Hyperparameter-Free Receiver for Grant-Free NOMA Systems With MIMO-OFDM

Takanori Hara, *Graduate Student Member, IEEE*, Hiroki Iimori, *Graduate Student Member, IEEE*, and Koji Ishibashi, *Senior Member, IEEE*

***Abstract*—We introduce a new grant-free non-orthogonal multiple access (GF-NOMA) scheme, which unlike conventional approaches, spreads over the frequency domain instead of the time domain to reduce access latency. Most conventional receivers for GF-NOMA require the pre-tuning of hyperparameters to achieve superior performance. To avoid such pre-tuning, we propose a hyperparameter-free receiver for the proposed architecture, exploiting the sparsity of channels in the delay domain. Simulation results demonstrate the superior performance of the proposed receiver in terms of the normalized mean-squared error and the activity error rate.**

***Index Terms*—Massive connectivity, grant-free, NOMA, active user detection, channel estimation, hyperparameter-free.**

I. Introduction

With the rapid development of Internet-of-Things (IoT), new wireless communication systems must be designed to cope with the exponential growth in the number of wireless devices. In this context, future systems must not only address sporadic data traffic but also meet the demands for both massive connectivity and low latency when IoT networks are applied to delay-sensitive applications such as robot control and automated driving. Conventional multiple access schemes use a grant-based approach wherein a base station (BS) assigns resource blocks to users ahead of data transmission, causing a huge control-signaling overhead.

As promising techniques to overcome this impediment, *grant-free* random access schemes such as grant-free non-orthogonal multiple access (GF-NOMA) have gained much attention [1]. A grant-free strategy does not exclusively assign radio resources to each user for data transmission, and thus, each active user in GF-NOMA systems can directly transmit its packet to the BS without waiting for any permission. However, there are fundamental challenges to realize massive connectivity and low latency, namely: 1) active user detection (AUD), 2) channel estimation (CE), and 3) multiuser detection (MUD).

Manuscript received September 26, 2020; accepted December 11, 2020. Date of publication December 16, 2020; date of current version April 9, 2021. This work was supported by the Ministry of Internal Affairs and Communications in Japan under Grant JPJ000254. The associate editor coordinating the review of this article and approving it for publication was Y. Huang. *(Corresponding author: Takanori Hara.)*

Takanori Hara and Koji Ishibashi are with the Advanced Wireless & Communication Research Center, The University of Electro-Communications, Tokyo 182-8585, Japan (e-mail: hara@awcc.uec.ac.jp; koji@ieee.org).

Hiroki Iimori is with the Focus Area Mobility, Department of Computer Science and Electrical Engineering, Jacobs University Bremen, 28759 Bremen, Germany (e-mail: h.iimori@ieee.org).

Digital Object Identifier 10.1109/LWC.2020.3045159

Recently, many studies have focused on utilizing sparsity to address the issues discussed above [2]–[10]. These works utilize the non-orthogonal sequences spread over the *time domain*, and a BS estimates the active users from the overlapped measurements via sparse-recovery techniques. However, this transmission scheme requires a sufficiently long sequence in the time domain to accommodate massive users efficiently, leading to the difficulties in meeting the latency requirements. In addition, as a waveform based on orthogonal division frequency multiplexing (OFDM) is employed for both downlink and uplink transmissions in fifth-generation new radio (5G NR) [11], the latency issue is more severe due to the extension of OFDM symbols, induced by the spreading over the time domain. To reduce the access latency in GF-NOMA based on massive multiple-input-multiple-output (MIMO)-OFDM, a scheme to perform AUD and CE in an alternating manner was proposed in [7]. However, the scheme proposed in [7] still requires dozens of OFDM symbols to perform accurate estimation. Hence, an alternative that makes use of the frequency domain while further reducing access latency and retaining the advantages of GF-NOMA is required.

Compressed sensing (CS)-based random access with multicarrier transmission and its associated pilot design was investigated in [12]. However, this letter made an impractical assumption in that the BS perfectly knows the number of potential active users and the maximum delay spread among all users in advance. In practice, the BS includes the parameters under consideration (*i.e.*, the number of active users and the maximum delay length), and cannot utilize prior channel statistics as the reference, owing to sporadic traffic and mobility of uplink users, *e.g.*, robots and cars. Hence, a receiver without the need for such knowledge is desired, and several approaches based on the message passing mechanism have been proposed [13], [14]. However, they require updates of many messages at each iteration, resulting in prohibitive computational complexity at the BS as the number of system parameters increases. Besides, the conventional receivers for GF-NOMA systems, *e.g.*, [3]–[5], require an exhaustive search of their design parameters called *hyperparameters* for all (time-varying) system parameters such as signal-to-noise ratio (SNR) and user activity ratio. Doubtlessly, it would enforce a huge overhead to obtain those system parameters and tune the corresponding hyperparameters before transmission in practice. Hence, a *hyperparameter-free* receiver, which does not require tuning of any hyperparameters, is strongly desired.

In this letter, we propose a *hyperparameter-free* receiver for GF-NOMA with massive MIMO-OFDM, incorporating the pilot design considered in [14] to cope with issues

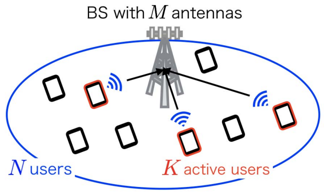

Fig. 1. Uplink GF-NOMA system.

described above. Our main contributions can be summarized as follows:

- Our proposed receiver based on [15] utilizes the sparsity of channels in the delay domain and avoids any pre-tuning of resultant hyperparameters.
- We reveal that the notable mathematical connection between the methods of [15] and [3] and propose a hyperparameter-free receiver based on [3].
- The proposed scheme is superior to the classical CS algorithms exploiting different sparseness and comparable to the state-of-the-art schemes for GF-NOMA systems.

II. System Model

We consider an uplink GF-NOMA system comprising N potential users, and a BS equipped with M antennas, as shown in Fig. 1. In addition, we assume that the BS is equipped with a one-dimensional uniform linear array (ULA) so that the antennas are separated by one-half wavelength and that all users are time-synchronized. Throughout this letter, $K < N$ users are active in each coherence time.

The uplink transmission is organized in OFDM symbols, where N_p pilot subcarriers are uniformly allocated to N_c subcarriers. All users share the same subcarrier locations for pilot transmission, and the subset of pilot subcarrier indices is denoted by $\mathcal{P} \subset \{1, 2, \ldots, N_\mathrm{c}\}$, with $|\mathcal{P}| = N_\mathrm{p}$. Let $\mathbf{Y} \in \mathbb{C}^{N_\mathrm{p} \times M}$ denote the received signals at all the receiver antennas in the subset $\mathcal{P}$ after cyclic prefix removal and discrete Fourier transform (DFT) modulation. Then, the received signals are given by

$$\begin{aligned} \mathbf{Y} &= \sum_{n \in \mathcal{A}} \mathrm{diag}(\mathbf{s}_n)\mathbf{G}_n + \mathbf{Z} \\ &\triangleq \sum_{n \in \mathcal{A}} \mathbf{S}_n \mathbf{G}_n + \mathbf{Z}, \end{aligned} \tag{1}$$

where $\mathcal{A}$ denotes the set of active users, $\mathbf{S}_n = \mathrm{diag}(\mathbf{s}_n)$ is a diagonal matrix based on the pilot sequence of user n, denoted by $\mathbf{s}_n \in \mathbb{C}^{N_\mathrm{p} \times 1}$, which is assumed to be unimodular, *i.e.*, $|s_{n,p}| = 1$ for $p = 1, \ldots, N_\mathrm{p}$. $\mathbf{G}_n = [\mathbf{g}_{n,1}, \ldots, \mathbf{g}_{n,N_\mathrm{p}}]^\mathrm{T} \in \mathbb{C}^{N_\mathrm{p} \times M}$ is the channel frequency response (CFR) between BS and user n over subcarriers $\mathcal{P}$. The matrix $\mathbf{Z} \in \mathbb{C}^{N_\mathrm{p} \times M}$ represents the noise, where the elements follow a complex Gaussian distribution with zero mean and variance σ_n^2. In this letter, we utilize the pilot design proposed in [14], assuming $N \leq N_\mathrm{p}$ and $N_\mathrm{p} < NN_\mathrm{cp}$, where N_cp denotes a cyclic prefix length.

For the p-th pilot subcarrier ($p \in \mathcal{P}$), the sub-channel of the n-th user can be modeled as follows [16]

$$\mathbf{g}_{n,p} = \sum_{\ell=1}^{L_\mathrm{path}} \beta_{n,\ell}\mathbf{a}(\phi_{n,\ell})\mathrm{e}^{-j2\pi\tau_{n,\ell}\left(-\frac{B_s}{2}+\frac{B_s(pN_\mathrm{c}/N_\mathrm{p}-1)}{N_\mathrm{c}}\right)} \in \mathbb{C}^{M \times 1}, \tag{2}$$

where $N_\mathrm{c}/N_\mathrm{p}$ is an integer, L_path represents the number of multi-path components (MPCs), $\beta_{n,\ell} \sim \mathcal{CN}(0, 1/L_\mathrm{path})$ and $\tau_{n,\ell} \in [0, N_\mathrm{cp}/B_s]$ are the complex path gain and the path delay of the ℓ-th MPC, respectively. B_s is the two-sided bandwidth, and the antenna array response vector is denoted by $\mathbf{a}(\phi_{n,\ell}) = [1, \mathrm{e}^{-j2\pi\phi_{n,\ell}}, \ldots, \mathrm{e}^{-j2\pi(M-1)\phi_{n,\ell}}]^\mathrm{T} \in \mathbb{C}^{M \times 1}$, with $\phi_{n,\ell} = d\sin(\psi_{n,\ell})/\lambda$, where $\psi_{n,\ell}$, λ, and d denote the angle of arrival of the n-th user's ℓ-th MPC, the wavelength, and the antenna spacing, respectively. Throughout this letter, we assume that $d = \lambda/2$ and $\psi_{n,\ell}$ is uniformly distributed in $[-\pi, \pi]$, resulting in $\phi_{n,\ell} \in [-1/2, 1/2]$.

The CFR can be represented by

$$\mathbf{G}_n = \sqrt{N_\mathrm{c}}\mathbf{F}_{N_\mathrm{p},N_\mathrm{cp}}\mathbf{H}_n, \tag{3}$$

where $\mathbf{H}_n = [\mathbf{h}_n^1, \ldots, \mathbf{h}_n^M] \in \mathbb{C}^{N_\mathrm{cp} \times M}$ denotes the channel impulse response (CIR) from the n-th user to the BS. The matrix $\mathbf{F}_{N_\mathrm{p},N_\mathrm{cp}} \in \mathbb{C}^{N_\mathrm{p} \times N_\mathrm{cp}}$ is the sub-matrix of the $N_\mathrm{c} \times N_\mathrm{c}$ DFT matrix $\mathbf{F}_{N_\mathrm{c}}$, and contains the N_p rows according to $\mathcal{P}$, and the first N_cp columns of $\mathbf{F}_{N_\mathrm{c}}$.

Let $\bar{\mathbf{F}}_{N_\mathrm{p},N_\mathrm{cp}} \in \mathbb{C}^{N_\mathrm{p} \times N_\mathrm{cp}}$ be the matrix $\mathbf{F}_{N_\mathrm{p},N_\mathrm{cp}}$, with all column vectors normalized. Then, the received signals can be expressed with the CIRs as follows:

$$\begin{aligned} \mathbf{Y} &= \sum_{n \in \mathcal{A}} \mathbf{S}_n(\sqrt{N_\mathrm{p}}\bar{\mathbf{F}}_{N_\mathrm{p},N_\mathrm{cp}}\mathbf{H}_n) + \mathbf{Z} \\ &= \sum_{n \in \mathcal{A}} \mathbf{S}_n\bar{\mathbf{F}}_{N_\mathrm{p},N_\mathrm{cp}}\mathbf{X}_n + \mathbf{Z}, \end{aligned} \tag{4}$$

where $\mathbf{X}_n \triangleq \sqrt{N_\mathrm{p}}\mathbf{H}_n$ and $\sqrt{N_\mathrm{p}}\bar{\mathbf{F}}_{N_\mathrm{p},N_\mathrm{cp}} = \sqrt{N_\mathrm{c}}\mathbf{F}_{N_\mathrm{p},N_\mathrm{cp}}$. We define $\mathbf{A} = [\mathbf{S}_1\bar{\mathbf{F}}_{N_\mathrm{p},N_\mathrm{cp}}, \ldots, \mathbf{S}_N\bar{\mathbf{F}}_{N_\mathrm{p},N_\mathrm{cp}}] \in \mathbb{C}^{N_\mathrm{p} \times NN_\mathrm{cp}}$ and $\mathbf{X} = [\mathbf{X}_1^\mathrm{T}, \ldots, \mathbf{X}_N^\mathrm{T}]^\mathrm{T} \in \mathbb{C}^{NN_\mathrm{cp} \times M}$, and (4) can be simplified as

$$\mathbf{Y} = \mathbf{AX} + \mathbf{Z}. \tag{5}$$

Here, all column vectors of $\mathbf{A}$ are normalized. For massive MIMO systems, the inherent sparsity in the angle domain [7], [17] can be exploited. However, we take advantage of the row sparsity of $\mathbf{X}$ in (5) to avoid an increase in the number of required parameters in the estimation scheme.

III. Proposed Method

In this section, we propose two hyperparameter-free receivers. Unlike the existing approaches, including that of [12], our proposals exploit the row-sparsity instead of the block sparsity of $\mathbf{X}$ in (5). After formulating the problem for the user activity and channel estimation as an maximum likelihood (ML) problem [3], we reveal a nontrivial mathematical connection between ML [3] and SPARROW [15], which is the $\ell_{2,1}$ mixed-norm minimization problem. Then, considering the connection, we propose hyperparameter-free receivers based on SPARROW or ML to perform the accurate estimation without any pre-tuning of algorithmic parameters.

A. ML-Based Problem Formulation

Based on [3], we define γ_n to represent the channel strengths involving the associated activity patterns, while

$\mathbf{\Gamma} \triangleq \operatorname{diag}(\boldsymbol{\gamma})$ with $\boldsymbol{\gamma} \triangleq [\gamma_1, \ldots, \gamma_{NN_{\mathrm{cp}}}]^{\mathrm{T}}$. Then, (5) can be rewritten as

$$\mathbf{Y} = \mathbf{A}\mathbf{\Gamma}^{\frac{1}{2}}\bar{\mathbf{X}} + \mathbf{Z}, \tag{6}$$

where $\mathbf{X} = \mathbf{\Gamma}^{\frac{1}{2}}\bar{\mathbf{X}}$, and $\bar{\mathbf{X}}$ is the matrix whose each row obeys the complex standard Gaussian distribution. Based on the model given by (6), the ML estimation problem can be represented as [3], [5]

$$\underset{\mathbf{\Gamma}\in\mathbb{D}_+}{\text{minimize}} \ \log|\mathbf{A}\mathbf{\Gamma}\mathbf{A}^{\mathrm{H}} + \sigma_n^2\mathbf{I}_{N_{\mathrm{p}}}| + \operatorname{tr}((\mathbf{A}\mathbf{\Gamma}\mathbf{A}^{\mathrm{H}} + \sigma_n^2\mathbf{I}_{N_{\mathrm{p}}})^{-1}\hat{\mathbf{R}}), \tag{7}$$

where $\hat{\mathbf{R}} = \mathbf{Y}\mathbf{Y}^{\mathrm{H}}/M$ denotes the sample covariance matrix, and $\mathbb{D}_+$ denotes the set of non-negative diagonal matrices.

B. Nontrivial Connection Between ML and SPARROW

In this subsection, a mathematical connection between ML and SPARROW is presented, demonstrating how to obtain an estimate of $\mathbf{X}$ from (6). As the matrix $\mathbf{A}\mathbf{\Gamma}\mathbf{A}^{\mathrm{H}} + \sigma_n^2\mathbf{I}_{N_{\mathrm{p}}}$ is positive-definite, the following inequality holds

$$\log|\mathbf{A}\mathbf{\Gamma}\mathbf{A}^{\mathrm{H}} + \sigma_n^2\mathbf{I}_{N_{\mathrm{p}}}| \leq \operatorname{tr}(\mathbf{A}\mathbf{\Gamma}\mathbf{A}^{\mathrm{H}} + \sigma_n^2\mathbf{I}_{N_{\mathrm{p}}} - \mathbf{I}_{N_{\mathrm{p}}}), \tag{8}$$

which readily yields the following relaxed problem of (7):

$$\underset{\mathbf{\Gamma}\in\mathbb{D}_+}{\text{minimize}} \ \operatorname{tr}(\mathbf{A}\mathbf{\Gamma}\mathbf{A}^{\mathrm{H}}) + \operatorname{tr}((\mathbf{A}\mathbf{\Gamma}\mathbf{A}^{\mathrm{H}} + \sigma_n^2\mathbf{I}_{N_{\mathrm{p}}})^{-1}\hat{\mathbf{R}}). \tag{9}$$

Since the matrix $\mathbf{A}\mathbf{\Gamma}\mathbf{A}^{\mathrm{H}}$ can be expressed as $\sum_{n=1}^{NN_{\mathrm{cp}}}\gamma_n\mathbf{a}_n\mathbf{a}_n^{\mathrm{H}}$, the first term of (9) can be transformed into $\sum_{n=1}^{NN_{\mathrm{cp}}}\gamma_n\|\mathbf{a}_n\|_2^2$ after some mathematical manipulations. Moreover, $\sum_{n=1}^{NN_{\mathrm{cp}}}\gamma_n\|\mathbf{a}_n\|_2^2$ is equal to $\sum_{n=1}^{NN_{\mathrm{cp}}}\gamma_n = \operatorname{tr}(\mathbf{\Gamma})$ owing to the fact that all columns of the measurement matrix $\mathbf{A}$ are normalized, *i.e.*, $\|\mathbf{a}_n\|_2^2 = 1$. Interestingly, when the first term of (9) is expressed by the trace of $\mathbf{\Gamma}$ according to the above, (9) can be regarded as a special case of SPARROW [15, Th. 1] with $\rho = \sigma_n^2$:

$$\underset{\mathbf{\Gamma}\in\mathbb{D}_+}{\text{minimize}} \ \operatorname{tr}(\mathbf{\Gamma}) + \operatorname{tr}((\mathbf{A}\mathbf{\Gamma}\mathbf{A}^{\mathrm{H}} + \rho\mathbf{I}_{N_{\mathrm{p}}})^{-1}\hat{\mathbf{R}}), \tag{10}$$

which mathematically demonstrates that SPARROW is indeed a relaxed variate of ML.

According to [15], the estimate of $\mathbf{X}$ in (6) is given by

$$\hat{\mathbf{X}} = \hat{\mathbf{\Gamma}}\mathbf{A}^{\mathrm{H}}(\mathbf{A}\hat{\mathbf{\Gamma}}\mathbf{A}^{\mathrm{H}} + \rho\mathbf{I}_{N_{\mathrm{p}}})^{-1}\mathbf{Y}, \tag{11}$$

where $\hat{\mathbf{\Gamma}}$ denotes an estimate of $\mathbf{\Gamma}$. Although SPARROW [15] and ML [3] can efficiently solve (7) and (10) by the coordinate descent (CD) method, they require a pre-determined parameter, such as the noise variance σ_n^2 or the pre-tuned hyperparameter ρ, to execute. Hence, an approach without any knowledge of such parameters is needed, which will therefore be described in the following subsection.

C. Hyperparameter-Free Activity and Channel Estimation

To avoid any pre-determined parameters, the proposed scheme utilizes a reformulation of the matrix $\mathbf{A}\mathbf{\Gamma}\mathbf{A}^{\mathrm{H}} + \rho\mathbf{I}_{N_{\mathrm{p}}}$ with the second term denoting the covariance matrix of the additive white Gaussian noise (AWGN) noise at the receiver models the covariance matrix of each column of $\mathbf{Y}$, *i.e.*, $\mathbf{y}_m$. Let us assume that each diagonal entry of the covariance matrix of the noise can be replaced with an arbitrary variable, unlike in the original SPARROW formulation. Then, the covariance matrix of $\mathbf{y}_m$ can be modeled as

$$\begin{aligned}\mathbf{\Sigma} &= \mathbf{A}\mathbf{\Gamma}\mathbf{A}^{\mathrm{H}} + \begin{bmatrix}\nu_1 & & & \\ & \nu_2 & & \\ & & \ddots & \\ & & & \nu_{N_{\mathrm{p}}}\end{bmatrix} \\ &\triangleq \bar{\mathbf{A}}\bar{\mathbf{\Gamma}}\bar{\mathbf{A}}^{\mathrm{H}},\end{aligned} \tag{12}$$

where $\bar{\mathbf{A}} = [\mathbf{A} \ \mathbf{I}_{N_{\mathrm{p}}}] \in \mathbb{C}^{N_{\mathrm{p}}\times(NN_{\mathrm{cp}}+N_{\mathrm{p}})}$, and $\bar{\mathbf{\Gamma}}$ is the following diagonal matrix:

$$\bar{\mathbf{\Gamma}} = \begin{bmatrix}\mathbf{\Gamma} & & & \\ & \nu_1 & & \\ & & \ddots & \\ & & & \nu_{N_{\mathrm{p}}}\end{bmatrix}. \tag{13}$$

Fortunately, even with this reformulation, $\mathbf{X}$ can be estimated from (11) as

$$\hat{\mathbf{X}} = \hat{\mathbf{\Gamma}}\mathbf{A}^{\mathrm{H}}\hat{\mathbf{\Sigma}}^{-1}\mathbf{Y}, \tag{14}$$

where $\hat{\mathbf{\Sigma}}$ denotes an estimate of the covariance matrix $\mathbf{\Sigma}$. It is worth noting that the above reformulation can be applied to the ML estimation since (7) involves a similarly structured covariance matrix, *i.e.*, $\mathbf{A}\mathbf{\Gamma}\mathbf{A}^{\mathrm{H}} + \sigma_n^2\mathbf{I}_{N_{\mathrm{p}}}$. Therefore, two different hyperparameter-free schemes according to ML and SPARROW can be presented for joint activity and channel estimation, respectively.

Following the CD method proposed in [15], the proposed scheme iteratively updates the variables in a coordinate fashion to minimize the objective function (7) or (10), where the

---

Algorithm 1 Hyperparameter-Free CD Method

Input: $\hat{\mathbf{R}} = \frac{1}{M}\mathbf{Y}\mathbf{Y}^{\mathrm{H}} \in \mathbb{C}^{N_{\mathrm{p}}\times N_{\mathrm{p}}}, \mathbf{A} \in \mathbb{C}^{N_{\mathrm{p}}\times NN_{\mathrm{cp}}}, \rho_0 = \frac{\|\mathbf{Y}\|_F^2}{N_{\mathrm{p}}M}$, The number of iterations T.

Initialize $\hat{\mathbf{\Sigma}}^{-1} = \frac{1}{\rho_0}\mathbf{I}_{N_{\mathrm{p}}}, \hat{\boldsymbol{\gamma}} = \mathbf{0}_{NN_{\mathrm{cp}}}, \hat{\nu}_1 = \cdots = \hat{\nu}_{N_{\mathrm{p}}} = \rho_0$.

for $k = 1, 2, \ldots, T$ **do**

  Randomly select a permutation $i_1, i_2, \ldots, i_{NN_{\mathrm{cp}}}$ of the coordinate indices $\{1, 2, \ldots, NN_{\mathrm{cp}}\}$ of $\hat{\boldsymbol{\gamma}}$.

  for $n = 1, 2, \ldots, NN_{\mathrm{cp}}$ **do**

  $$\delta = \begin{cases}\max\left\{\dfrac{\mathbf{a}_{i_n}^{\mathrm{H}}\hat{\mathbf{\Sigma}}^{-1}\hat{\mathbf{R}}\hat{\mathbf{\Sigma}}^{-1}\mathbf{a}_{i_n} - \mathbf{a}_{i_n}^{\mathrm{H}}\hat{\mathbf{\Sigma}}^{-1}\mathbf{a}_{i_n}}{(\mathbf{a}_{i_n}^{\mathrm{H}}\hat{\mathbf{\Sigma}}^{-1}\mathbf{a}_{i_n})^2}, -\hat{\gamma}_{i_n}\right\} & \text{(ML)} \\ \max\left\{\dfrac{\sqrt{\mathbf{a}_{i_n}^{\mathrm{H}}\hat{\mathbf{\Sigma}}^{-1}\hat{\mathbf{R}}\hat{\mathbf{\Sigma}}^{-1}\mathbf{a}_{i_n}} - 1}{\mathbf{a}_{i_n}^{\mathrm{H}}\hat{\mathbf{\Sigma}}^{-1}\mathbf{a}_{i_n}}, -\hat{\gamma}_{i_n}\right\} & \text{(SPARROW)}\end{cases}$$

  $\hat{\gamma}_{i_n} \leftarrow \hat{\gamma}_{i_n} + \delta$

  $\hat{\mathbf{\Sigma}}^{-1} \leftarrow \hat{\mathbf{\Sigma}}^{-1} - \delta\frac{\hat{\mathbf{\Sigma}}^{-1}\mathbf{a}_{i_n}\mathbf{a}_{i_n}^{\mathrm{H}}\hat{\mathbf{\Sigma}}^{-1}}{1+\delta\mathbf{a}_{i_n}^{\mathrm{H}}\hat{\mathbf{\Sigma}}^{-1}\mathbf{a}_{i_n}}$

  end for

  for $n = 1, 2, \ldots, N_{\mathrm{p}}$ **do**

  $$\delta = \begin{cases}\max\left\{\dfrac{\mathbf{e}_n^{\mathrm{H}}\hat{\mathbf{\Sigma}}^{-1}\hat{\mathbf{R}}\hat{\mathbf{\Sigma}}^{-1}\mathbf{e}_n - \mathbf{e}_n^{\mathrm{H}}\hat{\mathbf{\Sigma}}^{-1}\mathbf{e}_n}{(\mathbf{e}_n^{\mathrm{H}}\hat{\mathbf{\Sigma}}^{-1}\mathbf{e}_n)^2}, -\hat{\nu}_n\right\} & \text{(ML)} \\ \max\left\{\dfrac{\sqrt{\mathbf{e}_n^{\mathrm{H}}\hat{\mathbf{\Sigma}}^{-1}\hat{\mathbf{R}}\hat{\mathbf{\Sigma}}^{-1}\mathbf{e}_n} - 1}{\mathbf{e}_n^{\mathrm{H}}\hat{\mathbf{\Sigma}}^{-1}\mathbf{e}_n}, -\hat{\nu}_n\right\} & \text{(SPARROW)}\end{cases}$$

  $\hat{\nu}_n \leftarrow \hat{\nu}_n + \delta$

  $\hat{\mathbf{\Sigma}}^{-1} \leftarrow \hat{\mathbf{\Sigma}}^{-1} - \delta\frac{\hat{\mathbf{\Sigma}}^{-1}\mathbf{e}_n\mathbf{e}_n^{\mathrm{H}}\hat{\mathbf{\Sigma}}^{-1}}{1+\delta\mathbf{e}_n^{\mathrm{H}}\hat{\mathbf{\Sigma}}^{-1}\mathbf{e}_n}$

  end for

end for

Output: $\hat{\boldsymbol{\gamma}} \in \mathbb{R}^{NN_{\mathrm{cp}}\times 1}, \hat{\mathbf{\Sigma}}^{-1} \in \mathbb{C}^{N_{\mathrm{p}}\times N_{\mathrm{p}}}$.

covariance matrix including a hyperparameter is replaced with the matrix given in (12).[1] In contrast to the approach in [15], the proposed CD method based on the parameter-free formulation described above jointly estimates the noise variance, activity patterns, and channel strengths, without pre-tuning of σ_n^2 or ρ.

In light of all the above, an algorithmic flow of the proposed scheme is summarized in Algorithm 1. Notice that we assume that $\nu_1 = \cdots = \nu_{N_\mathrm{p}}$ in the initialization step, which implicitly sets $\hat{\mathbf{\Sigma}}^{-1}$ to $\rho_0^{-1}\mathbf{I}_{N_\mathrm{p}}$ with $\rho_0 = \|\mathbf{Y}\|_F^2/(N_\mathrm{p}M)$. In addition, in each outer iteration of the proposed algorithm, estimates of $\nu_1, \ldots, \nu_{N_\mathrm{p}}$, *i.e.*, $\hat{\nu}_1, \ldots, \hat{\nu}_{N_\mathrm{p}}$, are computed after the updates of the coordinates corresponding to $\boldsymbol{\gamma}$. The update of ν_n for $n = 1, 2, \ldots, N_\mathrm{p}$ requires the *n*-th canonical basis vector with 1 only at its *n*-th entry and zero elsewhere, which is denoted by $\mathbf{e}_{n_1}$.

The computational complexity of the proposed method is mainly owing to the matrix-vector multiplications, and the complexity order required for each outer iteration is $O((NN_\mathrm{cp} + N_\mathrm{p})N_p^2)$. While its complexity grows with the cubic of N_p, this can be acceptable for small N_p, and the proposed scheme can be considered efficient owing to the outstanding convergence rate of the CD method that was shown by numerical results in [15].

After processing the CD method, we obtain an estimate of $\mathbf{X}$ by (14). Furthermore, we determine the estimated set of $\mathcal{A}$ as follows

$$\hat{\mathcal{A}} = \left\{ n \mid \|\hat{\mathbf{H}}_n\|_F \geq \eta \|\hat{\mathbf{H}}_{\max}\|_F \text{ and } \hat{\mathbf{H}}_{\max} = \max_{n=1,\ldots,N} \|\hat{\mathbf{H}}_n\|_F \right\}, \tag{15}$$

with $\eta = 0.1$ denoting the ratio of the minimum and maximum Frobenius norms of the channel coefficients.[2]

IV. Numerical Results

In this section, we investigate the normalized mean-squared error (NMSE) performance and the activity error rate (AER), which is the probability of the activity pattern of each user to be detected incorrectly. In this letter, we define the NMSE as

$$\mathrm{NMSE} \triangleq \mathbb{E}\left[\frac{\|\mathbf{H} - \hat{\mathbf{H}}\|_F^2}{\|\mathbf{H}\|_F^2}\right]. \tag{16}$$

For the simulation results, the bandwidth is $B_s = 10$ MHz, and the number of significant paths L_path is 6. The system adopts OFDM, where $N_\mathrm{c} = 1024$ subcarriers and a cyclic prefix of length $N_\mathrm{cp} = 32$ are employed. The SNR is defined by the ratio of the pilot norm to the noise variance as $\mathrm{SNR} \triangleq \|\mathbf{s}_n\|_2^2/(N_\mathrm{p}\sigma_n^2) = 1/\sigma_n^2$. We assume that the BS is equipped with $M = 32$ antennas, and $N = 64$ potential users exist while $K = 8$ users are active unless otherwise specified. In addition, the number of iterations of the proposed scheme, denoted by T, is set to 10.

[1] In order to avoid redundancy, we omit details about the derivation of the update rules. Please refer to [3], [15] for details.

[2] Although η is a tunable parameter, the search for an optimal value is beyond the scope of this letter. If this search is conducted, a receiver operating characteristic needs to be evaluated [3].

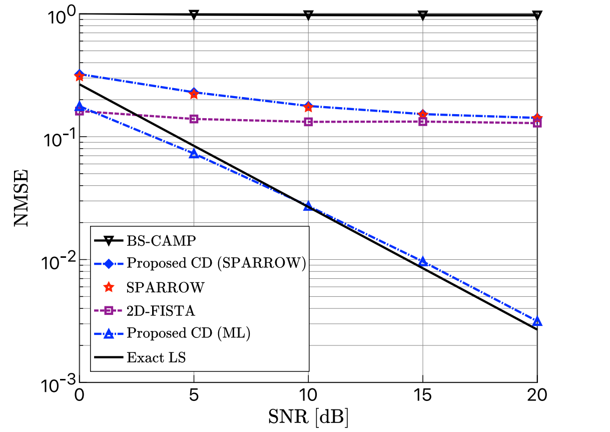

Fig. 2. NMSE performance of the proposed scheme and classical CS algorithms for $N_\mathrm{p} = 64$.

Fig. 2 shows the NMSE performance for $N_\mathrm{p} = 64$. As a benchmark, we evaluate the performance of a complex approximate message passing (CAMP) exploiting block-sparsity, namely, BS-CAMP, with 50 iterations. This scheme does not require prior impractical knowledge at the BS, unlike the approach in [12]. Moreover, we show the performance of the scheme of [18], *i.e.*, 2D-FISTA, with 100 iterations to evaluate the case that the sparsity in the angle domain is exploited in the estimation.[3] The performance of the least squares (LS) estimator, where the BS knows the indices of non-zero rows of $\mathbf{H}$, and that of the CD method of [15], with pre-tuned ρ and 10 iterations, are denoted by "Exact LS" and "SPARROW," respectively. In comparison, the proposed CD method with update rule of [15], denoted by Proposed CD (SPARROW), can outperform BS-CAMP, and its performance can approach that of SPARROW, whereas the one using update rule of [3], denoted by Proposed CD (ML), is significantly superior to BS-CAMP and 2D-FISTA in terms of performance. Our results thus imply that the $\ell_{2,1}$ mixed-norm minimization is a more suitable approach for GF-NOMA systems with MIMO-OFDM, compared to classical schemes utilizing block-sparsity and 2D-CS. Furthermore, while our proposed method does not require hyperparameters, its performance approaches that of Exact LS.

Moreover, the NMSE performance of the proposed and conventional schemes for $N_\mathrm{p} = 64$ is shown in Fig. 3. As conventional schemes, we evaluate the performance of the schemes proposed in [3], *i.e.*, non-negative least square (NNLS) and ML, where the number of iterations is set to 10. These results show that the proposed scheme can outperform NNLS significantly and is comparable to ML. Note that our objective is to achieve superior performance without any hyperparameters that need to be designed via an exhaustive search to yield a low estimation error. In light of the above, our proposed receiver can meet such an objective and is comparable to the state-of-the-art receiver for GF-NOMA systems. Next, we demonstrate the accuracy of activity detection of the proposed scheme.

[3] Although the Lipschitz constant is calculated by $\max_{\|\mathbf{X}\|_F=1} \|\mathbf{A}^\mathrm{H}\mathbf{A}\mathbf{X}\|_F$ in [18], it is computed via $\|\mathbf{A}^\mathrm{H}\mathbf{A}\|_F$ based on the relation $\|\mathbf{A}^\mathrm{H}\mathbf{A}\mathbf{X}\|_F \leq \|\mathbf{A}^\mathrm{H}\mathbf{A}\|_F\|\mathbf{X}\|_F$ since $\|\mathbf{X}\|_F = 1$ is not usually satisfied in this letter.

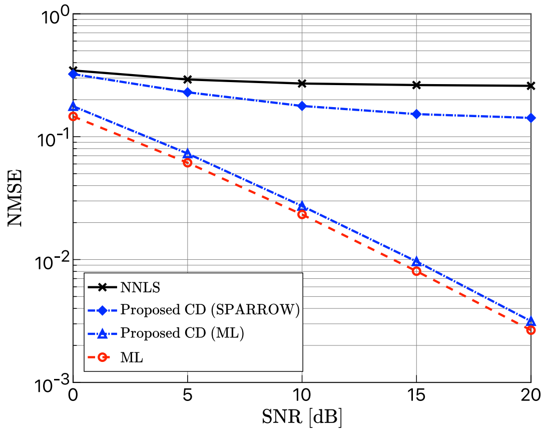

Fig. 3. NMSE performance of the proposed and conventional schemes for $N_{\mathrm{p}} = 64$.

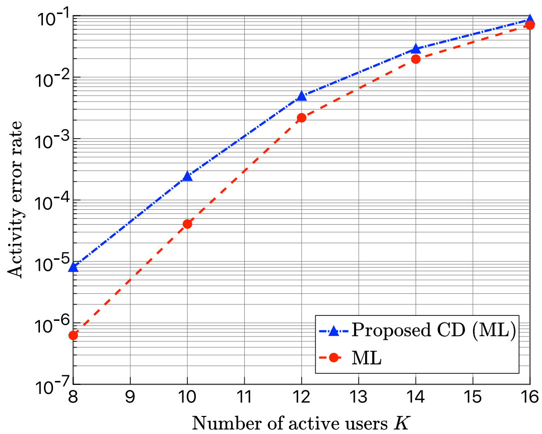

Fig. 4. AER of the proposed scheme for $N_{\mathrm{p}} = 64$ and SNR $= 5$ dB.

Fig. 4 represents the AER versus the number of active users K, where $N_{\mathrm{p}} = 64$ and SNR $= 5$ dB. As the activity detection depends on the accuracy of channel estimation, we focus on our proposal with update rule of [3] and ML. As seen from the figure, the accuracy of activity detection degrades as the number of active users increases. However, the AER of the proposed method can reach approximately 10^{-2}, even when 20% users are active, *i.e.*, $K/N = 13/64 \approx 0.2$. Although the conventional studies on GF-NOMA, *e.g.*, [2]–[4], consider the case where 10% of users or fewer are active, the proposed scheme can achieve low error rate even under tougher conditions. This indicates that our proposed scheme is robust against the variation of the traffic.

V. Conclusion

We introduce a new GF-NOMA spreading pilot sequence over the frequency domain to reduce the overhead of uplink transmissions. We also propose a hyperparameter-free receiver, which takes advantage of the sparsity of the channels in the delay domain. Unlike conventional schemes, our proposed method is an iterative algorithm requiring no pre-tuning of parameters to perform accurate estimation. Simulation results indicate that the proposed receiver can estimate the corresponding channels and the user-activity pattern with high precision in comparison with the classical algorithms utilizing a block-sparsity or a sparsity of channels in the angle domain.

References

[1] L. Liu, E. G. Larsson, W. Yu, P. Popovski, C. Stefanovic, and E. de Carvalho, "Sparse signal processing for grant-free massive connectivity: A future paradigm for random access protocols in the Internet of Things," *IEEE Signal Process. Mag.*, vol. 35, no. 5, pp. 88–99, Sep. 2018.

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

[3] S. Haghighatshoar, P. Jung, and G. Caire, "Improved scaling law for activity detection in massive MIMO systems," in *Proc. IEEE Int. Symp. Inf. Theory*, Vail, CO, USA, Jun. 2018, pp. 381–385.

[4] Y. Du *et al.*, "Joint channel estimation and multiuser detection for uplink grant-free NOMA," *IEEE Wireless Commun. Lett.*, vol. 7, no. 4, pp. 682–685, Aug. 2018.

[5] Z. Chen, F. Sohrabi, Y.-F. Liu, and W. Yu, "Covariance based joint activity and data detection for massive random access with massive MIMO," in *Proc. IEEE Int. Conf. Commun.*, Shanghai, China, May 2019, pp. 1–6.

[6] T. Hara and K. Ishibashi, "Grant-free non-orthogonal multiple access with multiple-antenna base station and its efficient receiver design," *IEEE Access*, vol. 7, pp. 175717–175726, 2019.

[7] M. Ke, Z. Gao, Y. Wu, X. Gao, and R. Schober, "Compressive sensing-based adaptive active user detection and channel estimation: Massive access meets massive MIMO," *IEEE Trans. Signal Process.*, vol. 68, pp. 764–779, Jan. 2020.

[8] V. K. Amalladinne, A. Vem, D. K. Soma, K. R. Narayanan, and J. Chamberland, "A coupled compressive sensing scheme for unsourced multiple access," in *Proc. IEEE Int. Conf. Acoust. Speech Signal Process.*, Calgary, AB, Canada, Apr. 2018, pp. 6628–6632.

[9] S. S. Kowshik, K. Andreev, A. Frolov, and Y. Polyanskiy, "Energy efficient random access for the quasi-static fading MAC," in *Proc. IEEE Int. Symp. Inf. Theory*, Paris, France, Jul. 2019, pp. 2768–2772.

[10] A. Vem, K. R. Narayanan, J. Chamberland, and J. Cheng, "A user-independent successive interference cancellation based coding scheme for the unsourced random access Gaussian channel," *IEEE Trans. Commun.*, vol. 67, no. 12, pp. 8258–8272, Dec. 2019.

[11] "Study on new radio (NR) access technology; Physical layer aspects, ver. 14.2.0, " 3rd Gener. Partnership Project, Sophia Antipolis, France, 3GPP Rep. TR 25.996, Sep. 2017.

[12] N. Y. Yu, K. Lee, and J. Choi, "Pilot signal design for compressive sensing based random access in machine-type communications," in *Proc. IEEE Wireless Commun. Netw. Conf.*, San Francisco, CA, USA, Mar. 2017, pp. 1–6.

[13] X. Lin, S. Wu, C. Jiang, L. Kuang, J. Yan, and L. Hanzo, "Estimation of broadband multiuser millimeter wave massive MIMO-OFDM channels by exploiting their sparse structure," *IEEE Trans. Wireless Commun.*, vol. 17, no. 6, pp. 3959–3973, Jun. 2018.

[14] X. Wu, L. Gu, W. Wang, and X. Gao, "Pilot design and AMP-based channel estimation for massive MIMO-OFDM uplink transmission," in *Proc. IEEE 27th Int. Symp. Pers. Indoor Mobile Radio Commun.*, Valencia, Spain, Sep. 2016, pp. 1–7.

[15] C. Steffens, M. Pesavento, and M. E. Pfetsch, "A compact formulation for $\ell_{2,1}$ mixed-norm minimization problem," *IEEE Trans. Signal Process.*, vol. 66, no. 6, pp. 1483–1497, Mar. 2018.

[16] Y. Zhou, M. Herdin, A. M. Sayeed, and E. Bonek, "Experimental study of MIMO channel statistics and capacity via the virtual channel representation," Dept. Electron. Commun. Eng., Univ. Wisconsin-Madison, Madison, WI, USA, Rep., Feb. 2007.

[17] C.-K. Wen, S. Jin, K.-K. Wong, J.-C. Chen, and P. Ting, "Channel estimation for massive MIMO using Gaussian-mixture Bayesian learning," *IEEE Trans. Wireless Commun.*, vol. 14, no. 3, pp. 1356–1368, Mar. 2015.

[18] S. Li, G. Zhao, W. Zhang, Q. Qiu, and H. Sun, "ISAR imaging by two-dimensional convex optimization-based compressive sensing," *IEEE Sensors J.*, vol. 16, no. 19, pp. 7088–7093, Oct. 2016.