m\right) \\ \text{s.t.} \quad & \frac{\operatorname{tr}\left(\mathbf{H}\mathbf{A}_m\mathbf{H}^H\right)}{\sum\limits_{n\neq m}^{M}\operatorname{tr}(\mathbf{H}\mathbf{A}_n\mathbf{H}^H)+\sum\limits_{m=1}^{M}\operatorname{tr}(\mathbf{A}_m\mathbf{R}_2)\sigma_o^2+\sigma^2} \\ & \geq \gamma_{g,m},\ m\in\mathcal{M}, \\ & \mathbf{A}_m \succeq \mathbf{0}, m\in\mathcal{M}, \end{aligned} \tag{28}$$

where $\mathbf{A}_m = [\mathbf{A}]_{\cdot m}[\mathbf{A}]_{\cdot m}^H, m\in\mathcal{M}$. We denote the optimum of problem $\mathcal{P}_{\text{INI}}$ as $\{\mathbf{A}_m^*\}$ and utilize $\{\mathbf{A}_m^*\}$ to deliver a feasible point as the initial starting point. Specifically, if $\{\mathbf{A}_m^*\}$ are all rank-one matrices, the eigenvalue decomposition (EVD) is applied on $\{\mathbf{A}_m^*\}$ to generate the starting point $\{[\mathbf{A}]_{\cdot m}\}$. Otherwise, the Gaussian randomization and scaling method in [39], [40] is used to generate the starting point. The complete CCP based sparse observation and coding algorithm is outlined in Algorithm 1.

Algorithm 1 The CCP Based Sparse Observation and Coding Algorithm.

Input: Set iteration index $i = 0$, smoothness factor $\theta = \theta_0$, decaying factor $0 < \beta < 1$, and small constants ϵ and ϵ_1.

Solve $\mathcal{P}_{\text{INI}}$ in (28) and denote its solution as $\{\mathbf{A}_m^*\}$.
If $\operatorname{rank}(\{\mathbf{A}_m^*\}) = 1, m\in\mathcal{M}$, utilize EVD to obtain the initial starting point $\mathbf{A}^{(0)}$.
Else, utilize the Gaussian randomization and scaling method to obtain the initial starting point $\mathbf{A}^{(0)}$.
repeat
  repeat
    Solve $\mathcal{P}_3$ in (27) at the starting point $\mathbf{A}^{(i)}$ and denote the solution as $\mathbf{A}^*$.
    Set $\mathbf{A}^{(i+1)} = \mathbf{A}^*$.
    Update iteration $i = i+1$.
  until $\frac{\left\|\mathbf{A}^{(i-1)}-\mathbf{A}^{(i)}\right\|_2}{\left\|\mathbf{A}^{(i-1)}\right\|_2} < \epsilon_1$.
  Update $\theta = \beta\theta$, $\mathbf{A}^{(0)} = \mathbf{A}^*$, and $i = 0$.
until $\theta < \epsilon$.

Output: $\mathbf{A}^*$.

Remark 2. (Computational complexity of Algorithm 1.) The worst convergence rate of CCP algorithm is linear [41], which means that each inner loop at most solves $\mathcal{O}\left(\log\left(\epsilon_1^{-1}\right)\right)$

[2]As long as problem $\mathcal{P}_0$ is feasible, the QoS requirements of the full observation case are surely satisfied. In this paper, we only discuss the feasible $\mathcal{P}_0$.

$$\begin{aligned} \mathcal{P}_3: \min_{\mathbf{A},\{t_{k,m}\}} \quad & \sum_{m=1}^{M}\sum_{k=1}^{K}\left(\nabla f_\theta\left(t_{k,m}^{(i)}\right)u_m+1+\sigma_o^2\right)t_{k,m}+\sum_{k=1}^{K}\nabla f_\theta\left(\sum_{m=1}^{M}t_{k,m}^{(i)}\right)^T\mathbf{t}_m P_k^g \\ \text{s.t.} \quad & |a_{k,m}|^2 - t_{k,m} \leq 0, k\in\mathcal{K}, m\in\mathcal{M}, \\ & \gamma_{g,m}\sum_{n\neq m}^{M}[\mathbf{A}]_{\cdot n}^H\left(\mathbf{H}^H\mathbf{H}+\sigma_o^2\mathbf{R}_2\right)[\mathbf{A}]_{\cdot n}+\gamma_{g,m}\sigma^2- \\ & \left(2\Re\left\{[\mathbf{A}]_{\cdot m}^{(i)H}\left(\mathbf{H}^H\mathbf{H}-\gamma_{g,m}\sigma_o^2\mathbf{R}_2\right)[\mathbf{A}]_{\cdot m}\right\}-[\mathbf{A}]_{\cdot m}^{(i)H}\left(\mathbf{H}^H\mathbf{H}-\gamma_{g,m}\sigma_o^2\mathbf{R}_2\right)[\mathbf{A}]_{\cdot m}^{(i)}\right)\leq 0, m\in\mathcal{M}. \end{aligned} \tag{27}$$

convex problems with $2KM$ variables. Furthermore, there are $(\theta_0-\epsilon)\beta^{-1}$ outer loops. Therefore, Algorithm 1 approximately needs to solve $\mathcal{O}\left(\log\left(\epsilon_1^{-1}\right)(\theta_0-\epsilon)\beta^{-1}\right)$ convex problems with $2KM$ variables. Using the interior point method with step length μ [42], the computational complexity of Algorithm 1 is $\mathcal{O}\left(\frac{(\theta_0-\epsilon)\log\left(\epsilon_1^{-1}\right)\log(2KM/\epsilon_1)}{\beta\log(\mu)}\right)$, which is much more efficient than the exhaustive search.

V. Power Consumption Minimization in Collaboration Scheme

In this section, we introduce the collaboration among nodes into the sparse observation and coding scheme to further improve the performance. The collaboration has a significant effect on the SINR of observed phenomena, the transmission duration, and the network power consumption, which leads to a more complicated trade-off between achievable rates and network power consumption. To achieve the optimal trade-off, the minimization problem of network power consumption for given rate constraints is studied.

A. Collaboration Scheme

Introducing collaboration, we propose a sparse observation and collaborative coding scheme for IoT to monitor multiple phenomena, where the observation diversity among nodes is utilized to improve the performance. This scheme consists of three phases: sensing, collaborative coding, and transmission. The sensing and transmission phases are the same as before, and we focus on the collaborative coding phase as illustrated in Fig. 2. For clarity, the sparse observation and coding scheme and the sparse observation and collaborative coding scheme are referred to as the distributed scheme and the collaboration scheme, respectively, in the later section. Different from the distributed scheme, the collaboration scheme requires the nodes belonging to the same group $\mathcal{Q}_m$ to share the observation of phenomenon s_m at the expense of increasing power consumption and occupying part of the transmission time[3].

The observation of nodes in the sensing phase is the same as (2). In the collaborative coding phase, the nodes in group $\mathcal{Q}_m$ share the observation of phenomenon s_m and obtain the corresponding collaborative observation by coherently combining the shared observations. The collaborative observation of nodes in group $\mathcal{Q}_m$ can be given by

$$[\mathbf{V}_c]_{\cdot m}=\mathbf{W}_m\left[\mathbf{V}\right]_{\cdot m},m\in\mathcal{M}, \tag{29}$$

where $\mathbf{W}_m\in\mathbb{R}^{K\times K}$ is the collaboration matrix of nodes in group $\mathcal{Q}_m$. The element value of $\mathbf{W}_m$ is dependent on the observed relation of nodes to phenomena. If node k observes phenomenon s_m, the elements of $\mathbf{W}_m$ at row k are all set to $\|\left[\mathbf{C}\right]_{\cdot m}\|_0^{-1}$, i.e., $[\mathbf{W}_m]_{k\cdot}=\|\left[\mathbf{C}\right]_{\cdot m}\|_0^{-1}$,

[3]The observation sharing of nodes in groups $\mathcal{Q}_m$ can be realized through the reliable multicast [26], where the collaboration power consumption is directly proportional to $|\mathcal{Q}_m|$ and the duration ratio of collaboration to transmission is τ. The focus of the present work is on the conceptual aspects of node collaboration and not the details of its physical implementation.

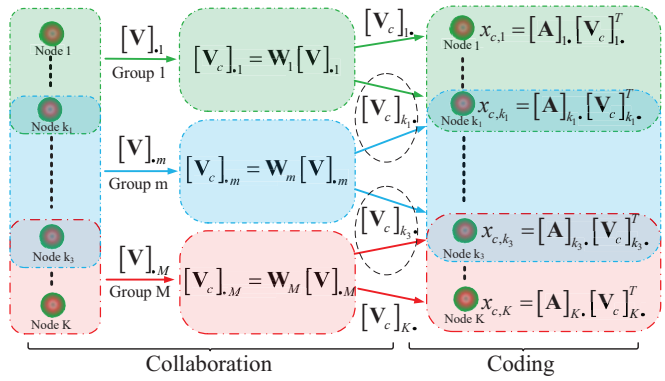


Figure 2: The process of collaborative coding phase.

otherwise $[\mathbf{W}_m]_{k\cdot}=\mathbf{0}$. The collaborative observations of different phenomena at node k are given as

$$\begin{aligned}[\mathbf{V}_c]_{k\cdot}^T&=\begin{bmatrix}\sum_{i=1}^{K}\frac{c_{i,1}}{\|[\mathbf{C}]_{\cdot 1}\|_0}(s_1+n_{i,1})\\ \vdots\\ \sum_{i=1}^{K}\frac{c_{i,M}}{\|[\mathbf{C}]_{\cdot M}\|_0}(s_M+n_{i,M})\end{bmatrix}\\&=[\mathbf{C}]_{k\cdot}^T\circ(\mathbf{s}+\mathbf{n}_{c,o}),k\in\mathcal{K},\end{aligned} \tag{30}$$

where $\mathbf{n}_{c,o}$ is the equivalent i.i.d Gaussian noise satisfying $\mathbf{n}_{c,o}\sim\mathcal{CN}(\mathbf{0},\mathbf{R}_3)$[4], $\mathbf{R}_3$ is a diagonal matrix and its m-th diagonal element is $\sigma_o^2\|\left[\mathbf{C}\right]_{\cdot m}\|_0^{-1}$. Then, each node linearly encodes the collaborative observations of different phenomena into a scalar for transmission. With a coding vector $[\mathbf{A}]_{k\cdot}$, the coded observation of node k is given by

$$x_{c,k}=[\mathbf{A}]_{k\cdot}\left([\mathbf{C}]_{k\cdot}^T\circ(\mathbf{s}+\mathbf{n}_{c,o})\right),k\in\mathcal{K}. \tag{31}$$

The coded observation of all nodes is given as

$$\begin{aligned}\mathbf{x}_c&=\begin{bmatrix}[\mathbf{A}]_{1\cdot}\left([\mathbf{C}]_{1\cdot}^T\circ(\mathbf{s}+\mathbf{n}_{c,o})\right)\\ \vdots\\ [\mathbf{A}]_{K\cdot}\left([\mathbf{C}]_{K\cdot}^T\circ(\mathbf{s}+\mathbf{n}_{c,o})\right)\end{bmatrix}\\&=(\mathbf{A}\circ\mathbf{C})(\mathbf{s}+\mathbf{n}_{c,o}).\end{aligned} \tag{32}$$

In the transmission phase, all nodes concurrently transmit their coded observations to the FC for recovery. The received observation of the FC can be given by

$$\begin{aligned}\mathbf{y}_c=&\mathbf{H}\left[\mathbf{A}\circ\mathbf{C}\right]_{\cdot m}s_m+\mathbf{H}\sum_{n\neq m}^{M}\left[\mathbf{A}\circ\mathbf{C}\right]_{\cdot n}s_n\\&+\mathbf{H}(\mathbf{A}\circ\mathbf{C})\mathbf{n}_{c,o}+\mathbf{n}_r.\end{aligned} \tag{33}$$

With the known CSI between nodes and the FC, the FC can recover phenomena $\mathbf{s}$ by processing the received observation

[4]The observation noise is suppressed by collaboration, and the suppression of noise power is directly proportional to the node number in collaboration, i.e., $\|\left[\mathbf{C}\right]_{\cdot m}\|_0$. Thus, the m-th element of suppressed noise $\mathbf{n}_{c,o}$ is an equivalent i.i.d Gaussian noise with mean 0 and variance $\sigma_o^2\|\left[\mathbf{C}\right]_{\cdot m}\|_0^{-1}$.

through the linear MMSE receiver. The SINR of phenomenon s_m $(m \in \mathcal{M})$ after linear MMSE receiver is given by

$$\text{SINR}_{c,m} = \mathbf{d}_m^H \left(\sum_{n \neq m}^{M} \mathbf{d}_n \mathbf{d}_n^H + \mathbf{D}\mathbf{R}_3\mathbf{D}^H + \sigma^2 \mathbf{I} \right)^{-1} \mathbf{d}_m, \quad (34)$$

where $\mathbf{D}_{N\times M} = [\mathbf{d}_1, \cdots, \mathbf{d}_M]$. The achievable rate of phenomenon s_m is given by

$$R_{c,m} = \frac{1}{1+\tau} \log_2 \left(1 + \text{SINR}_{c,m}\right), m \in \mathcal{M}, \quad (35)$$

where τ represents the duration ratio of collaboration to transmission. To guarantee the QoS requirements of phenomena $\mathbf{s}$, the achievable rates should exceed their generating rate $R_{g,m}, m \in \mathcal{M}$, i.e., $R_{c,m} \geq R_{g,m}, m \in \mathcal{M}$. The rate requirements are equivalent to $\text{SINR}_{c,m} \geq \tilde{\gamma}_{g,m}, m \in \mathcal{M}$, where $\tilde{\gamma}_{g,m} = 2^{(1+\tau)R_{g,m}} - 1$.

Due to introducing the collaboration among nodes, there is an extra collaboration power consumption in the collaboration scheme. Thus, the network power consumption of the collaboration scheme consists of four parts: sensing, collaboration, coding, and transmission power consumption. Except for the power consumption of collaboration and transmission, the others are the same as that of the distributed scheme. The collaboration power consumption is commonly regarded to be directly proportional to the number of collaboration nodes and the amount of collaborated data [26]. Therefore, the total collaboration power consumption of nodes can be given by

$$P_t^c = \sum_{m=1}^{M} \| [\mathbf{C}]_{\cdot m} \|_0 q_m R_{c,m}, \quad (36)$$

where q_m represents the collaboration power consumption to generate 1-bit collaborative observation of phenomenon s_m. And, according to (32), the total transmission power consumption is given as

$$\begin{aligned} \sum_{k=1}^{K} \mathbb{E}\left\{x_{c,k} x_{c,k}^H\right\} &= \sum_{k=1}^{K} [\mathbf{A} \circ \mathbf{C}]_{k\cdot} \left(\mathbf{I} + \mathbf{R}_3\right) [\mathbf{A} \circ \mathbf{C}]_{k\cdot}^H \\ &= \text{tr}\left(\mathbf{A} \circ \mathbf{C} \left(\mathbf{I} + \mathbf{R}_3\right) \mathbf{C}^H \circ \mathbf{A}^H\right). \end{aligned} \quad (37)$$

Thus, the network power consumption of the collaboration scheme can be given by

$$\begin{aligned} P = &\sum_{m=1}^{M} \| [\mathbf{C}]_{\cdot m} \|_0 u_{c,m} + \sum_{k=1}^{K} \left\| \| [\mathbf{C}]_{k\cdot} \|_2^2 \right\|_0 P_k^g + \\ &\sum_{k=1}^{K} P_k^{sl} + \text{tr}\left(\mathbf{A} \circ \mathbf{C} \left(\mathbf{I} + \mathbf{R}_3\right) \mathbf{C}^H \circ \mathbf{A}^H\right), \end{aligned} \quad (38)$$

where $u_{c,m} = (p_m + q_m) R_{g,m}$ represents the total power consumption of observation and collaboration of phenomenon s_m for the given generating rate $R_{g,m}$.

Remark 3. (Collaboration effects.) From (34), (35), and (38), collaboration has a heavy effect on the network power consumption, the received SINRs of observed phenomena, and the transmission duration, respectively. First, collaboration introduces an extra collaboration power consumption, which causes an increased network power consumption. Meanwhile, collaboration improves the received SINRs by utilizing the observation diversity of nodes to suppress observation noise and equalize bad observations. Moreover, collaboration shortens the transmission duration of nodes, which restricts the achievable rates of observed phenomena. Consequently, compared with the distributed scheme, introducing collaboration results in a more complicated trade-off between achievable rates and network power consumption.

Taking full consideration of the collaboration effects in Remark 3, the optimal trade-off between achievable rates and network power consumption should be studied by jointly designing the observation and coding matrices.

B. Problem Formulation

Again, the constant terms of network power consumption are ignored to give the optimization problem as

$$\begin{aligned} \mathcal{P}_{\text{c}}: \min_{\mathbf{C},\mathbf{A}} \quad & \sum_{m=1}^{M} \| [\mathbf{C}]_{\cdot m} \|_0 u_{c,m} + \sum_{k=1}^{K} \left\| \| [\mathbf{C}]_{k\cdot} \|_2^2 \right\|_0 P_k^g + \\ & \text{tr}\left(\mathbf{A} \circ \mathbf{C} \left(\mathbf{I} + \mathbf{R}_3\right) \mathbf{C}^H \circ \mathbf{A}^H\right) \\ \text{s.t.} \quad & C_2: c_{k,m} \in \{0,1\}, k \in \mathcal{K}, m \in \mathcal{M}, \\ & C_3: (1 - c_{k,m}) a_{k,m} = 0, k \in \mathcal{K}, m \in \mathcal{M}, \\ & C_6: \text{SINR}_{c,m} \geq \tilde{\gamma}_{g,m}, m \in \mathcal{M}. \end{aligned} \quad (39)$$

Constraints C_2 and C_3 represent the observed relation of nodes to phenomena and the coupled relation of observation matrix to coding matrix, respectively. Constraint C_6 ensures that the QoS requirements of phenomena are satisfied. Similar to problem $\mathcal{P}_0$, problem $\mathcal{P}_{\text{c}}$ is intractable. Compared with $\mathcal{P}_0$, $\mathcal{P}_{\text{c}}$ can be solved in a similar but more complex method.

C. Problem Solving

For brevity, we directly rewrite problem $\mathcal{P}_{\text{c}}$ as a solvable DC programming problem in Proposition 4.

Proposition 4. (DC programming problem of $\mathcal{P}_{\text{c}}$.) Problem $\mathcal{P}_{\text{c}}$ can be rewritten as the following DC programming problem

$$\begin{aligned} \mathcal{P}_4: \min_{\mathbf{A},\mathbf{T}} \quad & \sum_{m=1}^{M} \sum_{k=1}^{K} f_\theta(t_{k,m}) u_{c,m} + \sum_{k=1}^{K} f_\theta\left(\sum_{m=1}^{M} t_{k,m}\right) P_k^g \\ & + \left(1 + \sigma_o^2\right) \sum_{m=1}^{M} \sum_{k=1}^{K} t_{k,m} \\ \text{s.t.} \quad & C_4: |a_{k,m}|^2 - t_{k,m} \leq 0, k \in \mathcal{K}, m \in \mathcal{M}, \\ & C_7: \tilde{\gamma}_{g,m}\left(1+\sigma_o^2\right) \sum_{n \neq m}^{M} [\mathbf{A}]_{\cdot n}^H \mathbf{H}^H \mathbf{H} [\mathbf{A}]_{\cdot n} + \tilde{\gamma}_{g,m} \sigma^2 \\ & - \left(1 - \tilde{\gamma}_{g,m} \sigma_o^2\right) [\mathbf{A}]_{\cdot m}^H \mathbf{H}^H \mathbf{H} [\mathbf{A}]_{\cdot m} \leq 0, \ m \in \mathcal{M}. \end{aligned} \quad (40)$$

Proof: Please refer to Appendix E. ■

The CCP based sparse observation and collaborative coding algorithm to solve $\mathcal{P}_4$ is outlined in Algorithm 2. The starting point of Algorithm 2 is obtained by solving the full collaboration case where all nodes collaboratively observe all

phenomena[5], which is a special case of problem $\mathcal{P}_{\text{SOC}}$ in (51) and is termed as $\mathcal{P}_{\text{INI}}^{\text{C}}$. With the SDR method, problem $\mathcal{P}_{\text{INI}}^{\text{C}}$ is

$$\begin{aligned}\mathcal{P}_{\text{INI}}^{\text{C}}\text{:} \min_{\{\mathbf{A}_m\}} \quad & \sum_{m=1}^{M}\left(1+\frac{\sigma_o^2}{K}\right)\text{tr}\left(\mathbf{A}_m\right)\\ \text{s.t.} \quad & \frac{\text{tr}\left(\mathbf{H}\mathbf{A}_m\mathbf{H}^H\right)}{\sum\limits_{n\neq m}^{M}\text{tr}(\mathbf{H}\mathbf{A}_n\mathbf{H}^H)+\frac{\sigma_o^2}{K}\sum\limits_{m=1}^{M}\text{tr}(\mathbf{H}\mathbf{A}_m\mathbf{H}^H)+\sigma^2}\\ & \geq \gamma_{g,m},\ m\in\mathcal{M},\\ & \mathbf{A}_m\succeq\mathbf{0}, m\in\mathcal{M},\end{aligned} \tag{42}$$

where $\mathbf{A}_m = [\mathbf{A}]_{\cdot m}[\mathbf{A}]_{\cdot m}^H, m \in \mathcal{M}$. The computational complexity of Algorithm 2 is approximately to solve $\mathcal{O}\left(\log\left(\epsilon_1^{-1}\right)(\theta_0-\epsilon)\beta^{-1}\right)$ convex problems with $2KM$ variables. Using the interior point method with step length μ, the computational complexity of Algorithm 2 is $\mathcal{O}\left(\frac{(\theta_0-\epsilon)\log\left(\epsilon_1^{-1}\right)\log(2KM/\epsilon_1)}{\beta\log(\mu)}\right)$, which is much more efficient than the exhaustive search.

Algorithm 2 CCP Based Sparse Observation and Collaborative Coding Algorithm.

Input: Set iteration index $i=0$, smoothness factor $\theta=\theta_0$, decaying factor $0<\beta<1$, and small constants ϵ and ϵ_1.

Solve $\mathcal{P}_{\text{INI}}^{\text{C}}$ in (42) and denote the solution as $\{\mathbf{A}_m^*\}$.
If $\text{rank}(\{\mathbf{A}_m^*\})=1, m\in\mathcal{M}$, utilize EVD to obtain the initial starting point $\mathbf{A}^{(0)}$.
Else, utilize the Gaussian randomization and scaling method to obtain the initial starting point $\mathbf{A}^{(0)}$.
repeat
  repeat
    Solve $\mathcal{P}_5$ in (41) at the starting point $\mathbf{A}^{(i)}$ and denote the solution as $\mathbf{A}^*$.
    Set $\mathbf{A}^{(i+1)}=\mathbf{A}^*$.
    Update iteration $i=i+1$.
  until $\frac{\left\|\mathbf{A}^{(i-1)}-\mathbf{A}^{(i)}\right\|_2}{\left\|\mathbf{A}^{(i-1)}\right\|_2}<\epsilon_1$.
  Update $\theta=\beta\theta$, $\mathbf{A}^{(0)}=\mathbf{A}^*$, and $i=0$.
until $\theta<\epsilon$.

Output: $\mathbf{A}^*$.

[5]As long as problem $\mathcal{P}_{\text{SOC}}$ is feasible, the QoS requirements of the full collaboration case are surely satisfied.

Table I: Simulation parameters

Parameter	**Value**
Node number	$K=100$
Phenomenon number	$M=2\sim 25$
Antenna number at the FC	$N=30$
Noise power at the FC	10^{-3}
Observation noise power	10^{-2}
Duration ratio of collaboration to transmission	$\tau=1$
Power consumption to generate 1-bit data of s_m	$p_m=10^{-2}mJ$
Circuit power consumption of active node	$P^a=10^{-2}mW$
Circuit power consumption of sleep node	$P^{sl}=10^{-3}mW$
Power consumption to collaborate 1-bit data of s_m	$q_m=0.1mJ$
Initial value of θ	$\theta_0=10^3$
Update factor of θ	$\beta=0.1$
Acceptable maximum of θ	$\epsilon=10^{-3}$
Acceptable maximum error	$\epsilon_1=10^{-3}$

VI. Simulation Results And Discussions

In this section, simulation results are provided to show the performances of the proposed schemes. The channels between nodes and the FC are assumed to be the i.i.d Rayleigh block fading channel. The parameter settings are listed in Table I unless specified otherwise.

The network power consumptions of the IoT with conventional full observation scheme and the IoT with sparse observation and coding scheme versus the total achievable rate are shown in Fig. 3. From Fig. 3, we can see that the network power consumptions remain approximately constant at first, and then sharply increase with the increase in total achievable rate. This can be explained as follows. The IoT with different schemes is first in the power-limited regime and then in the bandwidth-limited regime. Moreover, for the same total achievable rate, the network power consumption of the IoT with full observation scheme is always greater than that of the IoT with sparse observation and coding scheme, which demonstrates that the sparse observation and coding scheme can improve energy efficiency. Fig. 4 shows the performance of the IoT with sparse observation and coding scheme for different node numbers. For the same total achievable rate, the network power consumption is gradually decreasing when the node number increases. This can be explained as follows. With sparse observation and coding scheme, IoT can select the

$$\begin{aligned}\mathcal{P}_5: \min_{\mathbf{A},\{t_{k,m}\}} \quad & \sum_{m=1}^{M}\sum_{k=1}^{K}\left(\nabla f_\theta\left(t_{k,m}^{(i)}\right)u_{c,m}+1+\sigma_o^2\right)t_{k,m}+\sum_{k=1}^{K}\nabla f_\theta\left(\sum_{m=1}^{M}t_{k,m}^{(i)}\right)^T\mathbf{t}_m P_k^g\\ \text{s.t.} \quad & |a_{k,m}|^2-t_{k,m}\leq 0, k\in\mathcal{K}, m\in\mathcal{M},\\ & \tilde{\gamma}_{g,m}\left(1+\sigma_o^2\right)\sum_{n\neq m}^{M}[\mathbf{A}]_{\cdot n}^H\mathbf{H}^H\mathbf{H}[\mathbf{A}]_{\cdot n}+\tilde{\gamma}_{g,m}\sigma^2-\\ & \left(1-\tilde{\gamma}_{g,m}\sigma_o^2\right)\left(2\Re\left\{[\mathbf{A}]_{\cdot m}^{(i)H}\mathbf{H}^H\mathbf{H}[\mathbf{A}]_{\cdot m}\right\}-[\mathbf{A}]_{\cdot m}^{(i)H}\mathbf{H}^H\mathbf{H}[\mathbf{A}]_{\cdot m}^{(i)}\right)\leq 0, m\in\mathcal{M}.\end{aligned} \tag{41}$$

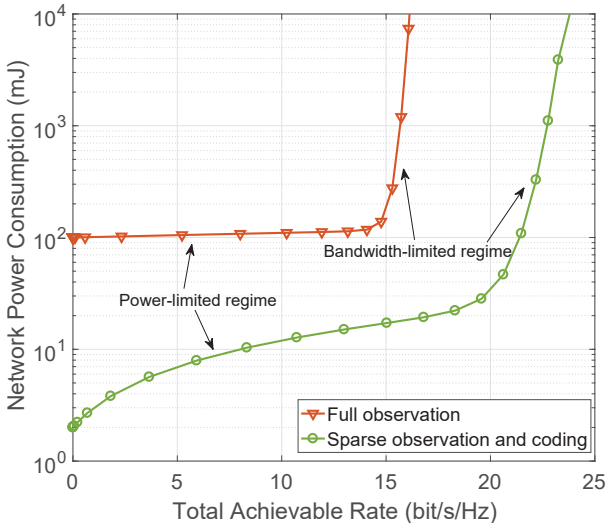


Figure 3: The comparison of network power consumption between the IoT with traditional full observation scheme and the IoT with sparse observation and coding scheme.

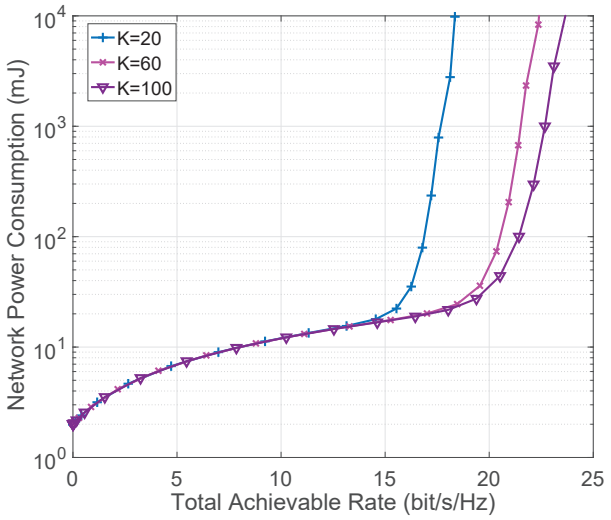


Figure 4: The network power consumption of the IoT with sparse observation and coding scheme for different node numbers.

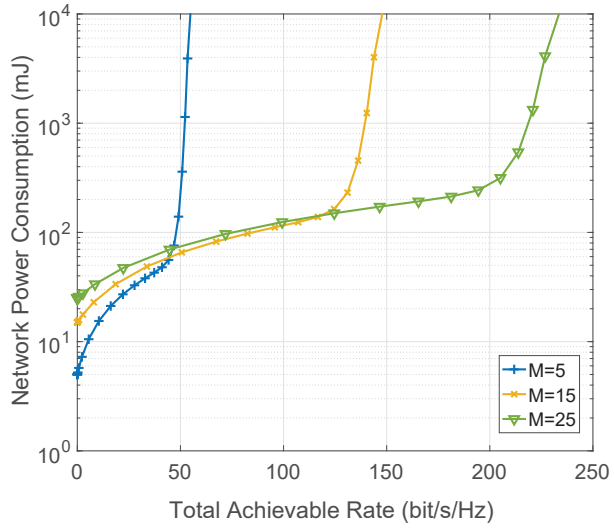


Figure 5: The network power consumption of the IoT with sparse observation and coding scheme for different phenomenon numbers.

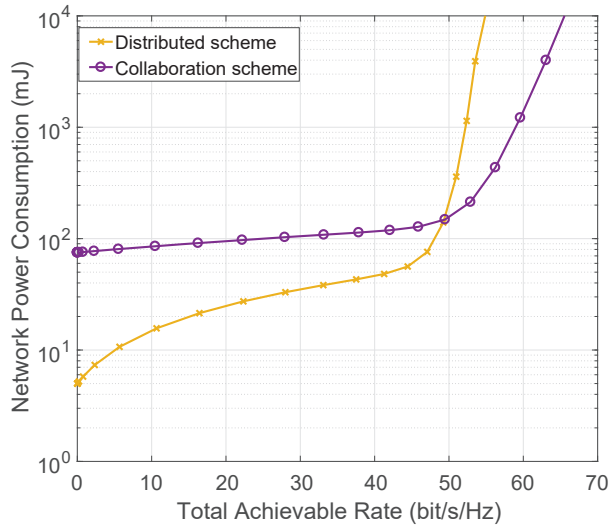


Figure 6: The comparison of network power consumption between the IoT with sparse observation and coding scheme and the IoT with sparse observation and collaborative coding scheme.

nodes with a better state of observation and channel to observe and encode phenomena when there are more nodes. In other words, when IoT deploys more nodes to monitor phenomena, the sparse observation and coding scheme can provide larger performance gains.

Fig. 5 shows the performance of the IoT with sparse observation and coding scheme for different phenomenon numbers. For low total achievable rates, to achieve the same total achievable rate, the network power consumption gradually increases when the phenomenon number increases. This can be interpreted as follows. Observing more phenomena needs to activate more nodes, which consumes more power in observation and coding. Moreover, the increased power consumption due to observing more phenomena is the dominant power consumption in the power-limited regime. For high total achievable rates, to achieve the same total achievable rate, the network power consumption gradually decreases as the phenomenon number increases. This can be interpreted as follows. Each node has more chances to choose the better phenomena for observation according to its observation state. As a result, the sparse observation and coding scheme can render a larger performance gain when IoT monitors more phenomena under high QoS requirements.

The network power consumptions of the IoT with distributed scheme and the IoT with collaboration scheme are

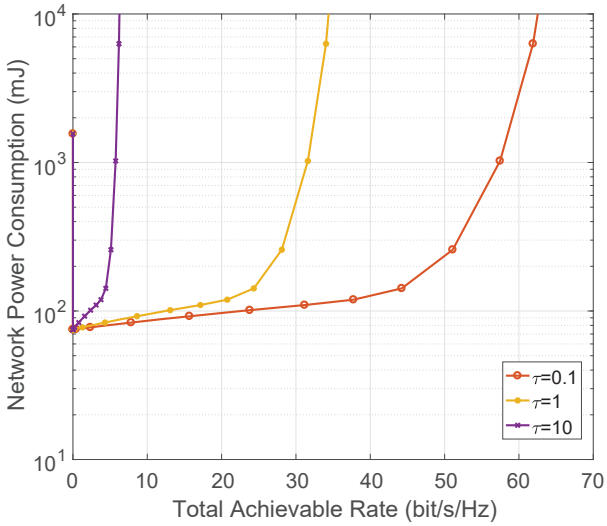


Figure 7: The network power consumption of the IoT with sparse observation and collaborative coding scheme for different duration ratios of collaboration to transmission.

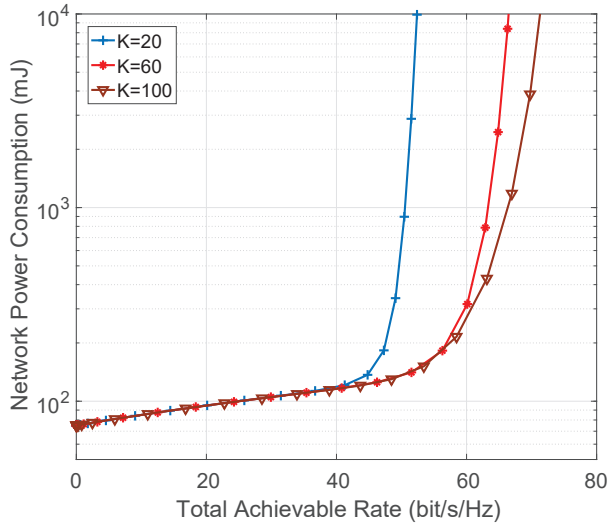


Figure 8: The network power consumption of the IoT with sparse observation and collaborative coding scheme for different node numbers.

compared in Fig. 6. For low total achievable rates, to achieve the same total achievable rate, the network power consumption of the IoT with collaboration scheme is greater than that of the IoT with distributed scheme. This can be explained as follows. The collaboration scheme requires extra power to achieve the collaboration of nodes in the same group, which is dominant power consumption in the power-limited regime. Moreover, the collaboration gain is not significant due to the small number of observations at low total achievable rates. For high total achievable rates, to achieve the same total achievable rate, the network power consumption of the IoT with distributed scheme is greater than that of the IoT with collaboration scheme. This interpretation is given as follows. The collaboration scheme can provide a significant performance gain by utilizing the observation diversity of more activated nodes, and the collaboration power consumption is not dominant for high total achievable rates. It demonstrates that compared with the distributed scheme, the collaboration scheme is more suitable for IoT to monitor phenomena under high QoS requirements.

Fig. 7 shows the performance of the IoT with collaboration scheme for different duration ratios of collaboration to transmission. For the same total achievable rate, the network power consumption gradually increases when the duration ratio increases. The reason is that the increased τ results in the shortened transmission duration. Due to the shortened transmission duration, the achievable rates of observed phenomena are decreased. Therefore, to transmit the same number of observations, nodes require more power for transmission, which is used to compensate for the loss of achievable rates.

Fig. 8 shows the performance of the IoT with collaboration scheme for different node numbers. For the same total achievable rate, the network power consumption gradually decreases when the node number increases. The interpretation is that the IoT consisting of more nodes can select the more suitable

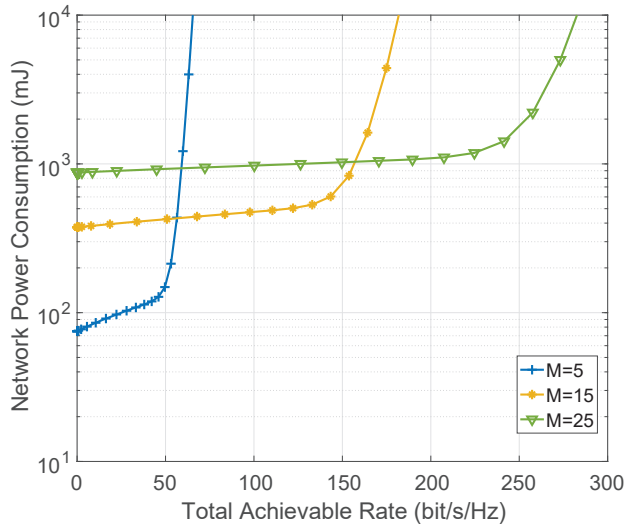


Figure 9: The network power consumption of the IoT with sparse observation and collaborative coding scheme for different phenomenon numbers.

nodes to achieve the observation and collaborative coding of phenomena. It also demonstrates that the collaboration scheme can provide significant collaboration gains for large-scale IoT to monitor multiple phenomena. Fig. 9 shows the performance of the IoT with collaboration scheme for different phenomenon numbers. For low total achievable rates, to achieve the same total achievable rate, the network power consumption gradually increases when the phenomenon number increases. This can be explained as follows. Monitoring more phenomena needs more power consumption for observation and collaborative coding and this power consumption is dominant in the power-limited regime. For high total achievable rates, to achieve the same total achievable rate, the network power consumption

gradually decreases when the phenomenon number increases. This can be interpreted as follows. The performance gain of the collaboration scheme is significant and its power consumption is not dominant for high total achievable rates. It demonstrates that the collaboration scheme is suitable for IoT to monitor multiple phenomena under high QoS requirements.

VII. Conclusion

This paper has studied a sparse observation and coding scheme for IoT to achieve the energy-efficient observation of multiple phenomena, where the nodes can dynamically select part of the phenomena for observation and coding. By analyzing outage performance, we demonstrate that the sparse observation and coding scheme can achieve a performance close to that of the full observation scheme with lower power consumption. Moreover, with the derived achievable rates and network power consumption, we find a performance trade-off between achievable rates and power consumption, which is closely associated with the observation and coding matrices. To optimize the trade-off in performance, we have studied the NP-hard optimization problem to minimize the network power consumption for given QoS constraints by jointly designing the observation and coding matrices. The CCP based algorithm with low-complexity has been used to provide an effective local optimum to this NP-hard problem. To further improve the performance, we have adopted the collaboration of nodes, where the observation diversity is utilized to suppress the observation noise and equalize the bad observation. Furthermore, taking full account of the collaboration effect in performance, the minimum power consumption of IoT to satisfy the given QoS requirements has been achieved through optimization.

Appendix A
Proof of Lemma 1

The SINR of phenomenon s_m $(m \in \mathcal{M})$ can be written as

$$\begin{aligned}\mathrm{SINR}_m &= \mathbf{d}_m^H \left(\sum_{n \neq m}^{M} \mathbf{d}_n \mathbf{d}_n^H + \mathbf{H}\mathbf{R}_1\mathbf{H}^H + \sigma^2 \mathbf{I} \right)^{-1} \mathbf{d}_m \\ &= \mu_m s_m + \nu_m.\end{aligned}$$

According to the distribution of the sum of scaled exponential random variables and eigenvalues on random matrix theory [43], the variance of ν_m is $\sigma_{\nu_m}^2 = \mu_m - \mu_m^2$ and SINR_m is $\mathrm{SINR}_m = \mu_m/(1-\mu_m)$, where

$$\mu_m = \frac{(\mathbf{H}\left[\mathbf{A}\circ\mathbf{C}\right]_{\cdot m})^H \mathbf{B}_m \mathbf{H}\left[\mathbf{A}\circ\mathbf{C}\right]_{\cdot m}}{1+(\mathbf{H}\left[\mathbf{A}\circ\mathbf{C}\right]_{\cdot m})^H \mathbf{B}_m \mathbf{H}\left[\mathbf{A}\circ\mathbf{C}\right]_{\cdot m}},$$

$$\mathbf{B}_m = \left(\sum_{n \neq m}^{M} \mathbf{d}_n \mathbf{d}_n^H + \mathbf{H}\mathbf{R}_1\mathbf{H}^H + \sigma^2 \mathbf{I} \right)^{-1}.$$

With the eigenvalue theory [44], the Hermitian matrix $\sum_{n\neq m}^{M} \mathbf{d}_n \mathbf{d}_n^H$ can be decomposed as $\sum_{n\neq m}^{M} \mathbf{d}_n \mathbf{d}_n^H = \mathbf{U}_m^H \mathbf{\Lambda}_m \mathbf{U}_m$, where $\mathbf{U}_m$ is the unitary matrix and $\mathbf{\Lambda}_m$ is the eigenvalue matrix defined as

$$\mathbf{\Lambda}_m = \mathrm{diag}\left[\lambda_1, \cdots, \lambda_{M-1}, \underbrace{0, \cdots, 0}_{N-M+1}\right].$$

According to [45, eq. (10)], SINR_m can be expressed as

$$\begin{aligned}\mathrm{SINR}_m &= (\mathbf{U}_m\mathbf{H}\left[\mathbf{A}\circ\mathbf{C}\right]_{\cdot m})^H (\mathbf{\Lambda}_m + \rho\mathbf{I}_N)^{-1} \mathbf{U}_m\mathbf{H}[\mathbf{A}\circ\mathbf{C}]_{\cdot m} \\ &= \underbrace{\sum_{n=1}^{M-1} \frac{|\hat{h}_{n,m}|^2}{\lambda_n+\rho}}_{Z} + \underbrace{\rho^{-1} \sum_{n=M}^{N} |\hat{h}_{n,m}|^2}_{W}\end{aligned} \tag{43}$$

where $\rho = \|\mathbf{A}\circ\mathbf{C}\|_2^2\sigma_o^2 + \sigma^2$, $\mathbf{U}_m\mathbf{H}[\mathbf{A}\circ\mathbf{C}]_{\cdot m} = [\hat{h}_{1,m}, \cdots, \hat{h}_{N,m}]$. Since $\mathbf{H}$ is the circularly symmetric complex Gaussian (CSCG) matrix, $\mathbf{U}_m\mathbf{H}\left[\mathbf{A}\circ\mathbf{C}\right]_{\cdot m}$ is a CSCG vector with variance matrix $\|\left[\mathbf{A}\circ\mathbf{C}\right]_{\cdot m}\|_2^2\mathbf{I}_N$, i.e., $\hat{h}_{n,m} \sim \mathcal{CN}(0, \|\left[\mathbf{A}\circ\mathbf{C}\right]_{\cdot m}\|_2^2)$. Thus, $|\hat{h}_{n,m}|^2$ is i.i.d. exponential random variable

$$f(z;\beta_m) = \begin{cases} \frac{1}{\beta_m}\mathrm{e}^{-\frac{z}{\beta_m}} & z \geq 0; \\ 1 & z < 0, \end{cases} \tag{44}$$

where $\beta_m = \|\left[\mathbf{A}\circ\mathbf{C}\right]_{\cdot m}\|_2^2$. The PDF of Z can be given by

$$f_Z(z|\lambda_1,\ldots,\lambda_{M-1}) = \sum_{n=1}^{M-1} \frac{(\lambda_n+\rho)^{3-M}\mathrm{e}^{-\frac{z(\lambda_n+\rho)}{\beta_m}}}{\prod_{j=1,j\neq n}^{M-1}\left[\left(\frac{1}{\lambda_n+\rho}\right)-\left(\frac{1}{\lambda_j+\rho}\right)\right]\beta_m}. \tag{45}$$

The PDF of W can be given by [43]

$$f_W(w) = \frac{w^{N-M}}{(N-M)!}\left(\frac{\rho}{\beta_m}\right)^{N-M+1} \exp\left(-\frac{\rho w}{\beta_m}\right). \tag{46}$$

The joint PDF of the ordered eigenvalues is

$$f_{\mathbf{\Lambda}}(\lambda_1,\ldots,\lambda_{M-1}) = \prod_{m=1}^{M-1} \frac{\lambda_m^{N-M+1}\mathrm{e}^{-\lambda_m}}{(M-i-1)(N-i)!} \prod_{j=1,j>m}^{M-1} (\lambda_m-\lambda_j)^2. \tag{47}$$

The PDF of SINR_m can be derived by marginalization as

$$\begin{aligned}f_{\mathrm{SINR}_m}(\gamma) &= \int_0^\infty \cdots \int_0^{\lambda_3}\int_0^{\lambda_2} f_{\mathrm{SINR}_m}(\gamma|\lambda_1,\ldots,\lambda_{M-1}) \\ &\cdot f_{\mathbf{\Lambda}}(\lambda_1,\ldots,\lambda_{M-1}) d\lambda_1\cdots d\lambda_{M-1} \\ &= \frac{1}{\beta_m}\frac{(\frac{\gamma}{\beta_m})^{N-1}\mathrm{e}^{-\frac{\rho\gamma}{\beta_m}}}{(N-1)!(1+\frac{\rho^N\gamma}{\beta_m})^M}\left[\sum_{k=0}^{M-1}\binom{M-1}{k}\frac{N!}{(N-k)!}\rho^{-k}\right. \\ &\left. + \frac{\gamma}{\beta_m}\sum_{k=0}^{M-1}\binom{M-1}{k}\frac{(N-1)!}{(N-k-1)!}\rho^{-k}\right].\end{aligned} \tag{48}$$

Appendix B
Proof of Proposition 1

With the integral result in [31], that

$$\begin{aligned}\mathcal{I}(a,b,z,\gamma) &= \frac{\mathrm{e}^z}{\Gamma(a)}\int_1^{\gamma+1} \mathrm{e}^{-zt}(t-1)^{a-1}t^{b-a-1}\mathrm{d}t \\ &= \frac{\mathrm{e}^z}{\Gamma(a)}\sum_{i=0}^{a-1}\binom{a-1}{i}(-1)^i\int_1^{\gamma+1}\mathrm{e}^{-zt}t^{b-i-2}\mathrm{d}t \\ &= \frac{\mathrm{e}^z}{\Gamma(a)}\sum_{i=0}^{a-1}\binom{a-1}{i}\Big[\Gamma(b-i-1,z) \\ &\quad - \Gamma\Big(b-i-1,(1+\gamma)z\Big)\Big](-1)^i z^{i-b+1},\end{aligned}$$

we can obtain the outage probability

$$\begin{aligned}
P_{\mathrm{o},m} &= \int_0^{\gamma_{\mathrm{T}}} f_{\mathrm{SINR}_m}(t)\mathrm{d}t \\
&= \frac{1}{\beta_m}\sum_{k=0}^{M-1}\binom{M-1}{k}\frac{N\gamma_0^{k-N}}{(N-k)!}\int_0^{\gamma_{\mathrm{T}}}\frac{(\frac{t}{\beta_m})^{N-1}\mathrm{e}^{-\frac{t}{\beta_m\gamma_0}}}{(1+\frac{t}{\beta_m})^M}\mathrm{d}t \\
&\quad + \frac{1}{\beta_m}\sum_{k=0}^{M-1}\binom{M-1}{k}\frac{\gamma_0^{k-N}}{(N-k-1)!}\int_0^{\gamma_{\mathrm{T}}}\frac{(\frac{t}{\beta_m})^{N}\mathrm{e}^{-\frac{t}{\beta_m\gamma_0}}}{(1+\frac{t}{\beta_m})^M}\mathrm{d}t \\
&= \sum_{k=0}^{M-1}\binom{M-1}{k}\frac{N\gamma_0^{k-N}}{(N-k)!}\mathcal{I}\Big(N,N-M+1,\gamma_0^{-1},\frac{\gamma_{\mathrm{T}}}{\beta_m}\Big) \\
&\quad + \sum_{k=0}^{M-1}\Bigg[\binom{M-1}{k}\frac{\gamma_0^{k-N}}{(N-k-1)!} \\
&\quad \cdot \mathcal{I}\Big(N+1,N-M+2,\gamma_0^{-1},\frac{\gamma_{\mathrm{T}}}{\beta_m}\Big)\Bigg],
\end{aligned} \tag{49}$$

where $\gamma_0 = \rho^{-1}$.

Appendix C
Proof of Proposition 2

Only considering dominant terms in (14), the diversity order is given by

$$\begin{aligned}
d &= -\lim_{\gamma_0\to\infty}\frac{\log(P_{\mathrm{o}}(\gamma_0))}{\log(\gamma_0)} \\
&= \lim_{\gamma_0\to\infty}\log\Bigg[\gamma_0^{M-1-N}\Big(\frac{N\mathcal{I}(N,N-M+1,\gamma_0^{-1},\frac{\gamma_{\mathrm{T}}}{\beta_m})}{(N-M+1)!} \\
&\quad + \frac{\mathcal{I}(N+1,N-M+2,\gamma_0^{-1},\frac{\gamma_{\mathrm{T}}}{\beta_m})}{(N-M)!}\Big)\Bigg]\Big/\log(\gamma_0) \\
&= N-M+1.
\end{aligned} \tag{50}$$

Appendix D
Proof of Lemma 5

The received SINRs through the MMSE receiver is greater than the received SINRs through the maximal ratio combining (MRC) receiver. Thus, we utilize the SINRs of MRC receiver $\gamma_{\mathrm{MRC},m}$ to provide the lower bounds of SINR_m $(m\in\mathcal{M})$, which are given as

$$\frac{\left|[\mathbf{A}]_{\cdot m}^H\mathbf{H}^H\mathbf{H}[\mathbf{A}]_m\right|^2}{\sum\limits_{n\neq m}^{M}\left|[\mathbf{A}]_{\cdot m}^H\mathbf{H}^H\mathbf{H}[\mathbf{A}]_n\right|^2+\left|[\mathbf{A}]_{\cdot m}^H\mathbf{H}^H\mathbf{H}\mathbf{n}_o\right|^2+\sigma^2\left\|[\mathbf{A}]_{\cdot m}^H\mathbf{H}\right\|_2^2}.$$

While substituting the SINRs through MRC receiver $\gamma_{\mathrm{MRC},m}, m\in\mathcal{M}$ into the optimization problem, the QoS constraints are still intractable as variables are complexly coupled with each other. In order to simplify QoS constraints, we provide the lower bounds for $\gamma_{\mathrm{MRC},m}, m\in\mathcal{M}$ by utilizing the induced norm inequality $\lambda_{\min}\le\|\mathbf{Qb}\|_2^2/\|\mathbf{b}\|_2^2\le\lambda_{\max}$ [44], where $\lambda_{\min}$ and $\lambda_{\max}$ are the minimum and maximum eigenvalue of $\mathbf{Q}^H\mathbf{Q}$, respectively. The lower bounds of SINRs γ_m $(m\in\mathcal{M})$ are given by

$$\frac{[\mathbf{A}]_{\cdot m}^H\mathbf{H}^H\mathbf{H}[\mathbf{A}]_{\cdot m}}{\sum\limits_{n\neq m}^{M}[\mathbf{A}]_{\cdot n}^H\mathbf{H}^H\mathbf{H}[\mathbf{A}]_{\cdot n}+\sum\limits_{k=1}^{K}\|[\mathbf{A}]_{k\cdot}\|_2^2\,\mathbf{h}_k^H\mathbf{h}_k\sigma_o^2+\sigma^2}.$$

Appendix E
Proof of Proposition 4

In short, to solve $\mathcal{P}_{\mathrm{c}}$, we first decouple the coding matrix $\mathbf{A}$ from the observation matrix $\mathbf{C}$ and equivalently rewrite $\mathcal{P}_{\mathrm{c}}$ as a sparse collaboration coding problem with smoothed l_0-norm. Then, we utilize the lower bounds of $\mathrm{SINR}_{c,m}, m\in\mathcal{M}$ and the upper bound of objective function to transform the sparse collaboration coding optimization problem into a solvable DC programming problem. For brevity, we straightway provide the sparse collaboration coding problem with smoothed l_0-norm

$$\begin{aligned}
\mathcal{P}_{\mathrm{SOC}}: \min_{\mathbf{A}} \quad & \sum_{m=1}^{M}\sum_{k=1}^{K} f_\theta\left(|a_{k,m}|^2\right)u_{c,m} \\
& + \sum_{k=1}^{K} f_\theta\left(\|[\mathbf{A}]_{k\cdot}\|_2^2\right)P_k^g \\
& + \mathrm{tr}\left(\mathbf{A}\left(\mathbf{I}+\mathbf{R}_3\right)\mathbf{A}^H\right) \\
\text{s.t.} \quad & C_6: \mathrm{SINR}_{c,m}\ge\tilde{\gamma}_{g,m}, m\in\mathcal{M}.
\end{aligned} \tag{51}$$

The problem $\mathcal{P}_{\mathrm{SOC}}$ is intractable to solve due to the non-convex objective function and QoS constraints C_6. In order to convert $\mathcal{P}_{\mathrm{SOC}}$ into a general DC programming, we first provide the lower bounds for the SINRs through MMSE receiver and the upper bound for the objective function. The SINRs through MMSE receiver $\mathrm{SINR}_{c,m}, m\in\mathcal{M}$ in the collaboration scheme are greater than the SINRs through MRC receiver $\gamma_{\mathrm{MRC},m}^c, m\in\mathcal{M}$, which are given as

$$\frac{[\mathbf{A}]_{\cdot m}^H\mathbf{H}^H\mathbf{H}[\mathbf{A}]_{\cdot m}}{\sum\limits_{n\neq m}^{M}[\mathbf{A}]_{\cdot m}^H\mathbf{H}^H\mathbf{H}[\mathbf{A}]_{\cdot n}+\sum\limits_{m=1}^{M}\frac{\sigma_o^2[\mathbf{A}]_{\cdot m}^H\mathbf{H}^H\mathbf{H}[\mathbf{A}]_{\cdot m}}{\|[\mathbf{C}]_{\cdot m}\|_0}+\sigma^2}.$$

Then, we can provide the lower bound for $\gamma_{\mathrm{MRC},m}^c, m\in\mathcal{M}$ by utilizing the induced norm inequality. The lower bounds $\gamma_{c,m}$ are given as

$$\frac{[\mathbf{A}]_{\cdot m}^H\mathbf{H}^H\mathbf{H}[\mathbf{A}]_{\cdot m}}{(1+\sigma_o^2)\sum\limits_{n\neq m}^{M}[\mathbf{A}]_{\cdot n}^H\mathbf{H}^H\mathbf{H}[\mathbf{A}]_{\cdot n}+\sigma_o^2[\mathbf{A}]_{\cdot m}^H\mathbf{H}^H\mathbf{H}[\mathbf{A}]_{\cdot m}+\sigma^2}.$$

According to the theory of induced norm inequality, there is an inequality

$$\mathrm{tr}\left(\mathbf{A}\mathbf{R}_3\mathbf{A}^H\right)\le\sigma_o^2\sum_{m=1}^{M}\|[\mathbf{A}]_{\cdot m}\|_2^2.$$

Substituting the inequality into the objective function, we can obtain the upper bound of objective function, which is given as

$$\begin{aligned}
&\sum_{m=1}^{M}\sum_{k=1}^{K} f_\theta\left(|a_{k,m}|^2\right)u_{c,m}+\sum_{k=1}^{K} f_\theta\left(\|[\mathbf{A}]_{k\cdot}\|_2^2\right)P_k^g \\
&+\left(1+\sigma_o^2\right)\sum_{m=1}^{M}\|[\mathbf{A}]_{\cdot m}\|_2^2.
\end{aligned}$$

With the upper bound of objective function and the lower bounds of SINRs, introducing the auxiliary variable matrix $\mathbf{T}$, the sparse collaboration coding problem $\mathcal{P}_{\text{SOC}}$ is rewritten as $\mathcal{P}_4$. For problem $\mathcal{P}_4$, its constraints C_4 are convex. And the objective function and constraints C_7 are DC functions. Thus, problem $\mathcal{P}_4$ falls into a general form of DC programming problem.

REFERENCES

[1] X. Deng, Y. Jiang, L. T. Yang, M. Lin, L. Yi, and M. Wang, "Data fusion based coverage optimization in heterogeneous sensor networks: A survey," *Inform. Fusion*, vol. 52, pp. 90–105, 2019.

[2] B. C. Csáji, Z. Kemény, G. Pedone, A. Kuti, and J. Váncza, "Wireless multi-sensor networks for smart cities: A prototype system with statistical data analysis," *IEEE Sens. J.*, vol. 17, no. 23, pp. 7667–7676, 2017.

[3] L. M. Oliveira and J. J. Rodrigues, "Wireless sensor networks: A survey on environmental monitoring." *JCM*, vol. 6, no. 2, pp. 143–151, 2011.

[4] A. Shirazinia, S. Dey, D. Ciuonzo, and P. S. Rossi, "Massive MIMO for decentralized estimation of a correlated source," *IEEE Trans. Signal Process.*, vol. 64, no. 10, pp. 2499–2512, 2016.

[5] F. Jiang, J. Chen, A. L. Swindlehurst, and J. A. López-Salcedo, "Massive MIMO for wireless sensing with a coherent multiple access channel," *IEEE Trans. Signal Process.*, vol. 63, no. 12, pp. 3005–3017, 2015.

[6] S. Joshi and S. Boyd, "Sensor selection via convex optimization," *IEEE Trans. Signal Process.*, vol. 57, no. 2, pp. 451–462, 2009.

[7] Y. Mo, R. Ambrosino, and B. Sinopoli, "Sensor selection strategies for state estimation in energy constrained wireless sensor networks," *Automatica*, vol. 47, no. 7, pp. 1330–1338, 2011.

[8] Y. Wang and D. Wang, "Energy-efficient node selection for target tracking in wireless sensor networks," *Int. J. Distrib. Sens. Netw.*, vol. 9, no. 1, p. 830950, 2013.

[9] S. P. Chepuri and G. Leus, "Sparse sensing for distributed detection," *IEEE Trans. Signal Process.*, vol. 64, no. 6, pp. 1446–1460, 2015.

[10] J. Luo, D. Wu, C. Pan, and J. Zha, "Optimal energy strategy for node selection and data relay in WSN-based IoT," *Mob. Netw. Appl.*, vol. 20, no. 2, pp. 169–180, 2015.

[11] P. Zhang, I. Nevat, G. W. Peters, F. Septier, and M. A. Osborne, "Spatial field reconstruction and sensor selection in heterogeneous sensor networks with stochastic energy harvesting," *IEEE Trans. Signal Process.*, vol. 66, no. 9, pp. 2245–2257, 2018.

[12] F. Kazemeyni, E. B. Johnsen, O. Owe, and I. Balasingham, "Formal modeling and validation of a power-efficient grouping protocol for WSNs," *J. Log. Algebr. Program.*, vol. 81, no. 3, pp. 284–297, 2012.

[13] K. Beydoun and V. Felea, "Energy-efficient wsn infrastructure," in *Proc. 2008 International Symposium on Collaborative Technologies and Systems*, 2008, pp. 58–65.

[14] X. Yu, P. Wu, W. Han, and Z. Zhang, "A survey on wireless sensor network infrastructure for agriculture," *Comput. Stand. Interfaces*, vol. 35, no. 1, pp. 59–64, 2013.

[15] L. Xu, R. Collier, and G. M. O'Hare, "A survey of clustering techniques in WSNs and consideration of the challenges of applying such to 5G IoT scenarios," *IEEE IoT J.*, vol. 4, no. 5, pp. 1229–1249, 2017.

[16] J. Fang and H. Li, "Power constrained distributed estimation with cluster-based sensor collaboration," *IEEE Trans. Wireless Commun.*, vol. 8, no. 7, pp. 3822–3832, 2009.

[17] S. Kar and P. K. Varshney, "Linear coherent estimation with spatial collaboration," *IEEE Trans. Inf. Theory*, vol. 59, no. 6, pp. 3532–3553, 2013.

[18] M. Fanaei, M. C. Valenti, A. Jamalipour, and N. A. Schmid, "Optimal power allocation for distributed blue estimation with linear spatial collaboration," in *Proc. 2014 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, 2014, pp. 5452–5456.

[19] S. Kar and P. K. Varshney, "Controlled collaboration for linear coherent estimation in wireless sensor networks," in *Proc. 2012 50th Annual Allerton Conference on Communication, Control, and Computing (Allerton)*, 2012, pp. 334–341.

[20] S. P. Chepuri and G. Leus, "Sparsity-promoting sensor selection for non-linear measurement models," *IEEE Trans. Signal Process.*, vol. 63, no. 3, pp. 684–698, 2014.

[21] H. Jamali-Rad, A. Simonetto, and G. Leus, "Sparsity-aware sensor selection: Centralized and distributed algorithms," *IEEE Signal Process Lett.*, vol. 21, no. 2, pp. 217–220, 2014.

[22] E. Masazade, M. Fardad, and P. K. Varshney, "Sparsity-promoting extended Kalman filtering for target tracking in wireless sensor networks," *IEEE Signal Process Lett.*, vol. 19, no. 12, pp. 845–848, 2012.

[23] S. Liu, E. Masazade, M. Fardad, and P. K. Varshney, "Sparsity-aware field estimation via ordinary kriging," in *Proc. 2014 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, 2014, pp. 3948–3952.

[24] S. Liu, M. Fardad, E. Masazade, and P. K. Varshney, "Optimal periodic sensor scheduling in networks of dynamical systems," *IEEE Trans. Signal Process.*, vol. 62, no. 12, pp. 3055–3068, 2014.

[25] I. D. Schizas, "Distributed informative-sensor identification via sparsity-aware matrix decomposition," *IEEE Trans. Signal Process.*, vol. 61, no. 18, pp. 4610–4624, 2013.

[26] S. Liu, S. Kar, M. Fardad, and P. K. Varshney, "Sparsity-aware sensor collaboration for linear coherent estimation," *IEEE Trans. Signal Process.*, vol. 63, no. 10, pp. 2582–2596, 2015.

[27] C. Alippi, G. Anastasi, M. Di Francesco, and M. Roveri, "Energy management in wireless sensor networks with energy-hungry sensors," *IEEE Instrum. Meas. Mag.*, vol. 12, no. 2, pp. 16–23, 2009.

[28] M. A. Razzaque and S. Dobson, "Energy-efficient sensing in wireless sensor networks using compressed sensing," *Sensors*, vol. 14, no. 2, pp. 2822–2859, 2014.

[29] T. Bouguera, J.-F. Diouris, J.-J. Chaillout, R. Jaouadi, and G. Andrieux, "Energy consumption model for sensor nodes based on LoRa and LoRaWAN," *Sensors*, vol. 18, no. 7, p. 2104, 2018.

[30] G. Auer, V. Giannini, C. Desset, I. Godor, P. Skillermark, M. Olsson, M. A. Imran, D. Sabella, M. J. Gonzalez, O. Blume *et al.*, "How much energy is needed to run a wireless network?" *IEEE Wirel. Commun.*, vol. 18, no. 5, pp. 40–49, 2011.

[31] I. S. Gradshteyn and I. M. Ryzhik, "Table of integrals, series and products," *Math. Comput.*, vol. 20, no. 96, p. 1157–1160, 2014.

[32] H. Lim and D. Yoon, "On the distribution of sinr for mmse MIMO systems," *IEEE Trans. Commun.*, vol. 67, no. 6, pp. 4035–4046, 2019.

[33] K. Zhai, Z. Ma, and X. Lei, "Closed-formed distribution for the sinr of mmse-detected MIMO systems and performance analysis," *AEU Int. J. Electron. Commun.*, vol. 97, pp. 16–24, 2018.

[34] H. A. Le Thi, T. P. Dinh, H. M. Le, and X. T. Vo, "DC approximation approaches for sparse optimization," *Eur. J. Oper. Res.*, vol. 244, no. 1, pp. 26–46, 2015.

[35] J. Weston, A. Elisseeff, B. Schölkopf, and M. Tipping, "Use of the zero-norm with linear models and kernel methods," *J. Mach. Learn. Res.*, vol. 3, no. Mar, pp. 1439–1461, 2003.

[36] F. Rinaldi, F. Schoen, and M. Sciandrone, "Concave programming for minimizing the zero-norm over polyhedral sets," *Comput. Optim. Appl.*, vol. 46, no. 3, pp. 467–486, 2010.

[37] T. Lipp and S. Boyd, "Variations and extension of the convex–concave procedure," *Optim. Eng.*, vol. 17, no. 2, pp. 263–287, 2016.

[38] M. Tao, E. Chen, H. Zhou, and W. Yu, "Content-centric sparse multicast beamforming for cache-enabled cloud RAN," *IEEE Trans. Wireless Commun.*, vol. 15, no. 9, pp. 6118–6131, 2016.

[39] E. Karipidis, N. D. Sidiropoulos, and Z.-Q. Luo, "Quality of service and max-min fair transmit beamforming to multiple cochannel multicast groups," *IEEE Trans. Signal Process.*, vol. 56, no. 3, pp. 1268–1279, 2008.

[40] Z. Xiang, M. Tao, and X. Wang, "Coordinated multicast beamforming in multicell networks," *IEEE Trans. Wireless Commun.*, vol. 12, no. 1, pp. 12–21, 2012.

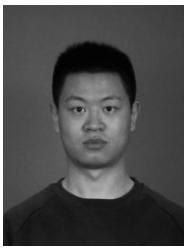
[41] H. A. Le Thi, T. P. Dinh *et al.*, "DC programming and DCA for general DC programs," in *Advanced Computational Methods for Knowledge Engineering*. Springer, 2014, pp. 15–35.

[42] S. Boyd, S. P. Boyd, and L. Vandenberghe, *Convex Optimization*. Cambridge university press, 2004.

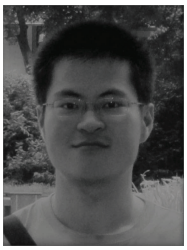
[43] A. M. Tulino, S. Verdú, and S. Verdu, *Random matrix theory and wireless communications*. Now Publishers Inc, 2004.

[44] R. A. Horn and C. R. Johnson, *Matrix Analysis*. Cambridge University Press, 2012.

[45] N. Kim, Y. Lee, and H. Park, "Performance analysis of mimo system with linear mmse receiver," *IEEE Trans. Wireless Commun.*, vol. 7, no. 11, pp. 4474–4478, 2008.



Chengcheng Han received the B.S. degree in electronic and information engineering from the University of Electronic Science and Technology of China, Chengdu, China, in 2016 and the Ph.D.degree in electrical engineering from the University of Science and Technology of China, Hefei, China, in 2021. He is currently a staff with the Department of Wireless Network Research of Huawei, Beijing. His research interest includes wireless communication and sensing.



Li Chen received the B.E. in electrical and information engineering from Harbin Institute of Technology, Harbin, China, in 2009 and the Ph.D.degree in electrical engineering from the University of Science and Technology of China, Hefei, China, in 2014. He is currently a faculty member with the Department of Electronic Engineering and Information Science, University of Science and Technology of China. His research interests include wireless IoT communications and wireless optical communications.

Nan Zhao (S'08-M'11-SM'16) is currently a Professor at Dalian University of Technology, China. He received the B.S. degree in electronics and information engineering in 2005, the M.E. degree in signal and information processing in 2007, and the Ph.D. degree in information and communication engineering in 2011, from Harbin Institute of Technology, Harbin, China. His recent research interests include UAV Communications, Interference Alignment, and Physical Layer Security.

Dr. Zhao is serving or served on the editorial boards of 7 SCI-indexed journals. He received Top Reviewer Award from IEEE Transactions on Vehicular Technology in 2016, and was nominated as an Exemplary Reviewer by IEEE Communications Letters in 2016. He won the best paper awards in IEEE VTC'2017-Spring and MLICOM 2017.

Yunfei Chen (Senior Member, IEEE) received his B.E. and M.E. degrees in electronics engineering from Shanghai Jiaotong University, Shanghai, P.R.China, in 1998 and 2001, respectively. He received his Ph.D. degree from the University of Alberta in 2006. He is currently working as an Associate Professor at the University of Warwick, U.K. His research interests include wireless communications, cognitive radios, wireless relaying and energy harvesting.

F. Richard Yu (S'00-M'04-SM'08-F'18) received the PhD degree in electrical engineering from the University of British Columbia (UBC) in 2003. From 2002 to 2006, he was with Ericsson (in Lund, Sweden) and a start-up in California, USA. He joined Carleton University in 2007, where he is currently a Professor. He received the IEEE Outstanding Service Award in 2016, IEEE Outstanding Leadership Award in 2013, Carleton Research Achievement Award in 2012, the Ontario Early Researcher Award (formerly Premiers Research Excellence Award) in 2011, the Excellent Contribution Award at IEEE/IFIP TrustCom 2010, the Leadership Opportunity Fund Award from Canada Foundation of Innovation in 2009 and the Best Paper Awards at IEEE ICNC 2018, VTC 2017 Spring, ICC 2014, Globecom 2012, IEEE/IFIP TrustCom 2009 and Int'l Conference on Networking 2005. His research interests include wireless cyber-physical systems, connected/autonomous vehicles, security, distributed ledger technology, and deep learning.

He serves on the editorial boards of several journals, including Co-Editor-in-Chief for Ad Hoc & Sensor Wireless Networks, Lead Series Editor for IEEE Transactions on Vehicular Technology, IEEE Transactions on Green Communications and Networking, and IEEE Communications Surveys & Tutorials. He has served as the Technical Program Committee (TPC) Co-Chair of numerous conferences. Dr. Yu is a registered Professional Engineer in the province of Ontario, Canada, a Fellow of the Institution of Engineering and Technology (IET), and a Fellow of the IEEE. He is a Distinguished Lecturer, the Vice President (Membership), and an elected member of the Board of Governors (BoG) of the IEEE Vehicular Technology Society.