Massive Uncoordinated Multiple Access for Beyond 5G

Mohammadkarimi, Mostafa; Dobre, Octavia A.; Win, Moe Z.

DOI
10.1109/TWC.2021.3117256

Publication date
2022

Document Version
Accepted author manuscript

Published in
IEEE Transactions on Wireless Communications

Massive Uncoordinated Multiple Access for Beyond 5G

Mostafa Mohammadkarimi, *Member, IEEE*, Octavia A. Dobre, *Fellow, IEEE*, and Moe Z. Win, *Fellow, IEEE*

***Abstract*—Existing wireless communication systems have been mainly designed to provide substantial gain in terms of data rates. However, 5G and Beyond will depart from this scheme, with the objective not only to provide services with higher data rates. One of the main goals is to support massive machine-type communications (mMTC) in the Internet-of-Things (IoT) applications. Supporting massive uplink communications for devices with sporadic traffic pattern and short-packet size, as it is in many mMTC use cases, is a challenging task, particularly when the control signaling is not negligible in size compared to the payload. In addition, channel estimation becomes challenging for sporadic and short-packet transmission due to the limited number of employed pilots. In this paper, a new uplink multiple access (MA) scheme is proposed for mMTC, which can support a large number of uncoordinated IoT devices with short-packet and sporadic traffic. The proposed uplink MA scheme removes the overheads associated with the device identifier as well as pilots and preambles related to channel estimation. An alternative mechanism for device identification (DI) is employed, where a unique spreading code is dedicated to each IoT device as identifier. This unique code is simultaneously used for the spreading purpose and DI. Two IoT DI algorithms which employ sparse signal reconstruction methods are proposed to determine the active IoT devices prior to data detection. Specifically, the Bayesian information criterion model order selection method is employed to develop an IoT DI algorithm for unknown and time-varying activity rate. Our proposed MA scheme benefits from a new non-coherent nonlinear multiuser detection algorithm designed on the basis of unsupervised machine learning techniques to enable data detection without *a priori* knowledge on channel state information. For performance improvement, an extension to multiple receive antennas through hard decision combining is proposed. The effectiveness of the proposed MA scheme for known and unknown activity rate and high overloading factor is supported by simulation results.**

***Index Terms*—Internet-of-Things (IoT), massive machine-type communications (mMTC), Beyond 5G, uplink multiple access, sparse signal reconstruction, nonlinear multiuser detection, sporadic transmission, machine learning, multiple antennas.**

I. Introduction

MASSIVE UPLINK connectivity is the key factor in the realization of the Internet-of-Things (IoT), as part of 5G and Beyond wireless communication systems [1]. In many IoT applications, massive machine-type communications (mMTC) services are required, where a large number of devices transmit very short packets sporadically. Typically, the number of IoT devices assigned to each base station (BS) in mMTC is in orders of magnitude above what current communication networks are capable to support. Moreover, IoT devices do not transmit continuously, rather updates are infrequently transmitted to the BS, whenever a measured value changes. Hence, small packets are expected to carry critical payload in mMTC [2].

The design of the current wireless communication systems relies on the assumption that the control signaling related to physical (PHY) and media access control (MAC) layers is of negligible size compared to the payload. Thus, heuristic design of control signaling is acceptable and does not affect the overall system performance. However, in mMTC with short-packet transmission, the control signaling can be similar in size with the payload; thus, inefficient design of control signaling leads to highly suboptimal transmission schemes. Excessive control signaling, e.g., the overheads, preambles, and pilots associated with device identifier, exploited for channel estimation, and used for random access procedure, hinders massive connectivity [3]. Thus, efficient multiple access (MA) schemes with highly limited (or non-existent) control signaling are required.

Moreover, channel estimation is another challenge for sporadic and short-packet transmission, especially for a massive number of non-orthogonal transmissions. Existing channel estimation approaches are often based on the assumption that devices are active over long periods so that channel estimation through pilots and preambles is feasible. However, if an IoT device only transmits occasionally, such an assumption cannot longer be valid. Instead, channel estimation has to rely on a single transmission that may be very short, which constrains the number of orthogonal pilots [4]. Channel estimation becomes more challenging in the grant-free uplink MA scheme, where resources are randomly selected by devices [5].

Motivated by these facts, a new uplink MA scheme for short-packet and sporadic traffic in mMTC is proposed in this paper. The main idea behind the proposed MA scheme is to reduce the control signaling while simultaneously supporting a massive number of uncoordinated IoT devices with a single BS. The proposed MA scheme is designed based on asynchronous direct-sequence spread spectrum (DS-SS) with non-orthogonal spreading codebook, and is capable of supporting undetermined DS-SS systems in static networks, where the BS and IoT devices are immobile.

To remove the control signaling associated with the IoT device identifier, a unique spreading code is dedicated to each IoT device which is simultaneously used for the spreading purpose and device identification (DI). In a nutshell, instead of allocating a fragment of the IoT packet to the signaling associated with the MAC address (device identifier), the unique spreading code is used as IoT device identifier. Our MA scheme also relies on an unsupervised machine learning technique to

This research was supported by the Natural Sciences and Engineering Research Council of Canada (NSERC) through its Discovery program.

M. Mohammadkarimi is with the Faculty of Electrical Engineering, Mathematics and Computer Science, Delft University of Technology, Delft, Netherlands (e-mail: m.mohammadkarimi@tudelft.nl).

O. A. Dobre is with the Faculty of Engineering and Applied Science, Memorial University, St. John's, NL, Canada, (e-mail: odobre@mun.ca).

M. Z. Win is with the Laboratory for Information and Decision Systems (LIDS), Massachusetts Institute of Technology, Boston, MA, USA (e-mail: moewin@mit.edu).

enable non-coherent data detection, thus removing the need of preambles and pilots used for channel estimation. The lack of preambles and pilots further reduces the control signaling.

Our proposed approach for removing the device identifier relies on sparsity-aware IoT DI at the BS to determine the active IoT devices before data detection. Based on the sporadic traffic pattern of the IoT devices as well as lack of knowledge about the channel state information (CSI) of the IoT devices, the squared ℓ_2-norm sparse signal reconstruction (SSR) and Bayesian information criterion (BIC) $\ell_1-\ell_2$ mixed-norm simultaneous sparse signal reconstruction (SSSR) IoT DI algorithms are developed. In the former algorithm, the IoT identification problem is formulated as an SSR using the generalized cross-validation (GCV) approach followed by parallel hypothesis testing. The latter algorithm formulates the IoT DI problem as a BIC model order selection SSSR problem.

The proposed uplink MA scheme is also equipped with a new non-coherent nonlinear multiuser detection (MUD) algorithm to detect data of the active IoT devices, applied after the IoT DI algorithms. We propose the non-coherent 2-mean clustering (2-MC)-MUD algorithm based on 2-MC unsupervised machine learning and differential coding to detect data without channel estimation at the BS.

A. Related Works

IoT device activity and data detection techniques can be categorized into three groups: 1) regularized, 2) greedy, and 3) iterative-thresholding based methods [6]. The regularized methods apply a regularization parameter into the cost function which balances both approximation error and sparsity level of the solution. Sparse maximum a posteriori probability (S-MAP) and its relaxed versions were the first algorithms for activity and data detection that took into account regularization parameter [7]. To achieve an acceptable performance with lower computational complexity compared to the optimal S-MAP detector, several detectors were then proposed, such as sparsity-aware successive interference cancellation (SA-SIC) [8], SA-SIC with sorted QR decomposition [9], activity-aware multiple feedback SIC [10], activity-aware recursive least squares with decision feedback [11], and direction method of multipliers (ADMM) [12].

In the greedy techniques, the goal is to solve the sparse representation with the ℓ_0-norm minimization. Because of the fact that this problem is NP-hard, the greedy technique provides an approximate solution to alleviate this difficulty. The greedy strategy searches for the best local optimal solution in each iteration with the goal of achieving the optimal holistic solution. The pioneering work in this category applied the orthogonal least squares and orthogonal matching pursuit (OMP) algorithms to perform joint detection of data and device activity for mMTC [13]. Other detection algorithms in this category are: group OMP (GOMP) [14], compressive sample matching pursuit, detection-based OMP [15], weighted GOMP [16], detecting-based GOMP, block-correlation SIC [17], simultaneous OMP with extrinsic information transfer, and the threshold aided block sparsity adaptive subspace pursuit [18].

Iterative-thresholding based methods are alternatives to convex optimization for large-scale problems. These methods are inspired by belief propagation (BP) in graphical models, where their foundation is Gaussian loopy BP with simplified message passing that assumes high dimensional signal vectors in order to factorize a multivariate distribution. There are several approaches in this category, such as approximate message passing (AMP) with non-separable denoiser [19], vector AMP with MMSE denoiser [20], bilinear generalized AMP [21], joint expected maximization and AMP [22], and mixture of compressive sensing and message passing [23]. A blind BP detection algorithm for non-coherent non-orthogonal MA with massive receive antennas was also proposed in [24].

Most of the mentioned approaches either assume perfect CSI or its estimate through pilot at the BS. While using pilot results in lower spectral efficiency and higher latency in these approaches compared to our non-coherent method, some of these approaches, such as [20], can offer lower packet error rate (PER) when a sufficient number of pilots is used for joint activity detection and channel estimation. On the other hand, the single phase joint activity detection, channel estimation, and data detection approaches without requiring pilot, such as the structured sparsity learning MUD algorithm in [21], exhibit higher PER compared to our non-coherent method due to lack of spreading. Moreover, unlike most of the above-mentioned approaches, which considered coordinated uplink MA with perfectly synchronized transmissions, our proposed MA scheme supports uncoordinated IoT devices. It is worth mentioning that the algorithm in [21] is very promising for symbol-level synchronized transmission, and [20] can analytically characterize channel estimation error. Besides, existing regularized algorithms do not propose any solution to set the tuning parameter in the optimization problem when the activity rate is unknown and time-varying. The degree of sparsity, and thus, the false alarm and correct identification rates depend on the value of the tuning parameter.

B. Contributions

The main contributions of this work are as follows:

- A new uplink MA scheme is proposed for mMTC. The proposed MA scheme exhibits the following advantages:
  - It is capable to support thousands of uncoordinated IoT devices;
  - It supports sporadic traffic pattern and short-packet;
  - It significantly reduces packet time on-air since it is designed for underdetermined DS-SS (number of devices is larger than the spreading factor);
  - It removes the control signaling associated with the device identifier as well as pilots and preambles employed for channel estimation to reduce uplink overhead;
  - It exhibits high scalability in terms of adding new IoT devices (high overloading factor) without negatively affecting the system performance.
- A new mechanism for the IoT DI at the BS is developed instead of using device identifier. Since the active IoT devices in the network do not use a device identifier in order to identify themselves to the BS, the squared ℓ_2-norm SSR and the BIC $\ell_1-\ell_2$ mixed-norm SSSR IoT

identification algorithms are proposed to detect active IoT devices. The proposed algorithms exhibit the following advantages:
  - They can detect active IoT devices without knowledge of the CSI;
  - They remove the need for matched-filter (MF) implementation for all spreading codes; thus reducing the complexity of the receiver;
  - The BIC $\ell_1 - \ell_2$ mixed-norm SSSR algorithm can identify active IoT devices when the activity rate is unknown and time-varying;
  - They take the advantage of optimal tuning parameter;
  - The BIC $\ell_1 - \ell_2$ mixed-norm SSSR algorithm can identify active IoT devices for non-identical activity rate due to the BIC model selection;
  - There is control over the false alarm and correct identification rates of the individual IoT devices in the ℓ_2-norm SSR IoT DI algorithm.
- The statistical performance analysis of the squared ℓ_2-norm SSR IoT DI algorithm is presented, and theoretical expressions for the correct identification and false alarm rates are derived.
- A new non-coherent nonlinear MUD algorithm, i.e., 2-MC-MUD in combination with differential coding is designed for short packet transmission. The proposed 2-MC-MUD algorithm exhibits the following advantages:
  - It supports both coordinated and uncoordinated DS-SS transmission irrespective of the traffic pattern;
  - It does not require knowledge of the CSI at the BS.
- An extension to multiple receive antennas through hard decision combining is proposed. This combination offers the following advantages:
  - The performance of the proposed uplink MA boosts because of spatial diversity;
  - Higher overloading factor can be supported.

C. Notations

The identity matrix and zero vector are shown by $\mathbf{I}$ and $\mathbf{0}$, and the indicator function is defined as $\mathbb{I}\{x\} = 1$ if x is true; otherwise, $\mathbb{I}\{x\} = 0$. The cardinality of a set, which measures the number of elements of the set, is denoted by $\mathbf{card}(\cdot)$. The ℓ_0 quasi-norm of vector $\mathbf{a}_j = [a_{0,j}\ a_{1,j}\ \ldots\ a_{m-1,j}]^\dagger$ and the $\ell_0 - \ell_0$ quasi-norm of matrix $\mathbf{A} \triangleq [\mathbf{a}_0\ \mathbf{a}_1\ \ldots\ \mathbf{a}_{n-1}]$ are respectively defined as $\|\mathbf{a}_j\|_0 \triangleq \mathbf{card}\big(\{i \in \mathcal{I} | a_{i,j} \neq 0\}\big)$, and $\|\mathbf{A}\|_0 \triangleq \mathbf{card}\big(\{i \in \mathcal{I} | \exists j, j = 0, 1, \ldots, n-1, a_{i,j} \neq 0\}\big)$, where $\mathcal{I} \triangleq \{0, 1, \ldots, m-1\}$. We use $\mathrm{tr}(\mathbf{B})$, $\mathbf{B}^{-1}$ and $\det(\mathbf{B})$ to show the trace, inverse, and determinant of an square matrix $\mathbf{B}$. We also employ $\mathrm{diag}(\mathbf{B})$ to represent the diagonal elements of $\mathbf{B}$ in vector form. Throughout the paper, $(\cdot)^*$, $(\cdot)^\dagger$, and $(\cdot)^{\mathrm{H}}$ show the complex conjugate, transpose, and Hermitian transpose, respectively. Also, $|\cdot|$, $\lfloor\cdot\rfloor$, and $\otimes$ represent the absolute value operator, floor function, and Kronecker product, respectively. $\mathbb{E}\{\cdot\}$ is the statistical expectation, $\hat{x}$ is an estimate of x. The complex Gaussian distribution with mean vector $\boldsymbol{\mu}$ and covariance matrix $\boldsymbol{\Sigma}$ is denoted by $\mathcal{CN}(\boldsymbol{\mu}, \boldsymbol{\Sigma})$.

The remaining of the paper is organized as follows. Section II introduces the system model. Section III describes the proposed IoT DI algorithms and presents their analytical performance evaluation. In Section IV, the data detection problem is discussed, and the nonlinear 2-MC-MUD algorithm is proposed. An extension to multiple receive antennas is discussed in V. Simulation results are provided in Section VI, and conclusions are drawn in Section VII.

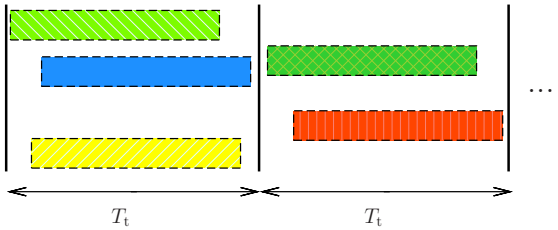


Fig. 1: Received packets at the BS.

II. System Model

Consider K_{u} IoT devices communicating with a single IoT BS in a single-hop communication. It is considered that the IoT devices transmit data in short packets over independent doubly block fading channel, where the fading channel is block fading in time and in frequency. The activity rate for each IoT device is assumed to be P_{a}.[1] The IoT devices transmit their packet after receiving a beacon signal transmitted by the IoT BS. This signal is periodically transmitted with period $T_{\mathrm{t}} = N_{\mathrm{s}}T_{\mathrm{s}} + \tau_{\max}$, where N_{s} is the number of symbols per IoT packet, T_{s} is the symbol duration, and $\tau_{\max}$ is the known maximum delay of the single-hop IoT network. It is assumed that T_{t} equals the coherence time of the fading channel.

We denote $\mathcal{X}_{\mathrm{u}} \triangleq \{0, 1, \ldots, K_{\mathrm{u}} - 1\}$ and $\mathcal{X}_{\mathrm{a}}$ the total and active IoT devices in the network, respectively. The round-trip delay of the kth IoT device is shown by $\tau_k \triangleq 2d_k/c$, $\tau_k \in [0, \tau_{\max}]$, where c is the speed of light, and d_k is the distance between the k-th IoT device and the BS. We consider that τ_k, $k, \in \mathcal{X}_{\mathrm{u}}$, is known at the receiver. Fig. 1 illustrates the received IoT packets at the BS.

As illustrated in Fig. 2, for each IoT device, the payload bits $\mathbf{d}_k$, $k \in \mathcal{X}_{\mathrm{a}}$, are encoded by the channel encoder to increase the reliability of packet transmission. Then, the encoded data is passed through the differential encoding block. Differential encoding is employed to remove the need of channel estimation in the MUD at the BS to enable non-coherent detection. After differential encoding, the data is binary phase-shift-keying (BPSK) modulated. Finally, the modulated signal is multiplied by a unique spreading waveform and then transmitted. It is considered that the spreading waveforms of the IoT devices do not change over time.

The impulse response of the doubly block fading channel for the kth IoT device is given as $g_k(t) \approx \breve{g}_k \delta(t - \tau_k)$, where $\breve{g}_k$ is the fading coefficient of the kth IoT device, which is constant during a packet but changes to an independent value for the next packet. Doubly block fading channel is a suitable model for sporadic traffic and short packet transmission. The received baseband signal over doubly block fading channel in each transmission period with respect to the timing reference of the BS is modeled as

[1] Both known and unknown activity rate are studied in this paper.

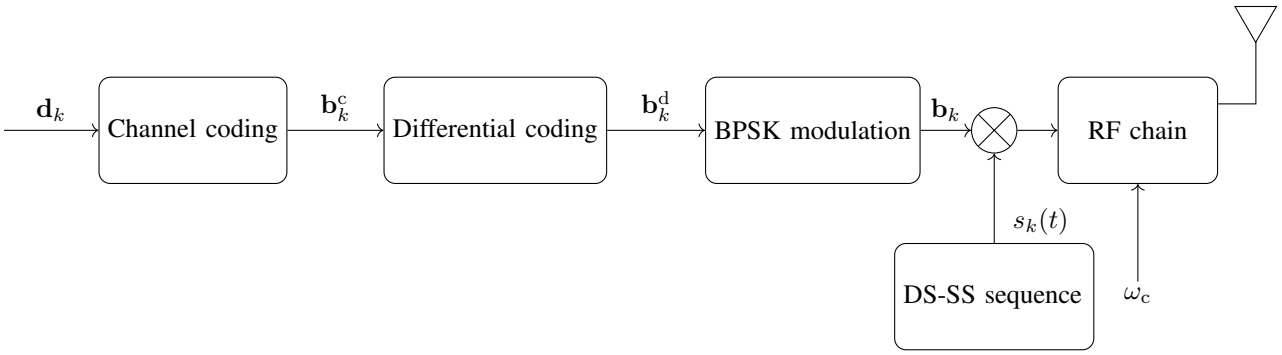


Fig. 2: Block diagram that illustrates packet transmission at IoT devices.

$$r(t) = \sum_{k=0}^{K_{\mathrm{u}}-1} \sum_{n=0}^{N_{\mathrm{s}}-1} \breve{g}_k \sqrt{\eta_k p_k} e^{j\phi_k} b_{k,n} s_k(t - nT_{\mathrm{s}} - \tau_k) + w(t)$$
$$= \sum_{k=0}^{K_{\mathrm{u}}-1} \sum_{n=0}^{N_{\mathrm{s}}-1} g_k b_{k,n} s_k(t - nT_{\mathrm{s}} - \tau_k) + w(t), \quad (1)$$

where $t \in [0, T_{\mathrm{t}}]$, $g_k \triangleq \breve{g}_k \sqrt{\eta_k p_k} e^{j\phi_k}$, and $\breve{g}_k$, ϕ_k, and $\{b_{k,n}, n = 0, 1, \ldots, N_{\mathrm{s}} - 1\}$ respectively denote the fading channel coefficient, carrier phase (CP), and symbol stream of the kth IoT device, which are unknown at the BS. Also, $\eta_k = \left(\frac{\lambda_{\mathrm{c}}}{4\pi d_k}\right)^2$ and p_k denote the pathloss and transmit power of the kth IoT device, respectively, where λ_{c} is the wavelength of the carrier signal. It is considered that $\breve{g}_k \sim \mathcal{CN}(\mu_k, \sigma_k^2)$, and the envelope of the CSI, i.e., $|\breve{g}_k|$ has a Rician distribution with K-factor $|\mu_k|^2/(\sigma_k^2)$. The symbol stream for the inactive IoT devices is modeled as transmitting zeros during the packet, i.e., $b_{k,n} = 0$, $n = 0, 1, \ldots, N_{\mathrm{s}} - 1$, while active IoT devices employ BPSK modulation with $\mathbb{E}\{|b_{k,n}|^2\} = 1$. The DS-SS signaling waveform of the kth IoT device, $s_k(t)$, is given by

$$s_k(t) = \sum_{m=0}^{N_{\mathrm{c}}-1} c_k^{(m)} \psi(t - mT_{\mathrm{c}}), \qquad t \in [0, T_{\mathrm{s}}], \quad (2)$$

where T_{c} is the chip duration, $\mathbf{c}_k = \left[c_k^{(0)} \; c_k^{(1)} \; \ldots \; c_k^{(N_{\mathrm{c}}-1)}\right]^\dagger$ is the spreading sequence of $\{+1, -1\}$ assigned to the kth IoT device, N_{c} is the spreading factor, and $\psi(t)$ is the chip waveform with unit power. It is assumed that $\psi(t)$ is a rectangular pulse confined within $[0, T_{\mathrm{c}}]$. To support massive connectivity, $K_{\mathrm{u}} > N_{\mathrm{c}}$, which leads to non-orthogonal transmission. The baseband additive complex Gaussian noise at the output of the receive filter with bandwidth $1/T_{\mathrm{c}}$ is denoted by $w(t)$ in (1).

Fig. 3 shows the block diagram of the proposed receiver at the IoT BS. As seen, the received baseband signal is passed through the chip MF and sampled at the chip rate. The output of the sampled chip MF for the ith chip at the jth observation symbol is obtained as

$$r_j^{(i)} \triangleq \int_{jT_{\mathrm{s}}+iT_{\mathrm{c}}}^{jT_{\mathrm{s}}+(i+1)T_{\mathrm{c}}} r(t)\psi\left(t - jT_{\mathrm{s}} - iT_{\mathrm{c}}\right) dt \quad (3)$$
$$= \sum_{k=0}^{K_{\mathrm{u}}-1} g_k u_{k,j}^{(i)} + w_j^{(i)} \qquad i = 0, 1, \ldots, N_{\mathrm{c}} - 1,$$

where

$$w_j^{(i)} \triangleq \int_{jT_{\mathrm{s}}+iT_{\mathrm{c}}}^{jT_{\mathrm{s}}+(i+1)T_{\mathrm{c}}} w(t)\psi\left(t - jT_{\mathrm{s}} - iT_{\mathrm{c}}\right) dt, \quad (4)$$

and

$$u_{k,j}^{(i)} \triangleq \int_{jT_{\mathrm{s}}+iT_{\mathrm{c}}}^{jT_{\mathrm{s}}+(i+1)T_{\mathrm{c}}} \sum_{n=0}^{N_{\mathrm{s}}-1} b_{k,n} s_k(t - nT_{\mathrm{s}} - \tau_k) \times \psi(t - jT_{\mathrm{s}} - iT_{\mathrm{c}}) dt. \quad (5)$$

By employing (4), one can show that the joint probability density function (PDF) of the corresponding noise vector associated with the jth observation vector, i.e., $\mathbf{w}_j \triangleq \left[w_j^{(0)} \; w_j^{(1)} \; \ldots \; w_j^{(N_{\mathrm{c}}-1)}\right]^\dagger$ is characterized by $\mathbf{w}_j \sim \mathcal{CN}\left(\mathbf{0}_{N_{\mathrm{c}}}, \sigma_{\mathrm{w}}^2 \mathbf{I}\right)$ with $\sigma_{\mathrm{w}}^2 \triangleq N_0/T_{\mathrm{c}}$, where $N_0/2$ is the power spectral density of the white noise. The integral in (5) represents the area under the received signal waveform of the kth IoT device during the ith chip-matched filtering duration at the jth observation symbol.

Let us write the delay of the kth IoT as

$$\tau_k \triangleq \alpha_k T_{\mathrm{s}} + \beta_k T_{\mathrm{c}} + \xi_k, \quad (6)$$

with $\alpha_k \triangleq \lfloor \tau_k/T_{\mathrm{s}} \rfloor$, $\beta_k \triangleq \lfloor \tau_k/T_{\mathrm{c}} \rfloor - \alpha_k N_{\mathrm{c}}$, and $\xi_k \in [0, T_{\mathrm{c}})$.

Based on the values of α_k, β_k, and ξ_k, $u_{k,j}^{(i)}$ in (5) is expressed as a function of $b_{k,j-\alpha_k}$ and $b_{k,j-\alpha_k-1}$ as [25]

$$u_{k,j}^{(i)} \triangleq \sum_{n=0}^{N_{\mathrm{s}}-1} \sum_{m=0}^{N_{\mathrm{c}}-1} c_k^{(m)} b_{k,n} \times \int_{jT_{\mathrm{s}}+iT_{\mathrm{c}}}^{jT_{\mathrm{s}}+(i+1)T_{\mathrm{c}}} \psi(t - nT_{\mathrm{s}} - mT_{\mathrm{c}} - \tau_k)\psi(t - jT_{\mathrm{s}} - iT_{\mathrm{c}}) dt$$
$$= b_{k,j-\alpha_k-1} x_k^{(i)}(1 - \xi_k) + b_{k,j-\alpha_k} x_k^{(i)}(\xi_k), \quad (7)$$

where

$$x_k^{(i)}(\nu) \triangleq \sum_{m=0}^{N_{\mathrm{c}}-1} c_k^{(m)} \int_{iT_{\mathrm{c}}}^{(i+1)T_{\mathrm{c}}} \psi(t - mT_{\mathrm{c}} - \nu T_{\mathrm{c}})\psi(t - iT_{\mathrm{c}}) dt, \quad (8)$$

with $\nu \in [0, 1)$. We can write (7) in vector form as follows

$$\mathbf{u}_{k,j} = b_{k,j-\alpha_k-1}\mathbf{x}_{k,0} + b_{k,j-\alpha_k}\mathbf{x}_{k,1} \quad (9)$$

where $b_{k,j} = 0$ when $j \notin [0, N_{\mathrm{s}} - 1]$, and

$$\mathbf{u}_{k,j} \triangleq \left[u_{k,j}^{(0)} \; u_{k,j}^{(1)} \; \ldots \; u_{k,j}^{(N_{\mathrm{c}}-1)}\right]^\dagger, \quad (10a)$$

$$\mathbf{x}_{k,1} \triangleq \left[x_k^{(0)}(\xi_k) \; x_k^{(1)}(\xi_k) \; \ldots \; x_k^{(N_{\mathrm{c}}-1)}(\xi_k)\right]^\dagger, \quad (10b)$$

$$\mathbf{x}_{k,0} \triangleq \left[x_k^{(0)}(1 - \xi_k) \; x_k^{(1)}(1 - \xi_k) \; \ldots \; x_k^{(N_{\mathrm{c}}-1)}(1 - \xi_k)\right]^\dagger.$$

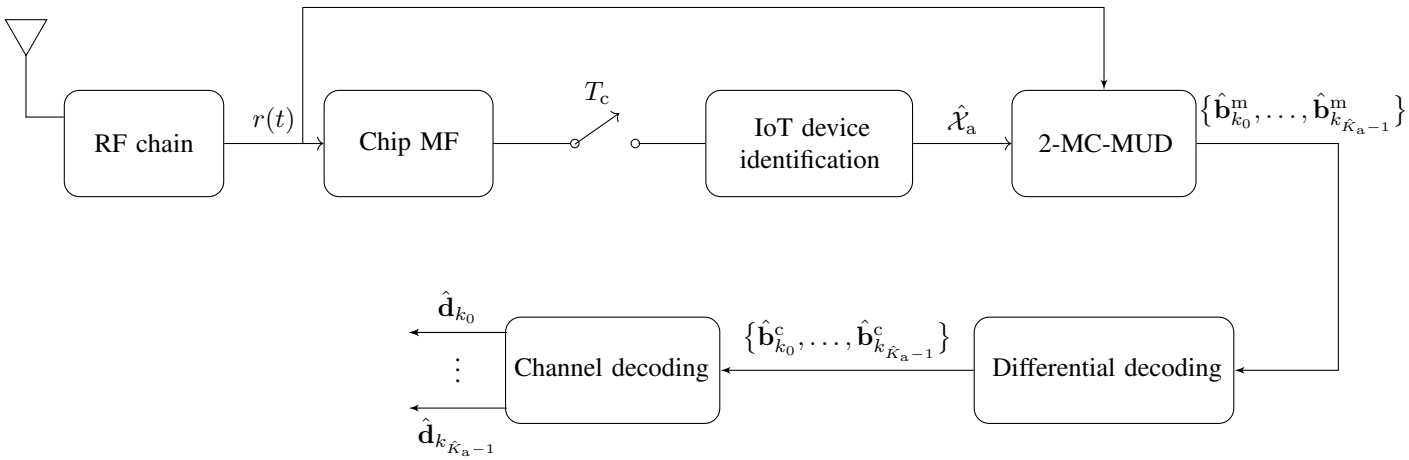


Fig. 3: Block diagram of the proposed receiver at the BS.

For the rectangular chip waveform $\psi(t)$, we can obtain

$$\begin{bmatrix}\mathbf{x}_{k,1}\\ \\ \mathbf{x}_{k,0}\end{bmatrix} = (1-\xi_k)\begin{bmatrix}\mathbf{0}_{\beta_k}\\ \mathbf{c}_k\\ \mathbf{0}_{N_{\rm c}-\beta_k}\end{bmatrix} + \xi_k\begin{bmatrix}\mathbf{0}_{\beta_k+1}\\ \mathbf{c}_k\\ \mathbf{0}_{N_{\rm c}-\beta_k-1}\end{bmatrix}. \quad (11)$$

Let us define $\mathbf{X}_k \triangleq \left[\mathbf{x}_{k,0}\ \mathbf{x}_{k,1}\right]$. By employing (3) and (9), the jth observation vector, i.e., $\mathbf{r}_j \triangleq \left[\mathbf{r}_j^{(0)}\ \mathbf{r}_j^{(1)}\ \ldots\ \mathbf{r}_j^{(N_{\rm c}-1)}\right]^\dagger$, is written as follows

$$\mathbf{r}_j = \mathbf{X}\mathbf{G}\mathbf{b}_j + \mathbf{w}_j = \mathbf{X}\mathbf{h}_j + \mathbf{w}_j, \quad (12)$$

where

$$\mathbf{X} \triangleq \left[\mathbf{X}_0\ \ \mathbf{X}_1\ \ \ldots\ \ \mathbf{X}_{K_{\rm u}-1}\right], \quad (13)$$

$$\mathbf{G} \triangleq \begin{bmatrix} g_0 & & & 0\\ & g_1 & & \\ & & \ddots & \\ 0 & & & g_{K_{\rm u}-1}\end{bmatrix} \otimes \mathbf{I}_2, \quad (14)$$

$$\mathbf{b}_j \triangleq \Big[b_{0,j-\alpha_0-1}\ \ b_{0,j-\alpha_0}\ \ b_{1,j-\alpha_1-1}\ \ b_{1,j-\alpha_1}\ \ \ldots\ \ b_{K_{\rm u}-1,j-\alpha_{(K_{\rm u}-1)}-1}\ \ b_{K_{\rm u}-1,j-\alpha_{(K_{\rm u}-1)}}\Big]^\dagger, \quad (15)$$

and

$$\mathbf{h}_j \triangleq \Big[h_{0,j,0}\ h_{0,j,1}\ h_{1,j,0}\ h_{1,j,1}\ \ldots\ h_{K_{\rm u}-1,j,0}\ h_{K_{\rm u}-1,j,1}\Big]^\dagger \quad (16)$$

with

$$h_{k,j,f} \triangleq g_k b_{k,j-\alpha_k-1+f}, \quad f\in\{0,1\}. \quad (17)$$

Finally, by staking the $N_{\rm t}$ observation vectors, the observation matrix is written as follows

$$\mathbf{R}_{\rm T} = \mathbf{X}\mathbf{G}\mathbf{B}_{\rm T} + \mathbf{W}_{\rm T} = \mathbf{X}\mathbf{H}_{\rm T} + \mathbf{W}_{\rm T}, \quad (18)$$

where $\mathbf{R}_{\rm T} \triangleq [\mathbf{r}_0\ \mathbf{r}_1\ \ldots\ \mathbf{r}_{N_{\rm t}-1}]$, $\mathbf{B}_{\rm T} \triangleq [\mathbf{b}_0\ \mathbf{b}_1\ \ldots\ \mathbf{b}_{N_{\rm t}-1}]$, $\mathbf{W}_{\rm T} \triangleq [\mathbf{w}_0\ \mathbf{w}_1\ \ldots\ \mathbf{w}_{N_{\rm t}-1}]$, and $\mathbf{H}_{\rm T} \triangleq [\mathbf{h}_0\ \mathbf{h}_1\ \ldots\ \mathbf{h}_{N_{\rm t}-1}]$. In (18), $\mathbf{X}$ is referred to as dictionary.

As seen in Fig. 3, after chip-matched filtering and sampling, the IoT DI algorithm is applied to the measurement matrix $\mathbf{R}_{\rm T}$ to detect the active IoT devices. The outcome of the IoT DI algorithm is a set of IoT devices $\hat{\mathcal{X}}_{\rm a}$. Then, the MUD algorithm is applied to detect data of the IoT devices in $\hat{\mathcal{X}}_{\rm a}$. After MUD, the bit streams related to the active IoT devices pass through differential and channel decoders, respectively. In the remaining of the paper, we propose different algorithms to realize the system in Fig. 3.

III. IoT DI

DI is the first step in uplink MA schemes that devices do not use control signaling in order to identify themselves to the BS. In this case, the BS needs to determine the active devices before data detection. In this section, different IoT DI algorithms are developed.

A. IoT DI: Problem Formulation

For the sake of decreasing the complexity, a portion of the observation window can be employed for IoT DI. Let us consider a truncated observation window of length L as

$$\mathbf{R} = \mathbf{X}\mathbf{G}\mathbf{B} + \mathbf{W} = \mathbf{X}\mathbf{H} + \mathbf{W}, \quad (19)$$

where $\mathbf{R} \triangleq [\mathbf{r}_{\bar\alpha}\ \mathbf{r}_{\bar\alpha+1}\ \ldots\ \mathbf{r}_{\bar\alpha+L-1}]$, $\mathbf{B} \triangleq [\mathbf{b}_{\bar\alpha}\ \mathbf{b}_{\bar\alpha+1}\ \ldots\ \mathbf{b}_{\bar\alpha+L-1}]$, $\mathbf{W} \triangleq [\mathbf{w}_{\bar\alpha}\ \mathbf{w}_{\bar\alpha+1}\ \ldots\ \mathbf{w}_{\bar\alpha+L-1}]$, and $\mathbf{H} \triangleq [\mathbf{h}_{\bar\alpha}\ \mathbf{h}_{\bar\alpha+1}\ \ldots\ \mathbf{h}_{\bar\alpha+L-1}]$ with $1 \leqslant L \leqslant N_{\rm s}+\alpha_{\min}-\bar\alpha$, where $\bar\alpha$ is an arbitrary positive integer, $\bar\alpha > \alpha_{\max} \triangleq \max\{\alpha_0,\alpha_1,\ldots,\alpha_{K_{\rm u}-1}\}$, and $\alpha_{\min} \triangleq \min\{\alpha_0,\alpha_1,\ldots,\alpha_{K_{\rm u}-1}\}$.[2] Fig. 4 shows the underdetermined system of linear equations in (18), and Fig. 5 illustrates truncated observation windows for IoT DI in (19).

The activity of an IoT device is defined for an entire packet, i.e, the rows of $\mathbf{H}$ corresponding to the active and inactive IoT devices are non-zero and zero, respectively. Thus, the problem of IoT DI for the kth IoT device, $k\in\mathcal{X}_{\rm u}$, can be expressed as the following binary hypothesis testing:

$$\begin{aligned} H_{1k}:&\quad \mathbf{h}_k^{\{\bar\alpha,L\}} \neq \mathbf{0}\\ H_{0k}:&\quad \mathbf{h}_k^{\{\bar\alpha,L\}} = \mathbf{0}, \end{aligned} \quad (20)$$

where

$$\mathbf{h}_k^{\{\bar\alpha,L\}} \triangleq \left[\mathbf{h}_{k,\bar\alpha}^\dagger \mathbf{h}_{k,\bar\alpha+1}^\dagger\ \ldots\ \mathbf{h}_{k,\bar\alpha+L-1}^\dagger\right]^\dagger, \quad (21a)$$

$$\mathbf{h}_{k,j} \triangleq \left[h_{k,j,0}\ h_{k,j,1}\right]^\dagger, \quad (21b)$$

[2]The number of non-zero elements of the observation window for the active IoT devices is considered the same to facilitate theoretical analysis.

This article has been accepted for publication in IEEE Transactions on Wireless Communications. This is the author's version which has not been fully edited and content may change prior to final publication. Citation information: DOI 10.1109/TWC.2021.3117256

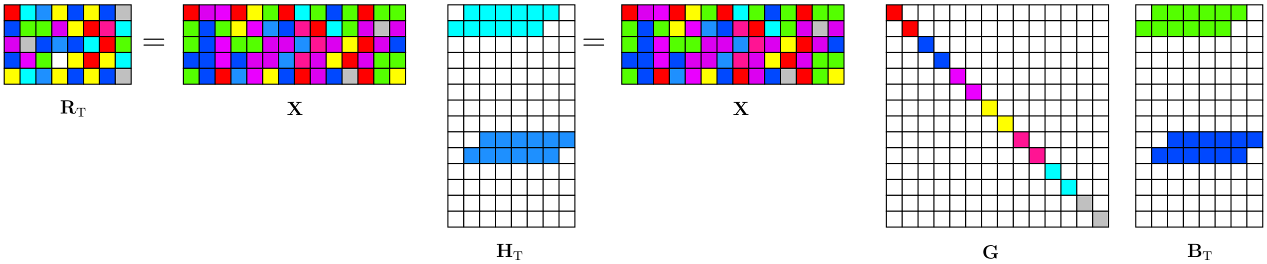


Fig. 4: Underdetermined system of linear equations for $K_{\mathrm{u}}=7$, $K_{\mathrm{a}}=2$, $N_{\mathrm{t}}=8$, $N_{\mathrm{c}}=5$, $\alpha_{\max}=1$, and $N_{\mathrm{s}}=6$. Due to the asynchronicity among the IoT devices, the matrix of the transmitted symbols $\mathbf{B}_{\mathrm{T}}$ includes two rows for each IoT device.

and H_{0k} and H_{1k} are the null and alternative hypotheses denoting that the kth IoT device is active and inactive, respectively. As seen in (20), the IoT DI problem is formulated as K_{u} parallel binary hypothesis testing problems. The first step in IoT DI is to reconstruct $\mathbf{h}_k^{\{\bar{\alpha},L\}}$, $k\in\mathcal{X}_{\mathrm{u}}$, from the truncated observation matrix in (19). .

Let us denote the number of active IoT devices by the random variable $k_{\mathrm{a}}=\mathbf{card}(\mathbf{X}_{\mathrm{a}})$. For $P_{\mathrm{a}}\ll 1$, $\mathbb{P}\{k_{\mathrm{a}}\ll K_{\mathrm{u}}\}=1$, and thus, $\mathbf{B}$ and $\mathbf{H}$ in (19) are sparse matrices. Moreover, the columns of $\mathbf{H}(\mathbf{B})$ share the same sparsity profiles. This sparse structure is referred to as block-sparse. The block-sparse structure of $\mathbf{H}$ can be observed in Fig. 4.

The sparse structure of $\mathbf{H}$ can be employed to reconstruct the columns of $\mathbf{H}$ from the underdetermined linear observation model in (19). When each column of $\mathbf{H}$ is individually reconstructed from its corresponding column in $\mathbf{R}$, it is referred to as SSR. The SSR for the columns of $\mathbf{H}$, i.e., $\mathbf{h}_j$, $\bar{\alpha}\leqslant j\leqslant\bar{\alpha}+L-1$, is formulated as follows

$$\hat{\mathbf{h}}_j=\arg\min_{\mathbf{h}_j}\ \frac{1}{2}\big\|\mathbf{r}_j-\mathbf{X}\mathbf{h}_j\big\|_2^2+\lambda_{\ell_0}\big\|\mathbf{h}_j\big\|_0,\tag{22}$$

where λ_{ℓ_0} is the tuning parameter which balances both approximation error and sparsity level of the solution.

The ℓ_0-minimization in (22) is both numerically unstable and NP-hard since the ℓ_0 quasi-norm is a discrete-value function. One approach to the SSR is to replace the ℓ_0 quasi-norm by a convex function with common sparsity profile that leads to a solution very close to the one of the original problem. Different convex functions can be employed to relax $\|\mathbf{h}_j\|_0$ in (22). A common family of convex functions is the ℓ_q norm, given as

$$\big\|\mathbf{h}_j\big\|_q=\left(\sum_{k=0}^{K_{\mathrm{u}}-1}\sum_{f=0}^{1}\Big|h_{k,j,f}\Big|^q\right)^{\frac{1}{q}}.\tag{23}$$

The recovered vectors by the ℓ_q norm minimization can be employed to infer the active IoT set $\mathcal{X}_{\mathrm{a}}$.

On the other hand, the block-sparse structure of $\mathbf{H}$ can be employed to improve the reconstruction of $\mathbf{H}$ in (22). This method of signal reconstruction is referred to as SSSR. Opposite to SSR, the SSSR simultaneously exploits the column sparsity along with the block-sparse structure in the optimization problem in order to reconstruct the matrix $\mathbf{H}$. The SSSR of $\mathbf{H}$, given $\mathbf{R}$ and the dictionary $\mathbf{X}$, is expressed as

$$\hat{\mathbf{H}}=\arg\min_{\mathbf{H}}\ \frac{1}{2}\big\|\mathbf{R}-\mathbf{X}\mathbf{H}\big\|_{\mathrm{F}}^2+\lambda_{\ell_0}^{\ell_0}\big\|\mathbf{H}\big\|_0,\tag{24}$$

where $\lambda_{\ell_0}^{\ell_0}$ is the tuning parameter which balances both approximation error and sparsity level of the solution. Similar to

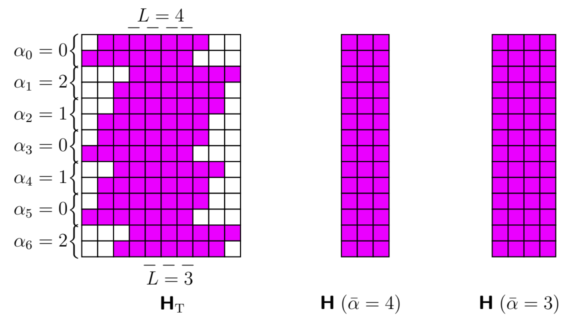


Fig. 5: Different observation windows for IoT DI ($K_{\mathrm{u}}=7$, $N_{\mathrm{s}}=7$, $\alpha_{\max}=2$, $\alpha_{\min}=0$, $1\leqslant L\leqslant 4$). The purple color is employed to show the packet of the IoT devices, which is zero for inactive and non-zero for active IoT devices.

the ℓ_0-minimization in (22), the $\ell_0-\ell_0$-minimization in (24) is unstable and NP-hard. Therefore, the quasi-norm $\big\|\mathbf{H}\big\|_0$ is replaced with the $\ell_p-\ell_q$ $(p,q\geqslant 1)$ mixed-norm as

$$J_{p,q}(\mathbf{H})=\sum_{k=0}^{K_{\mathrm{u}}-1}\big\|\mathbf{h}_k^{\{\bar{\alpha},L\}}\big\|_q^p\tag{25}$$

to convert the combinatorial problem in (24) into a convex optimization problem. The vector $\mathbf{h}_k^{\{\bar{\alpha},L\}}$ in (25) is defined in (21). The recovered matrix by the relaxed SSSR can also be employed to infer the active IoT set.

B. IoT DI for Known Activity Rate

Here, we propose an algorithm for IoT DI when the probability of activity P_{a} is known at the BS. Convex relation through squared ℓ_2-norm followed by a threshold setting mechanism is employed for IoT DI.

1) Squared ℓ_2-Norm SSR IoT DI: The squared ℓ_2-norm convex relaxation form of (22) is given by

$$\hat{\mathbf{h}}_j=\arg\min_{\mathbf{h}_j}\ \frac{1}{2}\big\|\mathbf{r}_j-\mathbf{X}\mathbf{h}_j\big\|_2^2+\lambda\big\|\mathbf{h}_j\big\|_2^2,\tag{26}$$

where $\mathbf{h}_j$ is given in (16) and (17). The squared ℓ_2-norm SSR algorithm formulates the IoT identification problem as a ridge regression (RD) estimation problem as in (26) followed by K_{u} parallel binary hypothesis testing problems. This is because the RD does not set the coefficients of $\hat{\mathbf{h}}_j$ to zero. The optimal solution of (26) is obtained as [26]

$$\hat{\mathbf{h}}_j=\Big(\mathbf{X}^{\dagger}\mathbf{X}+2\lambda\mathbf{I}\Big)^{-1}\mathbf{X}^{\dagger}\mathbf{r}_j,\tag{27}$$

which is a simple linear estimator of $\mathbf{r}_j$ that shrinks ordinary least-squares (LS) estimates towards zero. The tuning parameter λ for SSR can be obtained through cross-validation and GCV [27], [28]. The latter is a method of model selection that

© 2021 IEEE. Personal use is permitted, but republication/redistribution requires IEEE permission. See https://www.ieee.org/publications/rights/index.html for more information.

is widely employed; in this case, λ is obtained as follows [27]

$$\lambda_{\rm cv} = \arg\min_{\lambda} \frac{\left\|(\mathbf{I}-\mathbf{Q})\mathbf{r}_j\right\|_2^2}{\left[\,\mathrm{tr}\,(\mathbf{I}-\mathbf{Q})\right]^2}, \tag{28}$$

where $\mathbf{Q} \triangleq \mathbf{X}\big(\mathbf{X}^\dagger\mathbf{X}+2\lambda\mathbf{I}\big)^{-1}\mathbf{X}^\dagger$. In [29], it has been shown that the optimal tuning parameter of the RD estimator for $\mathbf{r}_j = \mathbf{X}\mathbf{h}_j + \mathbf{w}_j$ in terms of minimum mean squared error can be approximated as follows

$$\lambda_j^{\rm op} \approx \frac{\sigma_{\rm w}^2\,\mathrm{tr}\Big[\bar{\mathbf{\Sigma}}_{\mathbf{X}}^{-1}\Big]}{\mathbf{h}_j^{\rm H}\bar{\mathbf{\Sigma}}_{\mathbf{X}}^{-1}\mathbf{h}_j + 3\,\mathrm{tr}\Big[\bar{\mathbf{\Sigma}}_{\mathbf{X}}^{-2}\Big]}, \tag{29}$$

where $\bar{\mathbf{\Sigma}}_{\mathbf{X}} \triangleq \mathbf{X}^\dagger\mathbf{X}$. As observed, $\lambda_j^{\rm op}$ depends on $\mathbf{h}_j$ which is unknown and needs to be estimated by the RD estimator. In this case, for moderate and high signal-to-noise ratio (SNR) range, an approximation of (29) can be obtained by replacing $\mathbf{h}_j^{\rm H}\bar{\mathbf{\Sigma}}_{\mathbf{X}}^{-1}\mathbf{h}_j$ with its expected value [30]. Since the elements of $\mathbf{h}_j$ are uncorrelated, by employing $\mathbb{E}\big\{|h_{k,j,0}|^2\big\} = \mathbb{E}\big\{|h_{k,j,1}|^2\big\} = P_{\rm a}\eta_k p_k(\sigma_k^2+|\mu_k|^2)$, $k \in \mathcal{X}_{\rm u}$, we can show that $\mathbb{E}\big\{\mathbf{h}_j^{\rm H}\bar{\mathbf{\Sigma}}_{\mathbf{X}}^{-1}\mathbf{h}_j\big\} = P_{\rm a}(\mathbf{\Gamma}^\dagger\otimes\mathbf{1}_2^\dagger)\bar{\mathbf{\Lambda}}_{\mathbf{X}}$, where $\mathbf{\Gamma} \triangleq [\gamma_0\ \gamma_1\ \ldots\ \gamma_{K_{\rm u}-1}]^\dagger$, $\gamma_k \triangleq \eta_k p_k(\sigma_k^2+|\mu_k|^2)$, $\mathbf{1}_2 = [1\ 1]^\dagger$, and $\bar{\mathbf{\Lambda}}_{\mathbf{X}} \triangleq \mathrm{diag}(\bar{\mathbf{\Sigma}}_{\mathbf{X}}^{-1})$. Substituting $\mathbb{E}\big\{\mathbf{h}_j^{\rm H}\bar{\mathbf{\Sigma}}_{\mathbf{X}}^{-1}\mathbf{h}_j\big\} = P_{\rm a}(\mathbf{\Gamma}^\dagger\otimes\mathbf{1}^\dagger)\bar{\mathbf{\Lambda}}_{\mathbf{X}}$ into (29), results in

$$\lambda^{\rm opt} \approx \frac{\sigma_{\rm w}^2\mathrm{tr}\big[\bar{\mathbf{\Sigma}}_{\mathbf{X}}^{-1}\big]}{P_{\rm a}(\mathbf{\Gamma}^\dagger\otimes\mathbf{1}^\dagger)\bar{\mathbf{\Lambda}}_{\mathbf{X}} + 3\mathrm{tr}\big[\bar{\mathbf{\Sigma}}_{\mathbf{X}}^{-2}\big]}. \tag{30}$$

As seen in (30), $\lambda^{\rm opt}$ is inversely proportional to $P_{\rm a}$.

By substituting $\mathbf{r}_j = \mathbf{X}\mathbf{h}_j + \mathbf{w}_j$ in (12) into (27), $\hat{\mathbf{h}}_j$ can be written as a linear function of $\mathbf{h}_j$ as

$$\hat{\mathbf{h}}_j = \mathbf{\Omega}\mathbf{h}_j + \mathbf{w}_j', \tag{31}$$

where

$$\mathbf{\Omega} \triangleq \begin{bmatrix} \Omega_{0,0} & \Omega_{0,1} & \cdots & \Omega_{0,2K_{\rm u}-1} \\ \Omega_{1,0} & \Omega_{1,1} & \cdots & \Omega_{1,2K_{\rm u}-1} \\ \vdots & \vdots & \ddots & \vdots \\ \Omega_{2K_{\rm u}-1,0} & \Omega_{2K_{\rm u}-1,1} & \cdots & \Omega_{2K_{\rm u}-1,2K_{\rm u}-1} \end{bmatrix} = \mathbf{I} - 2\lambda^{\rm opt}\big(\bar{\mathbf{\Sigma}}_{\mathbf{X}} + 2\lambda^{\rm opt}\mathbf{I}\big)^{-1}, \tag{32}$$

and

$$\mathbf{w}_j' \triangleq \begin{bmatrix} w_{0,j,0}' \\ w_{0,j,1}' \\ \vdots \\ w_{K_{\rm u}-1,j,0}' \\ w_{K_{\rm u}-1,j,1}' \end{bmatrix} = \big(\bar{\mathbf{\Sigma}}_{\mathbf{X}} + 2\lambda^{\rm opt}\mathbf{I}\big)^{-1}\mathbf{X}^\dagger\mathbf{w}_j. \tag{33}$$

In (33), $\mathbf{w}_j'$ is zero-mean complex Gaussian colored noise vector with covariance matrix given by

$$\mathbf{\Sigma}^{w'} \triangleq \begin{bmatrix} \Sigma_{0,0}^{w'} & \Sigma_{0,1}^{w'} & \cdots & \Sigma_{0,2K_{\rm u}-1}^{w'} \\ \Sigma_{1,0}^{w'} & \Sigma_{1,1}^{w'} & \cdots & \Sigma_{1,2K_{\rm u}-1}^{w'} \\ \vdots & \vdots & \ddots & \vdots \\ \Sigma_{2K_{\rm u}-1,0}^{w'} & \Sigma_{2K_{\rm u}-1,1}^{w'} & \cdots & \Sigma_{2K_{\rm u}-1,2K_{\rm u}-1}^{w'} \end{bmatrix} = \mathbb{E}\big\{\mathbf{w}_j'\big(\mathbf{w}_j'\big)^{\rm H}\big\} = \sigma_{\rm w}^2\big(\bar{\mathbf{\Sigma}}_{\mathbf{X}} + 2\lambda^{\rm opt}\mathbf{I}\big)^{-2}\bar{\mathbf{\Sigma}}_{\mathbf{X}}, \tag{34}$$

where $\Sigma_{2k_1+f_1,2k_2+f_2}^{w'} = \mathbb{E}\big\{w_{k_1,j,f_1}'(w_{k_2,j,f_2}')^*\big\}$.

The elements of $\hat{\mathbf{h}}_j$ in (31) associated with the kth IoT device, i.e., $\hat{h}_{k,j,0}$ and $\hat{h}_{k,j,1}$ can be written as follows

$$\hat{h}_{k,j,f} = \Omega_{2k+f,2k+f}h_{k,j,f} + \Omega_{2k+f,2k+\bar{f}}h_{k,j,\bar{f}} + \sum_{n\neq k}\Big\{\Omega_{2k+f,2n+f}h_{n,j,f} + \Omega_{2k+f,2n+\bar{f}}h_{n,j,\bar{f}}\Big\} + w_{k,j,f}', \tag{35}$$

where $f,\bar{f} \in \{0,1\}$ and $\bar{f} \triangleq f+(-1)^f$. The second term on the right-hand side of (35) represents the effect of the multiuser interference caused by the active IoT devices in the network. Due to the central limit theorem (CLT), $\hat{h}_{k,j,f}$, $f \in \{0,1\}$, in (35) given hypothesis H_{0k} and H_{1k} can be accurately approximated by complex Gaussian random variables for sufficiently small values of K-factor $\kappa_k \triangleq |\mu_k|^2/\sigma_k^2$ and large enough $P_{\rm a}K_{\rm u}$. Simulation results show that for $\kappa_k \triangleq |\mu_k|^2/\sigma_k^2 < 0.2$, Gaussian assumption is valid. In fact, the lower κ_k, the more reliable the Gaussian assumption is. It should be mentioned that the random variables $\hat{h}_{k,j,0}$ and $\hat{h}_{k,j,1}$ are not joint Gaussian random variables as shown in Fig. 6. The mean, variance, and cross-correlation of $\hat{h}_{k,j,0}$ and $\hat{h}_{k,j,1}$ are given in Lemma 1.

Lemma 1. *First and second order statistics of the reconstructed signal for the kth IoT device in (35), i.e., $\hat{h}_{k,j,0}$ and $\hat{h}_{k,j,1}$, are given as follows*

$$\mathbb{E}\Big\{\hat{h}_{k,j,0}\big|H_{tk}\Big\} = \mathbb{E}\Big\{\hat{h}_{k,j,1}\big|H_{tk}\Big\} = 0, \qquad t \in \{0,1\} \tag{36}$$

$$\Sigma_{f,f}^{tk} \triangleq \mathbb{V}\mathrm{ar}\Big\{\hat{h}_{k,j,f}\big|H_{tk}\Big\} = \mathbb{E}\Big\{\big|\hat{h}_{k,j,f}\big|^2\big|H_{tk}\Big\} = t\gamma_k\Big(\Omega_{2k+f,2k+f}^2 + \Omega_{2k+f,2k+\bar{f}}^2\Big) + P_{\rm a}\sum_{n\neq k}\gamma_n\Big(\Omega_{2k+f,2n+f}^2 + \Omega_{2k+f,2n+\bar{f}}^2\Big) + \Sigma_{2k+f,2k+f}^{w'}, \tag{37}$$

and

$$\Sigma_{0,1}^{tk} = \mathbb{C}\mathrm{ov}\Big\{\hat{h}_{k,j,0},\hat{h}_{k,j,1}\big|H_{tk}\Big\} = \mathbb{E}\Big\{\hat{h}_{k,j,0}\hat{h}_{k,j,1}^*\big|H_{tk}\Big\} = t\gamma_k\Big(\Omega_{2k,2k}\Omega_{2k+1,2k} + \Omega_{2k+1,2k+1}\Omega_{2k,2k+1}\Big) + P_{\rm a}\sum_{n\neq k}\gamma_n\Big(\Omega_{2k,2n}\Omega_{2k+1,2n} + \Omega_{2k+1,2n+1}\Omega_{2k,2n+1}\Big) + \mathbf{\Sigma}_{2k,2k+1}^{\rm w'}, \tag{38}$$

where $\gamma_k = \eta_k p_k(\sigma_k^2+|\mu_k|^2)$, $\Sigma_{1,0}^{tk} = \Sigma_{0,1}^{tk}$, $t,f \in \{0,1\}$, $\bar{f} \triangleq f+(-1)^f$, and $\Sigma_{2k+f,2k+f}^{w'}$ is given in (34) (Proof in Appendix A) □

Since the joint PDF of $\hat{h}_{k,j,0}$ and $\hat{h}_{k,j,1}$ given H_{tk}, i.e., $p\big(\hat{h}_{k,j,0},\ \hat{h}_{k,j,1}|H_{tk}\big)$, $t \in \{0,1\}$, cannot be expressed in a tractable mathematical form and since there is high correlation between $\hat{h}_{k,j,0}$ and $\hat{h}_{k,j,1}$, we can either use $\hat{h}_{k,j,0}$ or $\hat{h}_{k,j,1}$ to identify the transmission state of the kth IoT device. Moreover, the in-phase and quadrature components of $\hat{h}_{k,j,f}$ can be accurately approximated by correlated joint Gaussian random variables due to the CLT for sufficiently small values of K-factor $\kappa_k \triangleq |\mu_k|^2/\sigma_k^2$ and large enough $P_{\rm a}K_{\rm u}$. To verify the credibility of Gaussian assumption, we evaluate the kurtosis and skewness for $\mathrm{Re}\{\hat{h}_{k,j,0}\}$ and $\mathrm{Re}\{\hat{h}_{k,j,1}\}$ in Table I.

Table I: Credibility of Gaussian assumption for $\hat{h}_{k,j,0}$ and $\hat{h}_{k,j,1}$.

Variable	Kurtosis	Skewness	Variance
Gaussian (theory)	3	0	0.002 {0.2657}
$\mathrm{Re}\{\hat{h}_{k,j,0}\}$	3.105	0.0224	0.001
$\mathrm{Re}\{\hat{h}_{k,j,1}\}$	3.014	−0.0233	{0.2643}

Similar to the proof of Lemma 1, we can show that the distribution of the reconstructed signal for the kth IoT device is given as follows

$$\begin{bmatrix}\mathrm{Re}\{\hat{h}_{k,j,0}\}\\ \mathrm{Im}\{\hat{h}_{k,j,0}\}\end{bmatrix} \sim \begin{cases}\mathcal{N}\left(\mathbf{0}, \mathbf{C}_{0,0}^{0k}\right), & H_{0k}\\ \mathcal{N}\left(\mathbf{0}, \mathbf{C}_{0,0}^{1k}\right), & H_{1k}\end{cases}, \tag{39}$$

and

$$\begin{bmatrix}\mathrm{Re}\{\hat{h}_{k,j,1}\}\\ \mathrm{Im}\{\hat{h}_{k,j,1}\}\end{bmatrix} \sim \begin{cases}\mathcal{CN}\left(\mathbf{0}, \mathbf{C}_{1,1}^{0k}\right), & H_{0k}\\ \mathcal{CN}\left(\mathbf{0}, \mathbf{C}_{1,1}^{1k}\right), & H_{1k}\end{cases}, \tag{40}$$

where

$$\mathbf{C}_{f,f}^{tk} = \begin{bmatrix}\bar{\Sigma}_{f,f}^{tk} & \rho_{f,f}^{tk}\\ \rho_{f,f}^{tk} & \tilde{\Sigma}_{f,f}^{tk}\end{bmatrix}, \tag{41}$$

$$\begin{aligned}\rho_{f,f}^{tk} &= \mathbb{E}\Big\{\mathrm{Re}\{\hat{h}_{k,j,f}\}\,\mathrm{Im}\{\hat{h}_{k,j,f}\}\big|H_{tk}\Big\} \\ &= t\bar{\mu}_k\tilde{\mu}_k\eta_k p_k\left(\Omega^2_{2k+f,2k+f} + \Omega^2_{2k+f,2k+\bar{f}}\right) \\ &\quad + P_{\mathrm{a}}\sum_{n\neq k}\bar{\mu}_n\tilde{\mu}_n\eta_n p_n\left(\Omega^2_{2k+f,2n+f} + \Omega^2_{2k+f,2n+\bar{f}}\right),\end{aligned} \tag{42}$$

$$\begin{aligned}\bar{\Sigma}_{f,f}^{tk} &\triangleq \mathbb{V}\mathrm{ar}\Big\{\mathrm{Re}\{\hat{h}_{k,j,f}\}\big|H_{tk}\Big\} = \mathbb{E}\Big\{\big(\mathrm{Re}\{\hat{h}_{k,j,f}\}\big)^2\big|H_{tk}\Big\} \\ &= t(\sigma_k^2/2 + |\bar{\mu}_k|^2)\eta_k p_k\left(\Omega^2_{2k+f,2k+f} + \Omega^2_{2k+f,2k+\bar{f}}\right) \\ &\quad + P_{\mathrm{a}}\sum_{n\neq k}(\sigma_n^2/2 + |\bar{\mu}_n|^2)\eta_n p_n\left(\Omega^2_{2k+f,2n+f} + \Omega^2_{2k+f,2n+\bar{f}}\right) \\ &\quad + \Sigma^{w'}_{2k+f,2k+f}/2,\end{aligned} \tag{43}$$

and

$$\begin{aligned}\tilde{\Sigma}_{f,f}^{tk} &\triangleq \mathbb{V}\mathrm{ar}\Big\{\mathrm{Im}\{\hat{h}_{k,j,f}\}\big|H_{tk}\Big\} = \mathbb{E}\Big\{\big(\mathrm{Im}\{\hat{h}_{k,j,f}\}\big)^2\big|H_{tk}\Big\} \\ &= t(\sigma_k^2/2 + |\tilde{\mu}_k|^2)\eta_k p_k\left(\Omega^2_{2k+f,2k+f} + \Omega^2_{2k+f,2k+\bar{f}}\right) \\ &\quad + P_{\mathrm{a}}\sum_{n\neq k}(\sigma_n^2/2 + |\tilde{\mu}_n|^2)\eta_n p_n\left(\Omega^2_{2k+f,2n+f} + \Omega^2_{2k+f,2n+\bar{f}}\right) \\ &\quad + \Sigma^{w'}_{2k+f,2k+f}/2,\end{aligned} \tag{44}$$

with $\bar{\mu}_k \triangleq \mathrm{Re}\{\mu_k\}$ and $\tilde{\mu}_k \triangleq \mathrm{Im}\{\mu_k\}$.

The larger the ratio of the variances in (39) and (40), i.e., $\Sigma^{1k}_{0,0}/\Sigma^{0k}_{0,0}$, and $\Sigma^{1k}_{1,1}/\Sigma^{0k}_{1,1}$, the better identification performance. Accordingly, we use the reconstructed signal in (35) for the identification of the kth IoT device as follows

$$\breve{\mathbf{h}}_{k,j} \triangleq \begin{bmatrix}\mathrm{Re}\{\hat{h}_{k,j,1}\}\\ \mathrm{Im}\{\hat{h}_{k,j,1}\}\end{bmatrix}\mathbb{I}\{\varrho_k < 0\} + \begin{bmatrix}\mathrm{Re}\{\hat{h}_{k,j,0}\}\\ \mathrm{Im}\{\hat{h}_{k,j,0}\}\end{bmatrix}\mathbb{I}\{\varrho_k \geqslant 0\}, \tag{45}$$

where $\breve{\mathbf{h}}_{k,j} = [\breve{h}_{k,j,0}, \breve{h}_{k,j,1}]^\dagger$, $k \in \mathcal{X}_{\mathrm{u}}$, and

$$\varrho_k = \left(\frac{\Sigma^{1k}_{0,0}}{\Sigma^{0k}_{0,0}} - \frac{\Sigma^{1k}_{1,1}}{\Sigma^{0k}_{1,1}}\right). \tag{46}$$

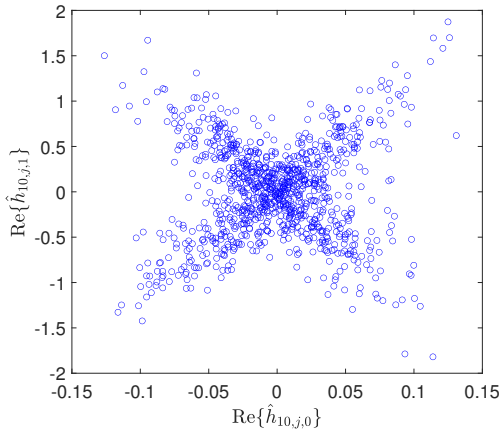


Fig. 6: Scatter plot of $\mathrm{Re}\{\hat{h}_{10,j,1}\}$ and $\mathrm{Re}\{\hat{h}_{10,j,0}\}$ when device $k = 10$ is active, $K_{\mathrm{u}} = 1536$, $N_{\mathrm{c}} = 512$, $P_{\mathrm{a}} = 0.01$, and SNR= 16 dB.

Using the reconstructed signal in the form of (45) enables us to derive closed-form expressions for the correct identification and false alarm rates. In order to identify the transmission state of the kth IoT device, $k \in \mathcal{X}_{\mathrm{u}}$, based on $\breve{\mathbf{h}}_{k,j}$, the maximum likelihood ratio (MLR) test can be used [31].

Lemma 2. *The optimal MLR decision rule for IoT DI based on the reconstructed signal $\breve{\mathbf{h}}_{k,j}$, $k \in \mathcal{X}_{\mathrm{u}}$, in (45) is given by*

$$d_k = \begin{cases}H_{1k}, & \phi\big(\breve{\mathbf{h}}_{k,j}\big) \geqslant \theta_k\\ H_{0k}, & \phi\big(\breve{\mathbf{h}}_{k,j}\big) < \theta_k\end{cases}, \tag{47}$$

where

$$\phi\big(\breve{\mathbf{h}}_{k,j}\big) = \sum_{n=0}^{1}\chi_{f,f}[n]z^2_{k,j}[n], \tag{48}$$

$$\chi_{f,f}[n] \triangleq \frac{\lambda_{f,f}[n]}{\lambda_{f,f}[n]+1}. \tag{49}$$

In (49), $\lambda_{f,f}[0]$ and $\lambda_{f,f}[1]$ are the eigenvalues of the symmetric matrix $\mathbf{B}^{1k}_{f,f} \triangleq (\mathbf{A}^{0k}_{f,f})^\dagger\mathbf{C}^{1k}_{f,f}\mathbf{A}^{0k}_{f,f}$ ($\mathbf{C}^{tk}_{f,f}$ is defined in (41)), and

$$[z_{k,j}[0], z_{k,j}[1]]^\dagger \triangleq (\mathbf{V}^{1k}_{f,f})^\dagger(\mathbf{A}^{0k}_{f,f})^\dagger\breve{\mathbf{h}}_{k,j}, \tag{50}$$

where $\mathbf{V}^{1k}_{f,f}$ is the modal matrix of $\mathbf{B}^{1k}_{f,f}$, i.e., $\left(\mathbf{V}^{1k}_{f,f}\right)^\dagger\mathbf{B}^{1k}_{f,f}\mathbf{V}^{1k}_{f,f} = \mathbf{\Lambda}^{1k}_{f,f}$ (matrix of eigenvalues for $\mathbf{B}^{1k}_{f,f}$), $\mathbf{A}^{0k}_{f,f} \triangleq \mathbf{V}^{0k}_{f,f}\left(\mathbf{\Lambda}^{0k}_{f,f}\right)^{\frac{-1}{2}}$, $\left(\mathbf{V}^{0k}_{f,f}\right)^\dagger\mathbf{C}^{0k}_{f,f}\mathbf{V}^{0k}_{f,f} = \mathbf{\Lambda}^{0k}_{f,f}$, and $f = 1$ for $\varrho_k < 0$, and $f = 0$ for $\varrho_k \geqslant 0$. The threshold value for the kth IoT device, i.e., θ_k, is set for a desirable false alarm rate $P_k^{(\mathrm{f})} \triangleq \mathbb{P}\{d_k = H_{1k}|H_{0k}\}$ as follows

$$P_k^{(\mathrm{f})} = \frac{1}{2\pi}\int_{\theta_k}^{+\infty}\int_{-\infty}^{+\infty}\prod_{n=0}^{1}\frac{\exp(-j\omega x)}{\sqrt{1-2j\chi_{f,f}[n]\omega}}\mathrm{d}\omega\mathrm{d}x. \tag{51}$$

By using (48), we can obtain the correct identification rate for the kth IoT device as follows

$$P_k^{(\mathrm{c})} = \frac{1}{2\pi}\int_{\theta_k}^{+\infty}\int_{-\infty}^{+\infty}\prod_{n=0}^{1}\frac{\exp(-j\omega x)}{\sqrt{1-2j\lambda_{f,f}[n]\omega}}\mathrm{d}\omega\mathrm{d}x. \tag{52}$$

(Proof in Appendix B) □

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TWC.2021.3117256, IEEE Transactions on Wireless Communications

Algorithm 1 Squared ℓ_2-norm SSR IoT DI

Input: $\mathbf{X}$, $\mathbf{R}$, $P_k^{(\mathrm{f})}$, n_k, $k \in \mathcal{X}_\mathrm{u}$
Output: Active IoT set $\hat{\mathcal{X}}_\mathrm{a}$
Initialization: $\hat{\mathcal{X}}_\mathrm{a} = \emptyset$

```
for k = 0, 1, ..., K_u − 1 do
    Obtain θ_k by using (51)
    Obtain h̆_{k,j} by employing (27) and (45)
    Compute φ(h̆_{k,j}) for j = 0, 1, ..., L − 1 using (48)
    Identify the transmission state of the kth IoT device
     by employing (47) and then (53)
    if D_k = H_{1k} then
        X̂_a ← {X̂_a, k}
    end if
end for
```

The decisions corresponding to the L measurements for the kth IoT device, i.e., $d_{k,\ell}$, $\ell = 1, 2, \ldots, L$, can be fused together as

$$D_k = \begin{cases} H_{1k}, & \sum_{\ell=1}^{L} d_{k,\ell} \geqslant n_k \\ H_{0k}, & \sum_{\ell=1}^{L} d_{k,\ell} < n_k, \end{cases} \tag{53}$$

where n_k is an integer value [32]. A formal description of the proposed squared ℓ_2-norm SSR IoT DI algorithm is given in Algorithm 1.

C. IoT DI for Unknown Activity Rate

Here, we develop an IoT DI algorithm for the case of unknown activity rate P_a at the BS. The convex relaxation through $\ell_1 - \ell_2$ mixed-norm is employed for signal reconstruction, which can directly identify the active IoT devices through the non-zero elements of the reconstructed signal. This is attributed to the $\ell_1 - \ell_2$ mixed-norm ability to provide sparse estimates. Note that the proposed $\ell_1 - \ell_2$ mixed-norm SSSR IoT DI algorithm does not require knowledge of transmit power by IoT devices.

1) BIC $\ell_1 - \ell_2$ Mixed-Norm SSSR IoT DI Algorithm: Let us consider $P_\mathrm{a} \in [0,\ P_\mathrm{max}]$, where P_a and the maximum activity rate P_max are unknown at the BS. Since in-phase and quadrature components tend to be either zero or non-zero simultaneously, this provides additional grouping in SSSR. By stacking the in-phase and quadrature components ($\mathbf{X}$ is a real-valued matrix), we can write (19) as

$$\mathbf{Y} = \mathbf{X}\mathbf{U} + \mathbf{V}, \tag{54}$$

where

$$\mathbf{Y} \triangleq \left[\mathrm{Re}\{\mathbf{R}\}\ \ \mathrm{Im}\{\mathbf{R}\}\right], \tag{55a}$$
$$\mathbf{U} \triangleq \left[\mathrm{Re}\{\mathbf{H}\}\ \ \mathrm{Im}\{\mathbf{H}\}\right], \tag{55b}$$
$$\mathbf{V} \triangleq \left[\mathrm{Re}\{\mathbf{W}\}\ \ \mathrm{Im}\{\mathbf{W}\}\right]. \tag{55c}$$

For block-sparse matrix $\mathbf{U}$ in (54), the $\ell_1 - \ell_2$ mixed-norm SSSR is given as follows

$$\hat{\mathbf{U}} = \arg\min_{\mathbf{U}}\ \frac{1}{2}\left\|\mathbf{Y} - \mathbf{X}\mathbf{U}\right\|_\mathrm{F}^2 + N_\mathrm{d}\lambda_\mathrm{g}\sum_{k=0}^{K_\mathrm{u}-1}\left\|\mathbf{u}_{\mathrm{G}_k}\right\|_2, \tag{56}$$

where $N_\mathrm{d} \triangleq 2LN_\mathrm{c}$

$$\mathbf{u}_{\mathrm{G}_k} \triangleq \left[\mathbf{U}_{2k,\cdot}\ \ \mathbf{U}_{2k+1,\cdot}\right], \tag{57}$$

and $\mathbf{U}_{k,\cdot}$ is the kth row of $\mathbf{U}$. In (56), λ_g represents the tuning parameter which is unknown and time-varying. The degrees of sparsity depends on λ_g; the larger λ_g is, the sparser the estimate is. For unknown λ_g, model order selection methods can be employed to identify active IoT devices. By extending the BIC model order selection method in [33] to multiple measurement vectors, the reconstructed matrix $\hat{\mathbf{U}}$ is given by

$$\hat{\mathbf{U}} = \hat{\mathbf{U}}^{(\hat{\lambda})}, \tag{58}$$

where

$$\hat{\lambda} = \arg\min_{\lambda \in [\lambda_\mathrm{L}, \lambda_\mathrm{U}]} C_\mathrm{BIC}(\lambda), \tag{59}$$

$$C_\mathrm{BIC}(\lambda) \triangleq \log\left(\frac{1}{N_\mathrm{d}}\left\|\mathbf{Y} - \mathbf{X}\hat{\mathbf{U}}^{(\lambda)}\right\|_\mathrm{F}^2\right) + \log(N_\mathrm{d})\frac{df}{N_\mathrm{d}}, \tag{60}$$

$$\hat{\mathbf{U}}^{(\lambda)} = \arg\min_{\mathbf{U}}\ \frac{1}{2}\left\|\mathbf{Y} - \mathbf{X}\mathbf{U}\right\|_\mathrm{F}^2 + N_\mathrm{d}\lambda\sum_{k=0}^{K_\mathrm{u}-1}\left\|\mathbf{u}_{\mathrm{G}_k}\right\|_2, \tag{61}$$

and df is the degree of freedom which is given as follows

$$df = \sum_{k=0}^{K_\mathrm{u}-1}\mathbb{I}\left\{\left\|\hat{\mathbf{u}}_{\mathrm{G}_k}\right\|_2 > 0\right\} + (2L-1)\sum_{k=0}^{K_\mathrm{u}-1}\frac{\left\|\hat{\mathbf{u}}_{\mathrm{G}_k}\right\|_2}{\left\|\hat{\mathbf{u}}_{\mathrm{G}_k}^\mathrm{LS}\right\|_2}, \tag{62}$$

where $\hat{\mathbf{u}}_{\mathrm{G}_k}^\mathrm{LS}$ is the LS estimate for the kth IoT device signal.

The Karush–Kuhn–Tucker (KKT) optimality conditions of the optimization problem in (61) is given as

$$-\mathbf{\Psi}_k + N_\mathrm{d}\lambda\frac{\mathbf{u}_{\mathrm{G}_k}}{\left\|\mathbf{u}_{\mathrm{G}_k}\right\|_2} = \mathbf{0} \quad \text{if } \mathbf{u}_{\mathrm{G}_k} \neq \mathbf{0}^\dagger \tag{63a}$$
$$\left\|\mathbf{\Psi}_k\right\|_2 \leqslant N_\mathrm{d}\lambda \quad \text{if } \mathbf{u}_{\mathrm{G}_k} = \mathbf{0}^\dagger, \tag{63b}$$

where

$$\begin{aligned}\mathbf{\Psi}_k &\triangleq \nabla_{\mathbf{u}_{\mathrm{G}_k}}\frac{1}{2}\left\|\mathbf{Y} - \mathbf{X}\mathbf{U}\right\|_\mathrm{F}^2 \\ &= \left[\mathbf{X}_{\cdot,2k}^\dagger(\mathbf{Y} - \mathbf{X}\mathbf{U})\ \ \mathbf{X}_{\cdot,2k+1}^\dagger(\mathbf{Y} - \mathbf{X}\mathbf{U})\right],\end{aligned} \tag{64}$$

and $\mathbf{X}_{\cdot,k}$ is the kth column of $\mathbf{X}$. Let us write $\mathbf{\Psi}_k$ as

$$\mathbf{\Psi}_k = \boldsymbol{\varphi}_k - \mathbf{u}_{\mathrm{G}_k}\mathbf{\Lambda}_k \tag{65}$$

where

$$\boldsymbol{\varphi}_k = \left[\mathbf{X}_{\cdot,2k}^\dagger(\mathbf{Y} - \mathbf{X}\mathbf{U}_{-\{2k,\cdot\}})\ \ \mathbf{X}_{\cdot,2k+1}^\dagger(\mathbf{Y} - \mathbf{X}\mathbf{U}_{-\{2k+1,\cdot\}})\right], \tag{66}$$

and

$$\begin{aligned}\mathbf{\Lambda}_k \triangleq \mathrm{diag}\{&\mathbf{X}_{\cdot,2k}^\dagger\mathbf{X}_{\cdot,2k}, \ldots, \mathbf{X}_{\cdot,2k}^\dagger\mathbf{X}_{\cdot,2k}, \\ &\mathbf{X}_{\cdot,2k+1}^\dagger\mathbf{X}_{\cdot,2k+1}, \ldots, \mathbf{X}_{\cdot,2k+1}^\dagger\mathbf{X}_{\cdot,2k+1}\}\end{aligned} \tag{67}$$

with $\mathbf{U}_{-\{i,\cdot\}}$ as the matrix $\mathbf{U}$ with the ith row being set to $\mathbf{0}^\dagger$. The dimension of the diagonal matrix $\mathbf{\Lambda}_k$ is $4L \times 4L$.

From (63a) and (65), we have $\mathbf{u}_{\mathrm{G}_k}\left(N_\mathrm{d}\lambda\mathbf{I}/\left\|\mathbf{u}_{\mathrm{G}_k}\right\|_2 + \mathbf{\Lambda}_k\right) = \boldsymbol{\varphi}_k$ when the kth IoT device is active. In contrast, when the kth IoT device is inactive, $\mathbf{\Psi}_k = \boldsymbol{\varphi}_k$. Hence, we can write

$$\mathbf{u}_{\mathrm{G}_k} = \mathbb{I}\left\{\|\boldsymbol{\varphi}_k\|_2 > N_\mathrm{d}\lambda\right\}\boldsymbol{\varphi}_k\left(\frac{N_\mathrm{d}\lambda}{\left\|\mathbf{u}_{\mathrm{G}_k}\right\|_2}\mathbf{I} + \mathbf{\Lambda}_k\right)^{-1}. \tag{68}$$

© 2021 IEEE. Personal use is permitted, but republication/redistribution requires IEEE permission. See https://www.ieee.org/publications/rights/index.html for more information.

Algorithm 2 BIC $\ell_1-\ell_2$ mixed-norm SSSR IoT DI algorithm

Input: $\mathbf{Y}$, $\mathbf{X}$, $\mathbf{\Lambda}_k$, $k=0,1,\ldots,K_{\rm u}-1$, $\lambda\in[\lambda_{\rm L},\lambda_{\rm U}]$, $M_{\rm C}$, and $M_{\rm G}$
Output: Active IoT set $\hat{\mathcal{X}}_{\rm a}$
Initialization: $\hat{\mathcal{X}}_{\rm a}=\emptyset$, $i=1$, $Golden=1$, $\lambda_1=\lambda_{\rm L}$
while $Golden=1$ **do**
  $\hat{\mathbf{U}}^{[0]}=\mathbf{0}$, $t=1$, $SSSR=1$
  while $SSSR=1$ **do**
    for $k=0,1,\ldots,K_{\rm u}-1$ **do**
      Obtain $\boldsymbol{\varphi}_k^{[t]}$ by employing (70)
      if $\|\boldsymbol{\varphi}_k^{[t]}\|_2\leqslant N_{\rm c}\lambda_i$ **then**
        $\hat{\mathbf{u}}_{{\rm G}_k}^{[t]}=\mathbf{0}$
      else
        Update $\mathbf{u}_{{\rm G}_k}^{[t]}$ as in (69)
      end if
    end for
    if $\left(\|\hat{\mathbf{U}}^{[t]}-\hat{\mathbf{U}}^{[t-1]}\|\geqslant\epsilon_{\rm c}\right)\cap(t<M_{\rm C})$ **then**
      $t\leftarrow t+1$
    else
      $\hat{\mathbf{U}}^{(\lambda_i)}=\hat{\mathbf{U}}^{[t]}$, $SSSR=0$
    end if
  end while
  $i\leftarrow i+1$
  if $i=2$ **then**
    $\lambda_i=\lambda_{\rm U}$, $Golden=1$
  else if $\left(|\lambda_{i-1}-\lambda_{i-2}|\geqslant\epsilon_{\rm g}\right)\cap(i<M_{\rm G}+1)$ **then**
    Find λ_i using the Golden selection search
    $Golden=1$
  else
    $\hat{\lambda}=\underset{\lambda\in\{\lambda_{i-1},\lambda_{i-2}\}}{\arg\min}\ C_{\rm BIC}(\lambda)$,
    $\hat{\mathbf{U}}=\hat{\mathbf{U}}^{(\hat{\lambda})}$, $Golden=0$,
  end if
end while
$\hat{\mathcal{X}}_{\rm a}=\left\{\forall k\in\{0,\ldots K_{\rm u}-1\}\middle|\|\hat{\mathbf{u}}_{{\rm G}_k}\|_2\neq 0\right\}$.

To solve the optimization (61), we can use block-coordinate descent algorithm, where consists of solving each $\mathbf{u}_{{\rm G}_k}$ in (57) at a time. By starting from a sparse solution like, $\hat{\mathbf{U}}=\mathbf{0}$, at each iteration, we check for a given k whether $\mathbf{u}_{{\rm G}_k}$ is optimal or not based on the conditions in (63). If $\|\boldsymbol{\varphi}_k\|_2\leqslant N_{\rm d}\lambda$, $\hat{\mathbf{u}}_{{\rm G}_k}=\mathbf{0}$; otherwise, $\mathbf{u}_{{\rm G}_k}$ at the tth iteration is iteratively updated as

$$\hat{\mathbf{u}}_{{\rm G}_k}^{[t]}=\mathbb{I}\Big\{\|\boldsymbol{\varphi}_k^{[t]}\|_2>N_{\rm d}\lambda\Big\}\boldsymbol{\varphi}_k^{[t]}\left(\frac{N_{\rm d}\lambda}{\left\|\hat{\mathbf{u}}_{{\rm G}_k}^{[t-1]}\right\|_2}\mathbf{I}+\mathbf{\Lambda}_k\right)^{-1},\quad(69)$$

where

$$\boldsymbol{\varphi}_k^{[t]}=\left[\mathbf{X}_{:,2k}^{\dagger}(\mathbf{Y}-\mathbf{X}\mathbf{U}_{-\{2k,:\}}^{[t-1]})\quad\mathbf{X}_{:,2k+1}^{\dagger}(\mathbf{Y}-\mathbf{X}\mathbf{U}_{-\{2k+1,:\}}^{[t-1]})\right].\quad(70)$$

This procedure continues until the absolute difference of successive iterations becomes smaller than the tolerance value $\epsilon_{\rm c}$.

2) Efficient One-dimensional Search: Efficient one dimensional iterative search algorithms can be used to solve the optimization problem in (59). In an iterative search method, the interval $[\lambda_{\rm L},\lambda_{\rm U}]$ is repeatedly reduced on the basis of function evaluations until a reduced bracket $[\lambda_{\rm L},\lambda_{\rm U}]$ is achieved which is sufficiently small. These methods can be applied to any function and differentiability of the function is not essential. An iterative search method in which iterations can be performed until the desired accuracy in either the minimizer or the minimum value of the objective function is achieved is the golden-section search method [34].

Convergence of the Optimization Problem in (61): It has been shown that for an optimization problem whose objective function is the sum of a smooth and convex function and a non-smooth but block-separable convex function, block-coordinate descent optimization converges towards the global minimum of the problem [35]. In (61), $\|\mathbf{Y}-\mathbf{X}\mathbf{U}\|_{\rm F}^2$ is a smooth and differentiable convex function and $\sum_{k=0}^{K_{\rm u}-1}\|\mathbf{u}_{{\rm G}_k}\|_2$ is a separable penalty function, where $\|\mathbf{u}_{{\rm G}_k}\|_2$ is a continuous and convex function with respect to $\mathbf{u}_{{\rm G}_k}$. Thus, block-coordinate descent converges to the global minimum.

A formal description of the $\ell_1-\ell_2$ mixed-norm SSSR IoT identification algorithm is summarized in Algorithm 2. In Algorithm 2, $M_{\rm G}$ and $M_{\rm C}$, denote the maximum number of iterations for the Golden selection search and the block-coordinate descent optimization, respectively.

IV. Data Detection

The next step after IoT DI is to detect the data of devices identified as active. Since CSI is unknown, the existing MUD algorithms, such as SIC cannot be employed. In this section, we propose a new nonlinear MUD algorithm which does not require CSI for data detection.

A. 2-MC-MUD Algorithm

The output of the IoT DI algorithm is a set of IoT devices $\hat{\mathcal{X}}_{\rm a}$. Since the delay of the IoT devices are known, we can apply sequence matched filtering to the small set of active IoT devices. Without loss of generality, we assume that $\hat{\mathcal{X}}_{\rm a}\triangleq\{k_0,k_1,\ldots,k_{\hat{K}_{\rm a}-1}\}$ and $\tau_{k_0}\leqslant\tau_{k_1}\leqslant\cdots\leqslant\tau_{k_{\hat{K}_{\rm a}-1}}$, where $\hat{K}_{\rm a}\triangleq\mathbf{card}(\hat{\mathcal{X}}_{\rm a})$.

We consider a bank of $\hat{K}_{\rm a}$ single-user MFs for the identified active IoT devices in $\hat{\mathcal{X}}_{\rm a}$. The output of the MF after synchronized sampling and normalization by $N_{\rm c}$ for the k_nth IoT device is expressed as [36]

$$\begin{aligned}y_{k_n,i}&\triangleq\frac{1}{N_{\rm c}}\int_{\tau_{k_n}+iT_{\rm s}}^{\tau_{k_n}+(i+1)T_{\rm s}}r(t)s_{k_n}(t-iT_{\rm s}-\tau_{k_n})dt\qquad(71)\\&=g_{k_n}b_{k_n,i}+\sum_{k_j<k_n}g_{k_j}b_{k_j,i+1}\rho_{k_nk_j}+\sum_{k_j<k_n}g_{k_j}b_{k_j,i}\rho_{k_jk_n}\\&\quad+\sum_{k_j>k_n}g_{k_j}b_{k_j,i}\rho_{k_nk_j}+\sum_{k_j>k_n}g_{k_j}b_{k_j,i-1}\rho_{k_jk_n}+w_{k_n,i},\end{aligned}$$

where $w_{k_n,i}\triangleq\sigma_{\rm w}\int_{\tau_{k_n}+iT_{\rm s}}^{\tau_{k_n}+(i+1)T_{\rm s}}w(t)s_{k_n}(t-iT_{\rm s}-\tau_{k_n})\mathrm{d}t$, $\rho_{k_nk_j}\triangleq\frac{1}{N_{\rm c}}\int_{\tau_{k_j}}^{T_{\rm s}}s_{k_n}(t)s_{k_j}(t-\tau_{k_j})\mathrm{d}t$, and $\rho_{k_jk_n}\triangleq\frac{1}{N_{\rm c}}\int_0^{\tau_{k_j}}s_{k_n}(t)s_{k_j}(t+T_{\rm s}-\tau_{k_j})\mathrm{d}t$.

The output of the single-user MF in (71) for the k_nth IoT device can be written as

$$y_{k_n,i}=g_{k_n}b_{k_n,i}+v_{k_n,i},\qquad i=0,1,\ldots,N_{\rm s}-1,\qquad(72)$$

where $v_{k_n,i}$ represents the effect of noise and multiuser interference on the k_nth IoT device, and $b_{k_n,i}\in\{-1,+1\}$.

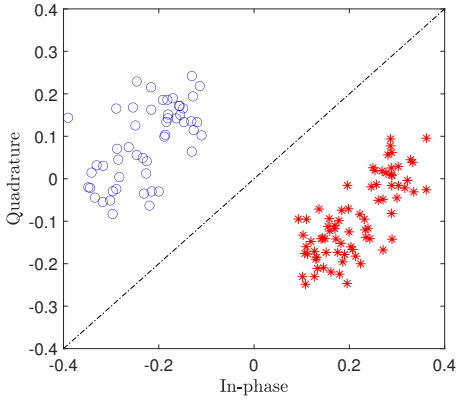


Fig. 7: Separated symbols by the proposed 2-MC-MUD algorithm for an active IoT device at 10 dB SNR, $K_{\rm u} = 768$, and $N_{\rm c} = 512$.

For data detection without any sign ambiguity, the phase of g_{k_n}, $k_n \in \hat{\mathcal{X}}_{\rm a}$, is leastwise required to be known at the BS. However, by employing differential coding at IoT devices, a MUD algorithm can be developed which removes the need for such *a priori* knowledge. Differential coding is a coding technique used for non-coherent data detection. Instead of encoding a bit sequence directly, it encodes the difference between the bit sequence as [30]

$$b^{\rm c}_{k_n,i} = b^{\rm c}_{k_n,i-1} \oplus b^{\rm d}_{k_n,i} \quad , \quad k_n \in \mathcal{X}_{\rm a}, \tag{73}$$

where $\oplus$ is the modulo-2 addition and $b^{\rm c}_{k_n,i} \in \{0,1\}$ is the ith bit at the output of the channel encoder of the k_nth IoT device as shown in Fig. 2. The BPSK modulated data for the k_nth IoT device in (72), i.e., $b_{k_n,i}$ can be mathematically expressed based on the differentially coded bit $b^{\rm d}_{k_n,i} \in \{0,1\}$ as $b_{k_n,i} = (-1)^{b^{\rm d}_{k_n,i}}$.

Since g_{k_n}, $k_n \in \mathcal{X}_{\rm a}$, remains unchanged during the short packet, the received symbols of the active IoT device k_n in (72) form two clusters corresponding to the transmitted bits 1 and 0. The main idea behind the proposed MUD algorithm is to extract these two clusters regardless of which cluster is labeled 1 or 0. By extracting the two clusters, the data stream of the active IoT device k_n can be detected without any prior knowledge about the CSI and CP due to differential coding.

To extract these two clusters for each active IoT device, the 2-MC algorithm can be employed. By applying the 2-MC algorithm to $y_{k_n,i}$, $i = 0,1,\ldots,N_{\rm s}-1$, in (72), the two clusters are separated based on the nearest mean criterion disregard to the label. The 2-MC minimizes the within-cluster sum of squares (WCSS), i.e, the sum of the squared Euclidean distance [37]. Let us define $\mathcal{U} \triangleq \{0,1,\ldots,N_{\rm s}-1\}$. The 2-MC algorithm partitions $\mathcal{U}$ into two sets $\mathcal{U}_{k_n,0}$ and $\mathcal{U}_{k_n,1}$ by minimizing the WCSS as follows

$$\begin{aligned} \arg\min_{\mathcal{U}} \; & \sum_{i\in\mathcal{U}_{k_n,0}} \left|y_{k_n,i} - \mu_{k_n,0}\right|^2 + \sum_{i\in\mathcal{U}_{k_n,1}} \left|y_{k_n,i} - \mu_{k_n,1}\right|^2, \\ \text{subject to} \quad & \mu_{k_n,0} = \frac{1}{\mathbf{card}(\mathcal{U}_{k_n,0})} \sum_{i\in\mathcal{U}_{k_n,0}} y_{k_n,i}, \\ & \mu_{k_n,1} = \frac{1}{\mathbf{card}(\mathcal{U}_{k_n,1})} \sum_{i\in\mathcal{U}_{k_n,1}} y_{k_n,i}. \end{aligned} \tag{74}$$

The minimization problem in (74) can be solved by different methods. One of the most common algorithm is the Lloyd's algorithm which uses an iterative refinement technique [38]. Given initial mean values[3] $\mu^{[0]}_{k_n,0}$ and $\mu^{[1]}_{k_n,1}$ for $\mu_{k_n,0}$ and $\mu_{k_n,1}$ in (74), the Lloyd's algorithm proceeds by alternating between the assignment and updating steps as follows:

Assignment Step: The element of $\mathcal{U}$ at iteration t, i.e., $\mathcal{U}^{[t]}$ is assigned to $\mathcal{U}^{[t]}_{k_n,0}$ when

$$\mathcal{U}^{[t]}_{k_n,0} = \left\{ i : \left|y_{k_n,i} - \mu^{[t]}_{k_n,0}\right|^2 \leqslant \left|y_{k_n,i} - \mu^{[t]}_{k_n,1}\right|^2 \right\}. \tag{75}$$

Otherwise, it is assigned to $\mathcal{U}^{[t]}_{k_n,1}$.

Updating Step: The mean of the the clusters $\mathcal{U}^{[t]}_{k_n,0}$ and $\mathcal{U}^{[t]}_{k_n,1}$ are updated as

$$\mu^{[t+1]}_{k_n,1} = \frac{1}{\mathbf{card}(\mathcal{U}^{[t]}_{k_n,1})} \sum_{i\in\mathcal{U}^{[t]}_{k_n,1}} y_{k_n,i}, \tag{76a}$$

$$\mu^{[t+1]}_{k_n,0} = \frac{1}{\mathbf{card}(\mathcal{U}^{[t]}_{k_n,0})} \sum_{i\in\mathcal{U}^{[t]}_{k_n,0}} y_{k_n,i}. \tag{76b}$$

The 2-MC algorithm converges when the assignment step does not change. Fig. 7 shows the output of the 2-MC algorithm for an active IoT device. As seen, the sequence at the output of the MF is portioned into two clusters regardless of the label.

After partitioning $\mathcal{U}$ into two clusters $\mathcal{U}_{k_n,0}$ and $\mathcal{U}_{k_n,1}$, $y_{k_n,i}$, $i = 0,1,\ldots,N_{\rm s}-1$, is mapped into a binary sequence $\mathbf{b}^{\rm m}_{k_n} \triangleq [b^{\rm m}_{k_n,0}\ b^{\rm m}_{k_n,1}\ \ldots\ b^{\rm m}_{k_n,N_{\rm s}-1}]^\dagger$ with elements as

$$b^{\rm m}_{k_n,i} = \mathbb{I}\{i \in \mathcal{U}_{k_n,1}\}. \tag{77}$$

Then, by applying differential decoding to the mapped binary sequence $\mathbf{b}^{\rm m}_{k_n}$, the channel coded data stream for the active IoT device k_n is obtained as follows

$$\hat{b}^{\rm c}_{k_n,i} = b^{\rm m}_{k_n,i} \oplus b^{\rm m}_{k_n,i-1}. \tag{78}$$

Finally, $\hat{\mathbf{b}}^{\rm c}_{k_n} \triangleq [\hat{b}^{\rm c}_{k_n,0}\ \hat{b}^{\rm c}_{k_n,1}\ \ldots\ \hat{b}^{\rm c}_{k_n,N_{\rm s}-2}]^\dagger$ is decoded by the channel decoder, and the data stream of the active IoT device k_n is obtained. The proposed 2-MC-MUD algorithm is summarized in Algorithm 3.

Algorithm 3 : 2-MC-MUD algorithm

Input: $r(t)$, $\hat{\mathcal{X}}_{\rm a}$, $\hat{K}_{\rm a} = \mathbf{card}(\hat{\mathcal{X}}_{\rm a})$
Output: $\hat{\mathbf{b}}_{k_n}$, $k_n \in \hat{\mathcal{X}}_{\rm a}$
for $n = 0,1,\ldots,\hat{K}_{\rm a}-1$ **do**
  Set initial value for $\mathcal{U}^{[0]}_{k_n,1}$ and $\mathcal{U}^{[0]}_{k_n,0}$
  Obtain $y_{k_n,i}$, $i = 0,1,\ldots,N_{\rm s}-1$, by employing (71)
  while $\mathcal{U}^{[t+1]}_{k_n,1} \neq \mathcal{U}^{[t]}_{k_n,1}$ **do**
    obtain $\mathcal{U}^{[t]}_{k_n,1}$ and $\mathcal{U}^{[t]}_{k_n,0}$ by employing (75)
    $\mu^{[t+1]}_{k_n,1} \leftarrow \mathcal{U}^{[t]}_{k_n,1}$ by employing (76a)
    $\mu^{[t+1]}_{k_n,0} \leftarrow \mathcal{U}^{[t]}_{k_n,0}$ by employing (76b)
  end while
  Obtain the binary mapped sequence $\mathbf{b}^{\rm m}_{k_n}$ through (77)
  Apply differential decoding in (78) to $\mathbf{b}^{\rm m}_{k_n}$ to obtain $\hat{\mathbf{b}}^{\rm c}_{k_n}$
  Apply channel decoding to $\hat{\mathbf{b}}^{\rm c}_{k_n}$ to obtain $\hat{\mathbf{d}}_{k_n}$
end for

[3]The Forgy method is used for intialization, where two observations from the dataset are used as the initial means [37].

This article has been accepted for publication in IEEE Transactions on Wireless Communications. This is the author's version which has not been fully edited and content may change prior to final publication. Citation information: DOI 10.1109/TWC.2021.3117256

B. Complexity Analysis

The complexity of the proposed squared ℓ_2-norm SSR IoT DI algorithm is $\mathcal{O}(K_{\mathrm{u}}LN_{\mathrm{c}}^2+LN_{\mathrm{c}}^3)$. The complexity of the proposed BIC $\ell_1-\ell_2$ mixed-norm SSSR IoT DI algorithm per each iteration is $\mathcal{O}(N_{\mathrm{c}}LK_{\mathrm{u}}^2)$, where the maximum number of iterations is $M_{\mathrm{G}}M_{\mathrm{C}}$ (M_{G} and M_{C} are the maximum number of iterations for the Golden selection search and the block-coordinate descent optimization, respectively. The complexity of the proposed 2-MC-MUD for single user matched filtering and clustering is $\mathcal{O}(N_{\mathrm{c}}^2N_{\mathrm{s}}k_{\mathrm{a}})$ and $\mathcal{O}(K_{\mathrm{u}}N_{\mathrm{s}}k_{\mathrm{a}})$, respectively, where $k_{\mathrm{a}}=\mathbf{card}(\mathsf{X}_{\mathrm{a}})$.

V. Multiple Receive Antennas

The performance of the proposed IoT DI algorithms drastically improve by employing multiple receive antennas at the BS due to spatial diversity. Let us consider that the BS is equipped with N_{r} receive antennas and fuses all 1-bit activity decisions made by each receive antennas according to following logic rule

$$Z_k=\begin{cases}H_{1k}, & \sum_{i=1}^{N_{\mathrm{r}}}D_{k,i}\geqslant m_k\\ H_{0k}, & \sum_{i=1}^{N_{\mathrm{r}}}D_{k,i}<m_k,\end{cases} \quad (79)$$

where $D_{k,i}\in\{0,1\}$ is the decision made by the ith receive antenna on the transmission state of the kth IoT device, and m_k is an arbitrary integer for the kth IoT device. For the suboptimal detector in (79) and $m_k=m, k=0,1,\ldots,K_{\mathrm{u}}-1$, the correct identification and false alarm rates are given by

$$Q_{\mathrm{C}}=\sum_{l=m}^{N_{\mathrm{r}}}\binom{N_{\mathrm{r}}}{l}P_{\mathrm{C}}^{l}\big(1-P_{\mathrm{C}}^{l}\big)^{N_{\mathrm{r}}-l}, \quad (80)$$

and

$$Q_{\mathrm{F}}=\sum_{l=m}^{N_{\mathrm{r}}}\binom{N_{\mathrm{r}}}{l}P_{\mathrm{F}}^{l}\big(1-P_{\mathrm{F}}^{l}\big)^{N_{\mathrm{r}}-l}, \quad (81)$$

where P_{C} and P_{F} are the correct identification and false alarm rates for a single receive antennas. Similarly, 1-bit decision fusion can be used for the 2-MC-MUD algorithm to improve data detection performance.

VI. Simulation Results

In this section, we examine the performance of the proposed IoT DI algorithms and the 2-MC-MUD algorithm through several simulation experiments.

A. Simulation Setup

Unless otherwise mentioned, we considered an IoT network with $K_{\mathrm{u}}=1024$ IoT devices. It is assumed that the spreading sequences of the IoT devices are random binary codes with spreading factor $N_{\mathrm{c}}=512$. Each IoT packet is 128 bits with payload length of 40 bits. The delay of the IoT devices was generated as uniform distributions $\alpha_k\sim\mathcal{U}_{\mathrm{d}}\big[0,5\big]$, $\beta_k\sim\mathcal{U}_{\mathrm{d}}\big[0,511\big]$, and $\xi_k\sim\mathcal{U}_{\mathrm{c}}\big[0,1\big)$. The effect of the unknown CSI and CP for each IoT device was modeled as independent complex Gaussian random variables with mean $\mu_k=\sqrt{0.1}+j\sqrt{0.1}$ and variance $\sigma_k^2=1$, $k\in\mathcal{X}_{\mathrm{u}}$, i.e., Rician fading with K-factor 0.2 was considered. The average system SNR was defined as $\vartheta\triangleq\bar{P}_{\mathrm{a}}\sum(|\mu_k|^2+\sigma_k^2)p_k\eta_k/\sigma_{\mathrm{w}}^2$, where $p_k=\varsigma/\eta_k$ (ς changes according to ϑ), $\bar{P}_{\mathrm{a}}=P_{\mathrm{a}}$ for Algorithm 1, $\bar{P}_{\mathrm{a}}=P_{\max}/2$ for Algorithm 2 (for the case of time-varying P_{a}), and $\sigma_{\mathrm{w}}^2=1$ is the variance of the additive noise. The range of tuning parameter for the BIC minimization in (59) was set as $\lambda=[0\ 500]$, $\epsilon_{\mathrm{g}}=2$, and $M_{\mathrm{G}}=50$. The performance of the proposed IoT DI algorithms were evaluated in terms of system correct identification P_{C} and system false alarm P_{F} rates for 10^6 Monte Carlo trials. Also, the performance of the proposed 2-MC-MUD algorithm was evaluated in terms of average PER in the presence IoT DI error.

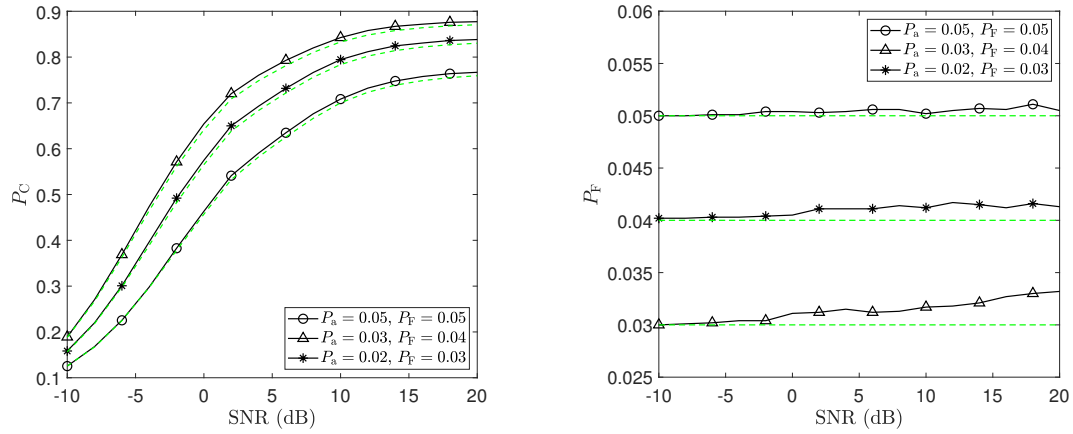
B. Simulation Results

Fig. 8 depicts P_{C} and P_{F} of the proposed squared ℓ_2-norm SSR IoT DI algorithm (Algorithm 1) versus SNR for different values of P_{a} and $P_{\mathrm{f}}^{(k)}$, $K_{\mathrm{u}}=1024$, and $L=1$. The threshold values θ_k, $k\in\mathcal{X}_{\mathrm{u}}$, are set by using (51). As seen, Algorithm 1 can offer high correct identification error rate even for a single observation vector. Also, there is an insignificant gap between P_{C} obtained in the simulation experiment and the theoretical result in (52). Similarly, P_{F} matches the preset false alarm rate, i.e., $P_{\mathrm{f}}^{(k)}\in\{0.03,0.04,0.05\}$, $k\in\mathcal{X}_{\mathrm{u}}$. We notice from Fig. 8 that the theoretical results more accurately match the simulation results at higher P_{a} and at lower SNRs since the CLT is more reliable.

In Fig. 9, we illustrate the performance of Algorithm 1 for $K_{\mathrm{u}}\in\{1024,4096\}$, $P_{\mathrm{a}}\in\{0.02,0.05,0.1\}$, and $P_{\mathrm{f}}^{(k)}\in\{0.05,0.07,0.1\}$ when multiple receive antennas are employed at the BS. Here, we consider the majority rule hard decision combining as a suboptimal detector for $N_{\mathrm{r}}=64$ receive antennas. As observed, miss identification ($P_{\mathrm{M}}=1-P_{\mathrm{C}}$) and false alarm rates substantially decrease by employing multiple receive antennas at the BS. Here, for 10^5 Monte Carlo trials, we obtain $P_{\mathrm{F}}=0$; this value is 4.43×10^{-26} for the theoretical result given in (81).

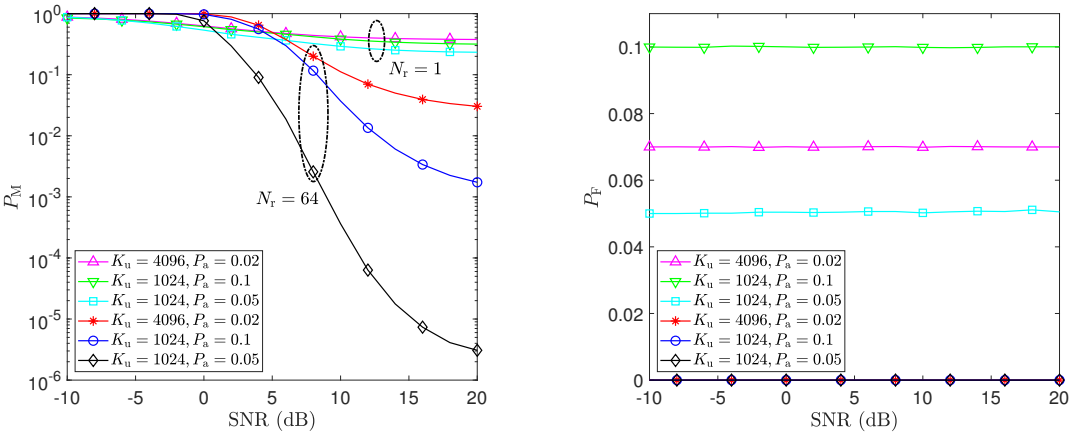
In Fig. 10, we compare the performance of Algorithm 1 and Algorithm 2 with the ADMM algorithm using GCV (ADMM-GCV) in [39], the AMP-based non-coherent activity detection algorithm in [40] and the GOMP method in [14] for $P_{\mathrm{a}}=0.02$, $K_{\mathrm{u}}=1024$, $L=21$, and $N_{\mathrm{r}}=1$. For Algorithm 1, the threshold values θ_k, $k\in\mathcal{X}_{\mathrm{u}}$, are set for $P_k^{(\mathrm{f})}=0.05$ in (51), and $n_k=5$ are used for hard decision combining in (53). As seen, our proposed algorithms offer acceptable correct identification and false alarm rates for a single receive antenna at the BS. Moreover, we observe that GOMP exhibits a larger correct identification rate at the expense of significantly higher false alarm rate. This is different for the algorithm in [40], which offers a lower false alarm rate compared with Algorithm 2 in the range of $[10,20]$ dB SNR at the cost of reduced correct identification rate. We also notice that the performance improvement of Algorithm 1 for $L=21$ is not very high compared to $L=1$ in Fig. 8 since $p\big(\hat{h}_{k,j_1,m},\hat{h}_{k,j_2,m}\ |H_{tk}\big)\neq p\big(\hat{h}_{k,j_1,m}|H_{tk}\big)p\big(\hat{h}_{k,j_2,m}|H_{tk}\big)$, $j_1\neq j_2$, $t\in\{0,1\}$. Moreover, as seen, the false alarm rate of Algorithm 1 increases slowly as the SNR increases. There are two reasons for this behaviour: 1) the validity of joint Gaussian PDF assumption decreases, and 2) the effectiveness of the hard decision combining decreases due to the high correlation among the reconstructed vectors.

© 2021 IEEE. Personal use is permitted, but republication/redistribution requires IEEE permission. See https://www.ieee.org/publications/rights/index.html for more information.



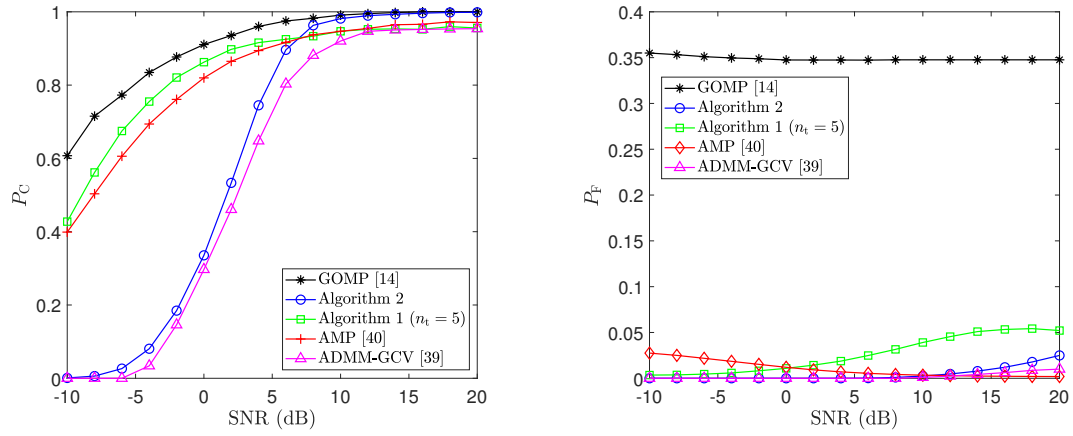
(a) $P_{\rm C}$ versus SNR **(b)** $P_{\rm F}$ versus SNR

Fig. 8: The system correct identification rate, $P_{\rm C}$, and false alarm rate, $P_{\rm F}$, of the proposed squared ℓ_2-norm SSR IoT DI algorithm (Algorithm 1) versus SNR for different values of $P_{\rm a}$, $P_{\rm f}^{(k)} \in \{0.03, 0.04, 0.05\}$, $K_{\rm u} = 1024$, and $L = 1$.



(a) $P_{\rm M}$ versus SNR. **(b)** $P_{\rm F}$ versus SNR.

Fig. 9: Performance of the proposed ℓ_2-norm SSR IoT DI algorithm for $P_{\rm f}^{(k)} \in \{0.05, 0.07, 0.1\}$, $K_{\rm u} \in \{1024, 4096\}$, $L = 1$, and $N_{\rm r} = \{1, 64\}$ with the majority hard decision combining.

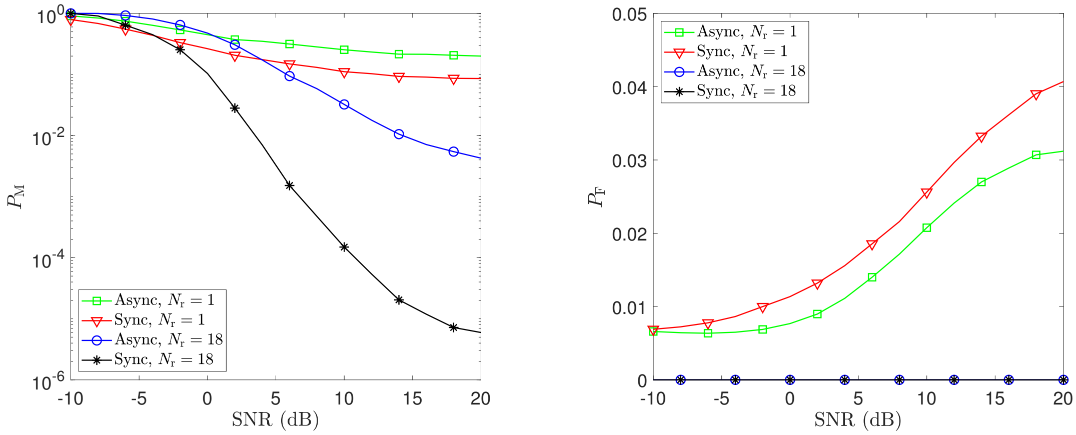


(a) $P_{\rm C}$ versus SNR **(b)** $P_{\rm F}$ versus SNR

Fig. 10: Performance comparison of the proposed IoT DI algorithms with the ADMM-GCV in [39], the AMP-based non-coherent activity detection algorithm in [40] and GOMP method in [14] for $P_{\rm a} = 0.02$, $K_{\rm u} = 1024$, $N_{\rm r} = 1$, $L = 21$, and $P_k^{(\rm f)} = 0.05$ (Algorithm 1).

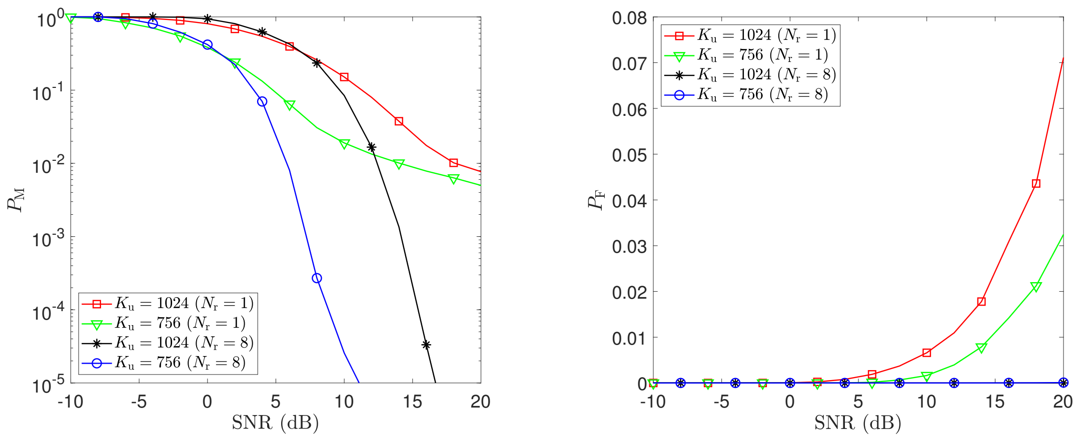
In Fig. 11, we show the effect of uncertainty in delay $\tau_k \triangleq \alpha_k T_{\rm s} + \beta_k T_{\rm c} + \xi_k$, $k \in \mathcal{X}_{\rm u}$, on the performance of Algorithm 1 for $P_{\rm a} = 0.05$, $K_{\rm u} = 1024$, $L = 21$, $n_{\rm t} = 5$, and $P_{\rm f}^{(k)} = 0.06$. We consider that packets arrive at the BS with chip delay uncertainty $\xi_k / T_{\rm c} \in \mathcal{U}_{\rm c}[0, 1)$ while the identification and detection are performed by employing the estimated delay based on distance, which is a priori known and fixed at the BS. For multiple receive antennas, we consider the majority rule hard decision combining as a suboptimal detector for $N_{\rm r} = 18$ receive antennas. As observed, the performance of Algorithm

This article has been accepted for publication in IEEE Transactions on Wireless Communications. This is the author's version which has not been fully edited and content may change prior to final publication. Citation information: DOI 10.1109/TWC.2021.3117256



(a) P_M versus SNR **(b)** P_F versus SNR

Fig. 11: Effect of uncertainty in delay τ_k, $k \in \mathcal{X}_u$, on the performance of the proposed ℓ_2-norm SSR IoT DI algorithm for $P_a = 0.05$, $K_u = 1024$, $N_r = \{1, 18\}$, $L = 21$, and $P_f^{(k)} = 0.06$.



(a) P_M versus SNR **(b)** P_F versus SNR

Fig. 12: The system miss identification rate, P_M, and false alarm rate, P_F, of the proposed BIC $\ell_1 - \ell_2$ mixed-norm SSSR IoT DI algorithm (Algorithm 2) versus SNR for random activity rate $P_a \in [0\ 0.06]$ and $L = 21$ with the majority rule hard decision combining.

1 degrades in the presence of the round-trip delay estimation error. However, by employing multiple receive antennas at the BS, this performance degradation can be reduced.

Fig. 12 illustrates P_C and P_F of the proposed BIC $\ell_1 - \ell_2$ mixed-norm SSSR IoT DI algorithm versus SNR when the activity rate varies uniformly in the range $P_a \in [0,\ 0.06]$. As seen, the proposed algorithm exhibits high correct identification rate for high overloading factors, such as $OF = 2$ (1024 devices), when P_a is unknown and time-varying. Also, the false alarm rate of the proposed algorithm is significantly low for the SNR values lower than 15 dB.

Fig. 13 compares the performance of the developed MA scheme when the proposed 2-MC-MUD and differentially coherent decorrelation (DCD)-MUD [36] algorithms are employed for $N_r = 12$. We also show the performance with estimated CSI using the joint activity and channel estimation method in [20]. As seen, the proposed MUD algorithm outperforms the DCD-MUD [36]. This superiority in performance is related to the capability of the 2-MC algorithm to accurately separate the two clusters of data. We also observe that data detection with estimated CSI can offer lower PER compared with our proposed non-coherent detection when a sufficient number of pilots N_p is used for joint activity detection and channel estimation. This is achieved at the expense of lower spectral efficiency and higher latency.

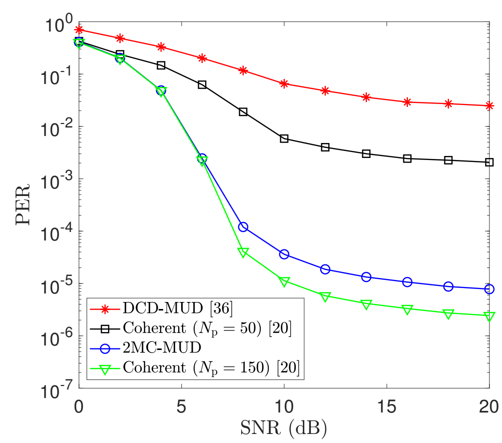


Fig. 13: PER of the proposed MA scheme when 2-MC-MUD and DCD-MUD [36] algorithms are employed for non-coherent data detection at the BS ($K_u = 768$, $OF = 1.5$, and $P_a \in [0,\ 0.06]$).

VII. Conclusion

A new uncoordinated uplink MA for mMTC with short-packet and sporadic traffic was proposed in this paper. The

© 2021 IEEE. Personal use is permitted, but republication/redistribution requires IEEE permission. See https://www.ieee.org/publications/rights/index.html for more information.

proposed MA scheme reduces the control signaling associated with the MAC and PHY layers. Instead of transmitting the device identifier using a portion of bits in a packet, the squared ℓ_2-norm SSR and BIC $\ell_1-\ell_2$ mixed-norm SSSR IoT DI algorithms were proposed to identify active IoT devices through the assigned unique non-orthogonal spreading code to each IoT device. To further reduce the overhead, we removed the preambles and pilots used for channel estimation by developing the non-coherent 2-MC-MUD algorithm based on unsupervised machine learning.

Appendix A

By applying the statistical expectation to (35) and employing $\mathbb{E}\{h_{k,j,1}|H_{tk}\}=0$ and $\mathbb{E}\{h_{k,j,0}|H_{tk}\}=0$, we can write $\mathbb{E}\{\hat{h}_{k,j,f}|H_{tk}\}=0$, $t,f\in\{0,1\}$. To obtain the variance of $\hat{h}_{k,j,f}$ given H_{tk}, i.e., $\Sigma^{tk}_{f,f}$, we use the variance sum law as

$$\mathbb{V}\text{ar}\Big\{\sum_i a_i z_i\Big\}=\sum_i\Big(|a_i|^2\mathbb{V}\text{ar}\{z_i\}+\sum_{j\neq i}a_i a_j^*\mathbb{C}\text{ov}\{z_i,z_j\}\Big). \quad (82)$$

Since $h_{k_1,j,f}$, $h_{k_2,j,\bar{f}}$, and $w'_{k,j,f}$, $k,k_1,k_2\in\mathcal{X}_{\text{u}}$, in (35), are zero-mean and uncorrelated random variables, by applying (82) to (35), we obtain (83) at the top of next page.

By using (17), $\mathbb{V}\text{ar}\{h_{k,j,f}|H_{1k}\}$, $f\in\{0,1\}$, in (83) can be written as

$$\begin{aligned}\mathbb{V}\text{ar}\big\{h_{k,j,f}|H_{tk}\big\}&=\mathbb{V}\text{ar}\big\{g_k b_{k,j-\alpha_k-1+f}|H_{tk}\big\} \quad (84)\\ &=\mathbb{E}\Big\{\mathbb{V}\text{ar}\big\{g_k b_{k,j-\alpha_k-1+f}|g_k,H_{tk}\big\}\Big\}\\ &\quad+\mathbb{V}\text{ar}\Big\{\mathbb{E}\big\{g_k b_{k,j-\alpha_k-1+f}|g_k,H_{tk}\big\}\Big\}.\end{aligned}$$

Since $\mathbb{V}\text{ar}\big\{b_{k,j-\alpha_k-1+f}|H_{tk}\big\}=t$, $t,f\in\{0,1\}$, we can write

$$\begin{aligned}&\mathbb{E}\Big\{\mathbb{V}\text{ar}\big\{g_k b_{k,j-\alpha_k-1+f}|g_k,H_{tk}\big\}\Big\} \quad (85)\\ &=\mathbb{E}\big\{|g_k|^2\big\}\mathbb{V}\text{ar}\big\{b_{k,j-\alpha_k-1+f}|H_{tk}\big\}=t(\sigma_k^2+|\mu_k|^2)\eta_k p_k.\end{aligned}$$

By substituting (85) and $\mathbb{E}\big\{g_k b_{k,j-\alpha_k-1+f}|g_k,H_{tk}\big\}=0$, $f\in\{0,1\}$, into (84), we obtain

$$\mathbb{V}\text{ar}\big\{h_{k,j,f}|H_{tk}\big\}=t(\sigma_k^2+|\mu_k|^2)\eta_k p_k. \quad (86)$$

Similar to (84), for $n\neq k$ and $f\in\{0,1\}$, we can write

$$\begin{aligned}\mathbb{V}\text{ar}\big\{h_{n,j,f}|H_{tk}\big\}&=\mathbb{V}\text{ar}\big\{g_n b_{n,j-\alpha_n-1+f}|H_{tk}\big\} \quad (87)\\ &=\mathbb{E}\Big\{\mathbb{V}\text{ar}\big\{g_n b_{n,j-\alpha_n-1+f}|g_n,H_{tk}\big\}\Big\}\\ &\quad+\mathbb{V}\text{ar}\Big\{\mathbb{E}\big\{g_n b_{n,j-\alpha_n-1+f}|g_n,H_{tk}\big\}\Big\}.\end{aligned}$$

Because $\mathbb{V}\text{ar}\big\{b_{n,j-\alpha_n-1+f}|H_{tk}\big\}=\mathbb{V}\text{ar}\big\{b_{n,j-\alpha_n-1+f}\big\}=P_{\text{a}}$, $n\neq k$, $f\in\{0,1\}$, we can write

$$\begin{aligned}&\mathbb{E}\Big\{\mathbb{V}\text{ar}\big\{g_n b_{k,j-\alpha_n-1+f}|g_n,H_{tk}\big\}\Big\} \quad (88)\\ &=\mathbb{E}\big\{|g_n|^2\big\}\mathbb{V}\text{ar}\big\{b_{n,j-\alpha_n-1+f}|H_{tk}\big\}=P_{\text{a}}(\sigma_n^2+|\mu_n|^2)\eta_n p_n.\end{aligned}$$

By substituting (88) and $\mathbb{E}\big\{g_n b_{k,j-\alpha_n-1+f}|g_n,H_{tk}\big\}=0$, $f\in\{0,1\}$, into (87), we obtain

$$\mathbb{V}\text{ar}\big\{h_{n,j,f}|H_{tk}\big\}=P_{\text{a}}(\sigma_n^2+|\mu_n|^2)\eta_n p_n. \quad (89)$$

Finally, by substituting (86), (89), $\mathbb{V}\text{ar}\{w'_{k,j,f}\}=\Sigma^{w'}_{2k+f,2k+f}$ into (83), (37) is derived. For the cross-correlation of $\hat{h}_{k,j,0}$ and $\hat{h}_{k,j,1}$, we obtain (90) at the top of next page, where by substituting (86) and (89) into (90), and then by using $\mathbb{E}\{w'_{k,j,0}(w'_{k,j,1})^*\,|H_{tk}\}=\Sigma^{w'}_{2k,2k+1}$, results in (38).

Appendix B

By employing the MLR test, the transmission state of the kth IoT device is identified as active, i.e., $d_k=H_{1k}$, if

$$\frac{p\big(\breve{\mathbf{h}}_{k,j}|H_{1k}\big)}{p\big(\breve{\mathbf{h}}_{k,j}|H_{0k}\big)}=\frac{2\pi|\mathbf{C}^{0k}_{f,f}|^{\frac{1}{2}}\exp\Big(-\frac{1}{2}\breve{\mathbf{h}}^{\dagger}_{k,j}(\mathbf{C}^{1k}_{f,f})^{-1}\breve{\mathbf{h}}_{k,j}\Big)}{2\pi|\mathbf{C}^{1k}_{f,f}|^{\frac{1}{2}}\exp\Big(-\frac{1}{2}\breve{\mathbf{h}}^{\dagger}_{k,j}(\mathbf{C}^{0k}_{f,f})^{-1}\breve{\mathbf{h}}_{k,j}\Big)}>\lambda,$$

where $\lambda=(1-P_{\text{a}})/P_{\text{a}}$. A canonical form of the above detector is given by [31]

$$\breve{\mathbf{h}}^{\dagger}_{k,j}(\mathbf{C}^{0k}_{f,f})^{-1}(\mathbf{C}^{1k}_{f,f})(\mathbf{C}^{1k}_{f,f}+\mathbf{C}^{0k}_{f,f})^{-1}\breve{\mathbf{h}}_{k,j}>\theta_k, \quad (91)$$

where θ_k is determined based on desirable false alarm rate for the kth IoT device. Let us write $\mathbf{C}^{0k}_{f,f}=\mathbf{V}^{0k}_{f,f}\mathbf{\Lambda}^{0k}_{f,f}(\mathbf{V}^{0k}_{f,f})^{-1}$, where $\mathbf{V}^{0k}_{f,f}$ is an square matrix whose columns are eigenvectors of $\mathbf{C}^{0k}_{f,f}$, and $\mathbf{\Lambda}^{0k}_{f,f}$ is a diagonal matrix where its ith diagonal element is the eigenvalue associated with the ith column of $\mathbf{C}^{0k}_{f,f}$. We define $\mathbf{A}^{0k}_{f,f}=\mathbf{V}^{0k}_{f,f}(\mathbf{\Lambda}^{0k}_{f,f})^{-\frac{1}{2}}$ and $\mathbf{B}^{1k}_{f,f}\triangleq(\mathbf{A}^{0k}_{f,f})^{\dagger}\mathbf{C}^{1k}_{f,f}\mathbf{A}^{0k}_{f,f}$. Taking into account $(\mathbf{V}^{0k}_{f,f})^{\dagger}\mathbf{V}^{0k}_{f,f}=\mathbf{I}$, we can show that $(\mathbf{A}^{0k}_{f,f})^{\dagger}\mathbf{C}^{0k}_{f,f}\mathbf{A}^{0k}_{f,f}=\mathbf{I}$. Then, using this result, the canonical detector in (91) can be written as follows

$$\breve{\mathbf{h}}^{\dagger}_{k,j}\mathbf{A}^{0k}_{f,f}\mathbf{B}^{1k}_{f,f}(\mathbf{B}^{1k}_{f,f}+\mathbf{I})^{-1}(\mathbf{A}^{0k}_{f,f})^{\dagger}\breve{\mathbf{h}}_{k,j}>\theta_k. \quad (92)$$

Based on eigenvalue decomposition of $\mathbf{B}^{0k}_{f,f}$, we have

$$\mathbf{B}^{1k}_{f,f}\triangleq(\mathbf{A}^{0k}_{f,f})^{\dagger}\mathbf{C}^{1k}_{f,f}\mathbf{A}^{0k}_{f,f}=\mathbf{V}^{1k}_{f,f}\mathbf{\Lambda}^{1k}_{f,f}(\mathbf{V}^{1k}_{f,f})^{-1}, \quad (93)$$

where $\mathbf{V}^{1k}_{f,f}$ and $\mathbf{\Lambda}^{1k}_{f,f}$ are the eigenvector and eigenvalue matrices of $\mathbf{B}^{1k}_{f,f}$, respectively. Since $\mathbf{B}^{1k}_{f,f}$ is a symmetric matrix, we have $(\mathbf{V}^{1k}_{f,f})^{\dagger}\mathbf{V}^{1k}_{f,f}=\mathbf{I}$. By letting $\mathbf{z}_{k,j}\triangleq[\mathbf{z}_{k,j}[0],\mathbf{z}_{k,j}[1]]^{\dagger}\triangleq(\mathbf{V}^{1k}_{f,f})^{\dagger}(\mathbf{A}^{0k}_{f,f})^{\dagger}\breve{\mathbf{h}}_{k,j}$ in (92), we obtain

$$\mathbf{z}^{\dagger}_{k,j}\mathbf{\Lambda}^{1k}_{f,f}(\mathbf{\Lambda}^{1k}_{f,f}+\mathbf{I})^{-1}\mathbf{z}_{k,j}>\theta_k, \quad (94)$$

which is equivalent to the test statistics in (48) and (49). Note that the matrix $\mathbf{A}^{0k}_{f,f}\mathbf{V}^{1k}_{f,f}$ diagonalizes both $\mathbf{C}^{0k}_{f,f}$ and $\mathbf{C}^{1k}_{f,f}$.

By employing (48) and (49), the false alarm rate for the kth IoT device is derived as follows

$$\begin{aligned}P_k^{(\text{f})}&=\mathbb{P}\big\{d_k=H_{1k}\big|H_{0k}\big\} \quad (95)\\ &=\mathbb{P}\Big\{\sum_{n=0}^{1}\chi_{f,f}[n]z^2_{k,j}[n]\geqslant\theta_k\big|H_{0k}\Big\}.\end{aligned}$$

To obtain the PDF of $U\triangleq\sum_{n=0}^{1}\chi_{f,f}[n]z^2_{k,j}[n]$ in (95), we need to derive its characteristic function (CF) and then express the PDF as the inverse Fourier transform. Since $z^2_{k,j}[n]$ in (48) under hypothesis H_{0k} follows the central Chi-squared (χ^2) distribution with 1 degrees of freedom and the fact that $z^2_{k,j}[0]$ and $z^2_{k,j}[1]$ are independent random variables, we obtain the CF of U as follows

$$\phi_U(\omega)\triangleq\mathbb{E}\{\exp(j\omega U)\}=\prod_{n=0}^{1}\frac{1}{\sqrt{1-2j\chi_{f,f}[n]\omega}}, \quad (96)$$

where $\chi_{f,f}[n]$ is given (49). Taking the inverse Fourier transform of $\phi_U(\omega)$, we have

$$p_U(u|H_{0k})=\frac{1}{2\pi}\int_{-\infty}^{+\infty}\prod_{n=0}^{1}\frac{\exp(-j\omega u)}{\sqrt{1-2j\chi_{f,f}[n]\omega}}\text{d}\omega. \quad (97)$$

This article has been accepted for publication in IEEE Transactions on Wireless Communications. This is the author's version which has not been fully edited and content may change prior to final publication. Citation information: DOI 10.1109/TWC.2021.3117256

$$\Sigma_{f,f}^{tk} = \mathbb{V}\text{ar}\{\hat{h}_{k,j,f}|H_{tk}\} = \mathbb{E}\{|\hat{h}_{k,j,f}|^2|H_{tk}\} = t\Omega^2_{2k+f,2k+f}\mathbb{V}\text{ar}\{h_{k,j,f}|H_{tk}\} + t\Omega^2_{2k+f,2k+\bar{f}}\mathbb{V}\text{ar}\{h_{k,j,\bar{f}}|H_{tk}\} + \sum_{n\neq k}\Omega^2_{2k+f,2n+f}\mathbb{V}\text{ar}\{h_{n,j,f}|H_{tk}\} + \sum_{n\neq n}\Omega^2_{2k+f,2n+\bar{f}}\mathbb{V}\text{ar}\{h_{n,j,\bar{f}}|H_{tk}\} + \mathbb{V}\text{ar}\{w'_{k,j,f}\}. \tag{83}$$

$$\Sigma_{0,1}^{1k} = \mathbb{C}\text{ov}\{\hat{h}_{k,j,0},\hat{h}_{k,j,1}|H_{tk}\} = \mathbb{E}\{\hat{h}_{k,j,0}\hat{h}^*_{k,j,1}|H_{tk}\} = t(\Omega_{2k,2k}\Omega_{2k+1,2k})\mathbb{E}\{|h_{k,j,0}|^2|H_{tk}\} + t(\Omega_{2k+1,2k+1}\Omega_{2k,2k+1})\mathbb{E}\{|h_{k,j,1}|^2|H_{tk}\} + \sum_{n\neq k}\Omega_{2k,2n}\Omega_{2k+1,2n}\mathbb{E}\{|h_{n,j,0}|^2|H_{tk}\} + \sum_{n\neq k}\Omega_{2k+1,2n+1}\Omega_{2k,2n+1}\mathbb{E}\{|h_{n,j,1}|^2|H_{tk}\} + \mathbb{E}\{w'_{k,j,0}(w'_{k,j,1})^*|H_{tk}\}. \tag{90}$$

Using (97), the false alarm rate for the kth IoT device is obtained as in (51). Following the same procedure, the correct identification rate for the kth IoT device in (52) is derived.

References

[1] C. Bockelmann, N. Pratas, H. Nikopour, K. Au, T. Svensson, C. Stefanovic, P. Popovski, and A. Dekorsy, "Massive machine-type communications in 5G: Physical and MAC-layer solutions," *IEEE Commun. Mag.*, vol. 54, no. 9, pp. 59–65, Sept. 2016.

[2] M. Zhong, Y. Yang, H. Yao, X. Fu, O. A. Dobre, and O. Postolache, "5G and IoT: Towards a new era of communications and measurements," *IEEE Instrum. Meas. Mag.*, vol. 22, no. 6, pp. 18–26, Dec. 2019.

[3] G. Durisi, T. Koch, and P. Popovski, "Toward massive, ultrareliable, and low-latency wireless communication with short packets," *Proc. IEEE*, vol. 104, no. 9, pp. 1711–1726, Sep. 2016.

[4] A. Zanella, M. Zorzi, A. F. dos Santos, P. Popovski, N. Pratas, C. Stefanovic, A. Dekorsy, C. Bockelmann, B. Busropan, and T. A. Norp, "M2M massive wireless access: challenges, research issues, and ways forward," in *Proc. IEEE Globecom*, 2013, pp. 151–156.

[5] M. Mohammadkarimi, M. A. Raza, and O. A. Dobre, "Signature-based nonorthogonal massive multiple access for future wireless networks: Uplink massive connectivity for machine-type communications," *IEEE Veh. Technol. Mag.*, vol. 13, no. 4, pp. 40–50, Oct. 2018.

[6] R. B. Di Renna, C. Bockelmann, R. C. de Lamare, and A. Dekorsy, "Detection techniques for massive machine-type communications: Challenges and solutions," *IEEE Access*, vol. 8, pp. 180 928–180 954, 2020.

[7] H. Zhu and G. B. Giannakis, "Exploiting sparse user activity in multiuser detection," *IEEE Trans. Commun.*, vol. 59, no. 2, pp. 454–465, Feb. 2011.

[8] B. Knoop, F. Monsees, C. Bockelmann, D. Wübben, S. Paul, and A. Dekorsy, "Sparsity-aware successive interference cancellation with practical constraints," in *Proc. VDE WSA*, 2013, pp. 1–8.

[9] J. Ahn, B. Shim, and K. B. Lee, "Sparsity-aware ordered successive interference cancellation for massive machine-type communications," *IEEE Commun. Lett.*, vol. 7, no. 1, pp. 134–137, Feb. 2018.

[10] R. B. Di Renna and R. C. de Lamare, "Activity-aware multiple feedback SIC for massive machine-type communications," in *Proc. ITG*. VDE, 2019, pp. 1–6.

[11] ——, "Adaptive activity-aware iterative detection for massive machine-type communications," *IEEE Wireless Commun. Lett.*, vol. 8, no. 6, pp. 1631–1634, 2019.

[12] A. C. Cirik, N. M. Balasubramanya, and L. Lampe, "Multi-user detection using ADMM-based compressive sensing for uplink grant-free NOMA," *IEEE Wireless Commun. Lett.*, vol. 7, no. 1, pp. 46–49, 2017.

[13] H. F. Schepker and A. Dekorsy, "Sparse multi-user detection for CDMA transmission using greedy algorithms," in *Proc. IEEE ISWCS*, Aachen, Germany, Nov. 2011, pp. 291–295.

[14] ——, "Compressive sensing multi-user detection with block-wise orthogonal least squares," in *Proc. IEEE VTC*, Yokohama, Japan, May 2012, pp. 1–5.

[15] W. Xiong, J. Cao, and S. Li, "Sparse signal recovery with unknown signal sparsity," *EURASIP Journal on Advances in Signal Processing*, vol. 2014, no. 1, pp. 1–8, 2014.

[16] H. F. Schepker, C. Bockelmann, and A. Dekorsy, "Improving greedy compressive sensing based multi-user detection with iterative feedback," in *Proc. VTC*. IEEE, 2013, pp. 1–5.

[17] ——, "Efficient detectors for joint compressed sensing detection and channel decoding," *IEEE Trans. Commun.*, vol. 63, no. 6, pp. 2249–2260, 2015.

[18] Y. Du, C. Cheng, B. Dong, Z. Chen, X. Wang, J. Fang, and S. Li, "Block-sparsity-based multiuser detection for uplink grant-free NOMA," *IEEE Trans. Wireless Commun.*, vol. 17, no. 12, pp. 7894–7909, 2018.

[19] Z. Tang, J. Wang, J. Wang, and J. Song, "Device activity detection and non-coherent information transmission for massive machine-type communications," *IEEE Access*, vol. 8, pp. 41 452–41 465, 2020.

[20] L. Liu and W. Yu, "Massive connectivity with massive MIMO—Part I: Device activity detection and channel estimation," *IEEE J. Sel. Topics Signal Process.*, vol. 66, no. 11, pp. 2933–2946, June 2018.

[21] T. Ding, X. Yuan, and S. C. Liew, "Sparsity learning-based multiuser detection in grant-free massive-device multiple access," *IEEE Trans. Wireless Commun.*, vol. 18, no. 7, pp. 3569–3582, 2019.

[22] C. Wei, H. Liu, Z. Zhang, J. Dang, and L. Wu, "Approximate message passing-based joint user activity and data detection for NOMA," *IEEE Commun. Lett.*, vol. 21, no. 3, pp. 640–643, 2016.

[23] B. Wang, L. Dai, Y. Yuan, and Z. Wang, "Compressive sensing based multi-user detection for uplink grant-free non-orthogonal multiple access," in *Proc. IEEE VTC*, 2015, pp. 1–5.

[24] T. Wang, L. Shi, K. Cai, L. Tian, and S. Zhang, "Non-coherent NOMA with massive MIMO," *IEEE Wireless Commun. Lett.*, vol. 9, no. 2, pp. 134–138, 2019.

[25] D. S. Pham, A. M. Zoubir, R. F. Brcic, and Y. H. Leung, "A nonlinear m-estimation approach to robust asynchronous multiuser detection in Non-Gaussian noise," *IEEE Trans. Signal Process.*, vol. 55, no. 5, pp. 1624–1633, May 2007.

[26] A. E. Hoerl and R. W. Kennard, "Ridge regression: Biased estimation for nonorthogonal problems," *Technometrics*, vol. 12, no. 1, pp. 55–67, Feb. 1970.

[27] G. H. Golub, M. Heath, and G. Wahba, "Generalized cross-validation as a method for choosing a good ridge parameter," *Technometrics*, vol. 21, no. 2, pp. 215–223, May 1979.

[28] R. Ward, "Compressed sensing with cross validation," *IEEE Trans. Inf. Theory*, vol. 55, no. 12, pp. 5773–5782, Dec. 2009.

[29] P. S. Boonstra, B. Mukherjee, and J. M. Taylor, "A small-sample choice of the tuning parameter in ridge regression," *Statistica Sinica*, vol. 25, no. 3, p. 1185, July 2015.

[30] A. Bhattacharya, *Digital Communication*. Tata McGraw-Hill, 2005.

[31] S. M. Kay, *Fundamentals of Statistical Signal Processing, Vol. II: Detection Theory*, 1998.

[32] M. A. Al-Jarrah, M. A. Yaseen, A. Al-Dweik, O. A. Dobre, and E. Alsusa, "Decision fusion for IoT-based wireless sensor networks," *IEEE Internet Things J.*, vol. 7, no. 2, pp. 1313–1326, Feb. 2019.

[33] M. Yuan and Y. Lin, "Model selection and estimation in regression with grouped variables," *Journal of the Royal Statistical Society*, vol. 68, no. 1, pp. 49–67, Aug. 2006.

[34] A. Antoniou and W.-S. Lu, *Practical optimization: algorithms and engineering applications*. Springer Science & Business Media, 2007.

[35] P. Tseng, "Convergence of a block coordinate descent method for nondifferentiable minimization," *Journal of optimization theory and applications*, vol. 109, no. 3, pp. 475–494, 2001.

[36] S. Verdu, *Multiuser Detection*. Cambridge University Press, 1998.

[37] K. Teknomo, "K-means clustering tutorial," *Medicine*, vol. 100, no. 4, p. 3, 2006.

[38] C. Tang and C. Monteleoni, "On lloyd's algorithm: New theoretical insights for clustering in practice," in *Proc. Artificial Intelligence and Statistics*, 2016, pp. 1280–1289.

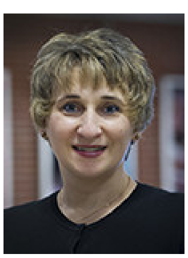
[39] S. Boyd, N. Parikh, and E. Chu, *Distributed optimization and statistical learning via the alternating direction method of multipliers*. Now Publishers Inc, 2011.

[40] K. Senel and E. G. Larsson, "Grant-free massive MTC-enabled massive MIMO: A compressive sensing approach," *IEEE Trans. Commun.*, vol. 66, no. 12, pp. 6164–6175, 2018.

© 2021 IEEE. Personal use is permitted, but republication/redistribution requires IEEE permission. See https://www.ieee.org/publications/rights/index.html for more information.

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TWC.2021.3117256, IEEE Transactions on Wireless Communications

Mostafa Mohammadkarimi (S'14, M'18) received the M.Sc. degree in electrical engineering from the K.N. Toosi University of Technology, Tehran, Iran, in 2011 and the Ph.D. degree in computer engineering from Memorial University, St. John's, NL, Canada, in 2017. From March 2016 to April 2017, he was a visiting Ph.D. student with the Laboratory for Information and Decision Systems, Massachusetts Institute of Technology (MIT), Cambridge, MA, USA. From February 2018 to April 2019, he was a Postdoctoral Fellow with the University of Alberta, Edmonton, AB, Canada. He was also an electronic design engineer at NexGen Control Solutions in 2019. From April 2019 to March 2020, he was a Postdoctoral Fellow with the University of British Columbia, Vancouver, BC, Canada. Then, from April 2020 to June 2021, he was a Postdoctoral Researcher with Institute for Digital Communications, Friedrich-Alexander-Universität Erlangen-Nürnberg, Erlangen, Nürnberg, Germany. Since September 2021, he has been with the Circuits and Systems group in the Department of Microelectronics at Delft University of Technology, CD Delft, Netherlands. His research interests include wireless communications, statistical signal processing, machine learning, and information theory. Dr. Mohammadkarimi has experience in designing and developing FPGA, DSP, SDR, and AVR microcontroller designs in various fields. He was a reviewer for numerous IEEE journals and a TPC member of international conferences. In 2014, he was the recipient of the Exemplary Reviewer Certificate of the IEEE COMMUNICATIONS LETTERS.



Octavia A. Dobre (M'05–SM'07–F'20) received the Dipl. Ing. and Ph.D. degrees from the Polytechnic Institute of Bucharest, Romania, in 1991 and 2000, respectively. Between 2002 and 2005, she was with New Jersey Institute of Technology, USA. In 2005, she joined Memorial University, Canada, where she is currently a Professor and Research Chair. She was a Visiting Professor with Massachusetts Institute of Technology, USA and Université de Bretagne Occidentale, France. Her research interests encompass wireless, optical and underwater communication technologies. She has (co-)authored over 400 refereed papers in these areas. Dr. Dobre serves as the Editor-in-Chief (EiC) of the IEEE OPEN JOURNAL OF THE COMMUNICATIONS SOCIETY. She was the EiC of the IEEE COMMUNICATIONS LETTERS, Senior Editor, Editor, and Guest Editor for various prestigious journals and magazines. She also served as General Chair, Technical Program Co-Chair, Tutorial Co-Chair, and Technical Co-Chair of symposia at numerous conferences. Dr. Dobre was a Fulbright Scholar, Royal Society Scholar, and Distinguished Lecturer of the IEEE Communications Society. She obtained Best Paper Awards at various conferences, including IEEE ICC, IEEE GLOBECOM, IEEE WCNC, and IEEE PIMRC. Dr. Dobre is a Fellow of the Engineering Institute of Canada and a Fellow of the Canadian Academy of Engineering.

Moe Z. Win Moe Z. Win (S'85-M'87-SM'97-F'04) is a Professor at the Massachusetts Institute of Technology (MIT) and the founding director of the Wireless Information and Network Sciences Laboratory. Prior to joining MIT, he was with AT&T Research Laboratories and with NASA Jet Propulsion Laboratory. His research encompasses fundamental theories, algorithm design, and network experimentation for a broad range of real-world problems. His current research topics include network localization and navigation, network interference exploitation, and quantum information science. He has served the IEEE Communications Society as an elected Member-at-Large on the Board of Governors, as elected Chair of the Radio Communications Committee, and as an IEEE Distinguished Lecturer. Over the last two decades, he held various editorial positions for IEEE journals and organized numerous international conferences. Recently, he has served on the SIAM Diversity Advisory Committee.

Dr. Win is an elected Fellow of the AAAS, the EURASIP, the IEEE, and the IET. He was honored with two IEEE Technical Field Awards: the IEEE Kiyo Tomiyasu Award (2011) and the IEEE Eric E. Sumner Award (2006, jointly with R. A. Scholtz). His publications, co-authored with students and colleagues, have received several awards. Other recognitions include the IEEE Communications Society Edwin H. Armstrong Achievement Award (2016), the Cristoforo Colombo International Prize for Communications (2013), the Copernicus Fellowship (2011) and the *Laurea Honoris Causa* (2008) from the Università degli Studi di Ferrara, and the U.S. Presidential Early Career Award for Scientists and Engineers (2004). He is an ISI Highly Cited Researcher.

© 2021 IEEE. Personal use is permitted, but republication/redistribution requires IEEE permission. See https://www.ieee.org/publications/rights/index.html for more information.