Low-Complexity Equalization and Detection for OTFS-NOMA

Stephen McWade*, Mark F. Flanagan†, and Arman Farhang*
*Department of Electronic and Electrical Engineering, Trinity College Dublin, Ireland
†School of Electrical and Electronic Engineering, University College Dublin, Ireland
Email: {smcwade, arman.farhang}@tcd.ie, mark.flanagan@ieee.org

Abstract—Orthogonal time frequency space (OTFS) modulation has recently emerged as a potential 6G candidate waveform which provides improved performance in high-mobility scenarios. In this paper we investigate the combination of OTFS with non-orthogonal multiple access (NOMA). Existing equalization and detection methods for OTFS-NOMA, such as minimum-mean-squared error with successive interference cancellation (MMSE-SIC), suffer from poor performance and high complexity. Motivated by this, in this paper we propose a low-complexity method for equalization and detection for OTFS-NOMA. The proposed method uses an equalizer based on the least-squares with QR factorization (LSQR) algorithm and a novel reliability zone (RZ) detection scheme which estimates the reliable symbols of the users and then uses interference cancellation to remove multi-user interference (MUI). We present numerical results which demonstrate the superiority of our proposed method, in terms of symbol error rate (SER), to the existing MMSE-SIC benchmark scheme. Additionally, we present results which illustrate that a judicious choice of RZ thresholds is important for optimizing the performance of the proposed algorithm.

I. INTRODUCTION

The sixth generation (6G) of mobile networks is expected to support communications in high-mobility environments such as high-speed rail, vehicle-to-everything (V2X) and unmanned aerial vehicle (UAV) communications [1]. Orthogonal frequency division multiplexing (OFDM) has been the waveform utilized in the 4th and 5th generation of wireless networks. However, it is well-known that in high-mobility scenarios, OFDM performs poorly due to the Doppler effect [2]. In recent years, a new waveform called orthogonal time frequency space (OTFS) has been proposed to address this drawback of OFDM in time-varying channels. In contrast to OFDM, which transmits data symbols in the time-frequency domain, OTFS places the data symbols in the delay-Doppler domain [3]. OTFS uses a transformation to spread each information symbol over the time-frequency plane. This means that the symbols are all equally affected by the time and frequency selectivity of the channel which converts the time-varying channel to a time-invariant one in the delay-Doppler domain.

For a multi-user OTFS system, the multiple access (MA) technique utilized is an important consideration. The methods

proposed thus far in the literature can be broadly categorized as orthogonal multiple access (OMA) or non-orthogonal multiple access (NOMA). In OTFS-OMA, users are multiplexed either in the delay or the Doppler domain and only one user can occupy a given resource block [4]. However, the users suffer from multi-user interference (MUI) due to the Doppler spread. The resulting MUI can be mitigated by inserting guard bands between the users, as was done in [5]. However, this leads to a spectral efficiency (SE) loss [6].

An alternative approach is OTFS-NOMA, where the users are allowed to occupy the same resource block and are multiplexed in either the power domain or the code domain. A multi-user detection (MUD) scheme such as successive interference cancellation (SIC) is then used to detect the user symbols [7]. NOMA is a well-known technique which can provide improved SE compared to OMA. A number of OTFS-NOMA schemes have been proposed in the literature in recent years which use either power-domain [8], [9] or code-domain [10] multiplexing. In this paper, we focus on power-domain OTFS-NOMA.

With regard to the existing work on power-domain OTFS-NOMA, the authors of [8] considered a single high-mobility OTFS user multiplexed with multiple low-mobility OFDM users. However, this system model is restricted to a single OTFS user and hence cannot accommodate multiple high-mobility users. The authors of [9] addressed this issue and proposed an OTFS-NOMA scheme which utilizes a rectangular pulse shape where multiple users overlap in the delay-Doppler domain and are multiplexed in the power domain. However, the system proposed in [9] used minimum-mean-squared-error (MMSE) equalization in combination with SIC for equalization and detection. The problem with this scheme is that direct implementation of MMSE for equalization is prohibitively computationally complex and thus impractical for real-world scenarios.

As of yet, to the best of our knowledge, there is no low-complexity equalization and detection method for power-domain OTFS-NOMA. In addition, the low-complexity equalization and detection method of [11] for single-user OTFS is not directly applicable to a NOMA scenario due to the presence of MUI. In this paper, we address these gaps in the literature with the following contributions: (i) we propose a low-complexity receiver for a 2-user downlink OTFS-NOMA

system where both users deploy OTFS and are multiplexed in the power domain; (ii) we propose an iterative method for equalization and detection which uses the LSQR algorithm to equalize the channel and an RZ detector to detect reliable symbols from both users; (iii) our proposed method uses a novel symbol-wise interference cancellation process which considers MUI as well as inter-Doppler interference (IDI) and inter-symbol interference (ISI). We present numerical results which show that the proposed method provides symbol error rate (SER) performance gains of up to 6 dB over MMSE equalization with SIC. Additionally, we present numerical results which demonstrate that detection performance can be improved further by optimizing the choice of thresholds within the RZ detector.

The rest of this paper is organized as follows. Section II describes the system model for a 2-user OTFS-NOMA system. In Section III, the proposed equalization and detection algorithm is presented. Section IV presents numerical results. Finally, Section V concludes the paper.

Notations: Superscripts $(\cdot)^{\mathrm{T}}$ and $(\cdot)^{\mathrm{H}}$ denote transpose and Hermitian transpose, respectively. Bold lower-case characters are used to denote vectors and bold upper-case characters are used to denote matrices. $x[n]$ denotes the n-th element of the vector $\mathbf{x}$. The function $\mathrm{vec}(\mathbf{X})$ vectorizes the matrix $\mathbf{X}$ by stacking its columns to form a single column vector, and $\otimes$ represents the Kronecker product. The $p \times p$ identity matrix and $p \times q$ all-zero matrix are denoted by $\mathbf{I}_p$ and $\mathbf{0}_{p \times q}$, respectively. Finally, $j = \sqrt{-1}$ represents the imaginary unit.

II. SYSTEM MODEL

For ease of exposition, in the following sections we describe the system model and the proposed detector for the case of a 2-user OTFS-NOMA system; however, note that the proposed method is generalizable to any number of users. We consider a downlink OTFS-NOMA system where both users occupy the same delay-Doppler domain resources and are multiplexed in the power domain. For User $i \in \{1, 2\}$, let the $M \times N$ matrix $\mathbf{X}_i$ contain the MN quadrature amplitude modulation (QAM) data symbols placed in the delay-Doppler domain. The elements of $\mathbf{X}_i$ are assumed to be independent and identically distributed (i.i.d.) complex random variables. Additionally, a normalized (unit-energy) square QAM constellation is assumed for each user.

The delay-time domain transmit signal of User i can be expressed as $\mathbf{S}_i = \mathbf{A}_{\mathrm{cp}}\mathbf{X}_i\mathbf{F}_N^{\mathrm{H}}$, where $\mathbf{F}_N$ is the N-point unitary discrete Fourier transform (DFT) matrix in which the (l, k) element is $\frac{1}{\sqrt{N}}e^{-j\frac{2\pi}{N}lk}$ and $\mathbf{A}_{\mathrm{cp}} = [\mathbf{J}_{\mathrm{cp}}, \mathbf{I}_N]$ is the CP addition matrix (here $\mathbf{J}_{\mathrm{cp}}$ is composed of the last N_{cp} rows of $\mathbf{I}_N$). After parallel to serial conversion, the time-domain symbols for User i can be written as

$$\mathbf{s}_i = \mathrm{vec}(\mathbf{S}_i) = (\mathbf{F}_N^{\mathrm{H}} \otimes \mathbf{A}_{\mathrm{cp}})\mathbf{x}_i. \quad (1)$$

The users are then multiplexed in the power domain and their signals are superimposed before transmission. The superimposed transmit signal is given by $\mathbf{s} = \sqrt{p_1}\mathbf{s}_1 + \sqrt{p_2}\mathbf{s}_2$, where p_i is the power allocation coefficient for User i and

$p_1 + p_2 = 1$. After digital to analog conversion, the continuous-time signal $s(t)$ is transmitted through the linear time-varying (LTV) channel. The received signal at the receiver for User $i \in \{1, 2\}$ can be written as

$$r_i(t) = \int \int h_i(\tau, \nu)s(t - \tau)e^{j2\pi\nu(t-\tau)}d\tau d\nu + \omega_i(t), \quad (2)$$

where $h_i(\tau, \nu) = \sum_{p=0}^{P_i-1} h_{i,p}\delta(\tau - \tau_{i,p})\delta(\nu - \nu_{i,p})$, is the delay-Doppler channel impulse response (CIR) for User i, which consists of P_i channel paths, and $\omega_i(t)$ is the complex AWGN with variance σ_i^2. The parameters $h_{i,p}$, $\tau_{i,p}$ and $\nu_{i,p}$ represent the channel gain, delay and Doppler shift, respectively, associated with path p of User i's channel. The power delay profile (PDP) of the channel of User i is given by $\boldsymbol{\lambda}_i = [\lambda_i(0), \ldots, \lambda_i(0)]$ and is assumed to be normalized such that $\sum_{p=0}^{P_i-1} \lambda_i(p) = 1$. Each channel path gain is modeled as a complex Gaussian random variable with mean zero and variance $\lambda_i(p)$.

The received signal is then sampled with sampling period T_{s} and the discrete received signal samples can be expressed as

$$r_i[n] = \sum_{l=0}^{L-1} h_i[n, l]s[n - l] + \omega_i[n], \quad (3)$$

where $h_i[n, l]$ is the CIR for User i's channel at time instant n and delay l. The discrete-time received signal for User i can be written in matrix form as $\mathbf{r}_i = \mathbf{H}_i\mathbf{s} + \boldsymbol{\omega}_i$, where $\boldsymbol{\omega}_i$ is the $MN \times 1$ complex AWGN vector and $\mathbf{H}_i$ is the $MN \times MN$ time-domain channel matrix of User i constructed from the CIRs. The received signal is then demodulated and converted back to the delay-Doppler domain by performing an N-point DFT operation across the time-domain samples. Thus, the received signal is given by $\mathbf{y}_i = (\mathbf{F}_N \otimes \mathbf{R}_{\mathrm{cp}})\mathbf{r}_i$. This can alternatively be written as $\mathbf{y}_i = \mathbf{G}_i\mathbf{x}_{\mathrm{sup}} + \mathbf{w}_i$, where $\mathbf{G}_i = (\mathbf{F}_N \otimes \mathbf{R}_{\mathrm{cp}})\mathbf{H}_i(\mathbf{F}_N^{\mathrm{H}} \otimes \mathbf{A}_{\mathrm{cp}})$ is the effective channel matrix, $\mathbf{x}_{\mathrm{sup}} = \sqrt{p_1}\mathbf{x}_1 + \sqrt{p_2}\mathbf{x}_2$ is the superimposed delay-Doppler symbol vector and $\mathbf{w}_i = (\mathbf{F}_N \otimes \mathbf{R}_{\mathrm{cp}})\boldsymbol{\omega}_i$ is the noise vector. In this work, we assume that the users are ordered in ascending order of their average signal-to-noise ratio (SNR), i.e., $\mathrm{SNR}_1 < \mathrm{SNR}_2$, where $\mathrm{SNR}_i = \frac{p_1 + p_2}{\sigma_i^2}$.

III. PROPOSED EQUALIZATION AND DETECTION TECHNIQUE

The proposed method is inspired by the method proposed in [11] for single-user OTFS which utilized an iterative LSMR based method with RZ detection and interference cancellation. However, if the method in [11] is applied directly to OTFS-NOMA with SIC to detect the signals of User 1 and User 2, we can expect poor performance due to the MUI present in the system. Therefore, in the proposed method, we perform SIC at a symbol level rather than a packet level. This allows for the decoding of symbols from both users as soon as they become reliable and also allows for the incorporation of MUI cancellation to improve the detection performance. The proposed algorithm uses an iterative process in which the LSQR algorithm is used to equalize the channel and an RZ

Algorithm 1 LSQR Algorithm

1: **Input**: $\mathbf{G}$, $\mathbf{y}$ and σ^2
2: **Initialize**: $\mathbf{b} = \begin{pmatrix} \mathbf{y} \\ \mathbf{0} \end{pmatrix}$, $\mathbf{A} = \begin{pmatrix} \mathbf{G} \\ \sigma \mathbf{I} \end{pmatrix}$, $\beta_0 = \|\mathbf{b}\|$, $\mathbf{u}_0 = \mathbf{b}/\beta_0$,
   $\alpha_0 = \|\mathbf{A}^H \mathbf{u}_0\|$, $\mathbf{v}_0 = \mathbf{A}^H \mathbf{u}_0/\alpha_0$, $\mathbf{w}_0 = \mathbf{v}_0$, $\bar{\phi}_0 = \beta_0$,
   $\bar{\rho}_0 = \alpha_0$, $\mathbf{x}_0 = \mathbf{0}_{MN \times 1}$
3: **for** $u = 1 : U$ **do**
4:      $\beta_u = \|\mathbf{A}\mathbf{v}_{u-1} - \alpha_{u-1}\mathbf{u}_{u-1}\|$
5:      $\mathbf{u}_u = (\mathbf{A}\mathbf{v}_{u-1} - \alpha_{u-1}\mathbf{u}_{u-1})/\beta_u$
6:      $\alpha_u = \|\mathbf{A}^H \mathbf{u}_u - \beta_u \mathbf{v}_{u-1}\|$
7:      $\mathbf{v}_u = (\mathbf{A}^H \mathbf{u}_u - \beta_u \mathbf{v}_{u-1})/\alpha_u$
8:      $\rho_u = \|[\bar{\rho}_{u-1} \ \beta_u]\|$, $c_u = \frac{\bar{\rho}_{u-1}}{\rho_u}$, $s_u = \frac{\beta_u}{\rho_u}$
9:      $\theta_u = s_u \alpha_u$, $\phi_u = c_u \bar{\phi}_{u-1}$
10:     $\tau_u = \frac{\phi_u}{\rho_u}$, $\mu_u = \frac{\theta_u}{\rho_u}$
11:     $\bar{\phi}_u = -s_u \bar{\phi}_{u-1}$, $\bar{\rho}_u = -c_u \alpha_u$
12:     $\mathbf{x}_u = \mathbf{x}_{u-1} + \tau_u \mathbf{w}_{u-1}$
13:     $\mathbf{w}_u = \mathbf{v}_u - \mu_u \mathbf{w}_{u-1}$
14:     **if** $\|\mathbf{y} - \mathbf{G}\mathbf{x}_u\| \leq \epsilon$, break
15: **end for**
16: **Output**: $\tilde{\mathbf{x}} = \mathbf{x}_u$

Fig. 1. Illustration of the unreliable zone $\mathcal{V}_i(T)$ in the case where User i employs a 16-QAM constellation.

detector is used to detect the reliable symbols of both User 1 *and* User 2 within each iteration. Interference cancellation is then used to remove ISI, IDI *and* MUI from the undetected symbols of both users, which improves the detection quality in subsequent iterations. In the next subsection, we give a brief overview of the LSQR algorithm.

A. LSQR algorithm

The proposed method uses the LSQR algorithm, which is listed in Algorithm 1, to equalize the channel. LSQR is a well-known iterative algorithm for solving equalization problems of the form $\mathbf{y} = \mathbf{G}\mathbf{x} + \mathbf{w}$, where $\mathbf{x}$ is the transmitted vector, $\mathbf{y}$ is the received vector, $\mathbf{G}$ is the sparse channel matrix and $\mathbf{w}$ is the complex AWGN noise vector with variance per dimension σ^2 [12]. At iteration u, LSQR constructs the vector $\mathbf{x}_u$ in the Krylov subspace

$$\mathcal{K}(\mathbf{G}^H\mathbf{G}, \mathbf{G}^H\mathbf{y}, u) = \text{span}\{\mathbf{G}^H\mathbf{y}, (\mathbf{G}^H\mathbf{G})\mathbf{G}^H\mathbf{y}, \dots,$$
$$(\mathbf{G}^H\mathbf{G})^{u-1}\mathbf{G}^H\mathbf{y}\}$$

which minimizes the norm of the residual, $\|\mathbf{y} - \mathbf{G}\mathbf{x}_k\|$. At each iteration, the LSQR procedure consists of 2 stages: Golub-Kahan bidiagonalization and QR decomposition [12]. In lines $4-7$, the Golub-Kahan bidiagonalization process is used to form the vectors $\mathbf{u}_u$, $\mathbf{v}_u$ and the constants α_u, β_u which form the following bidiagonal least-squares problem

$$\min_{\mathbf{w}_u} \|\mathbf{B}_u \mathbf{w}_u - [\beta_0, 0, 0, \dots, 0]^T\|,$$

where $\mathbf{B}_u$ is a $(u+1) \times u$ matrix with elements $[\alpha_0, \alpha_1, \dots, \alpha_{u-1}]$ on the main diagonal and $[\beta_1, \beta_2, \dots, \beta_{u-1}]$ on the first subdiagonal and β_0 is initialized as $\|\mathbf{y}\|$. This bidiagonal least-squares problem can then be solved at low computational cost using the QR factorization of $\mathbf{B}_u$. Lines $8-13$ of Algorithm 1 show
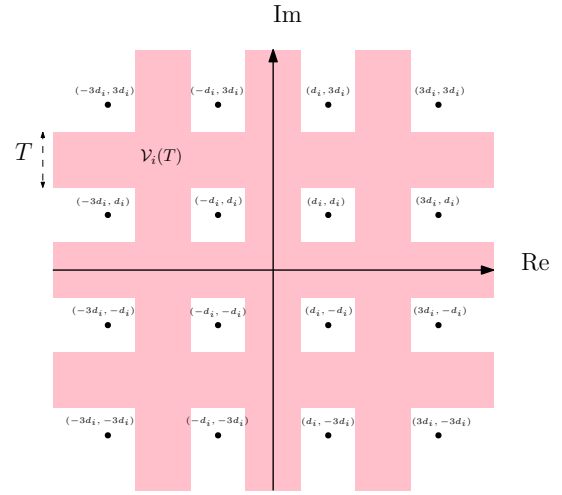
a simple recursive method for obtaining this solution as outlined in [13]. The iterative process continues until either the norm of the residual reaches a pre-determined tolerance, ϵ, or the maximum number of iterations U is reached.

LSQR can also be regularized by including the noise variance per dimension σ^2 as a damping parameter to improve the semi-convergence property of the algorithm. After several iterations, LSQR provides performance similar to MMSE but with lower complexity [12]. Once LSQR is used to obtain the estimate of the transmitted signal, the proposed algorithm uses an RZ detector to obtain the reliable symbols of each user. In the next subsection, we describe the RZ detector.

B. Reliability zone detector

Here, we first introduce some relevant notation. Each User $i \in \{1, 2\}$ uses A_i-ary QAM modulation, where the QAM symbol constellation is defined as

$$\mathcal{A}_i = \{u + vj \ : \ u, v \in \{(2a-1)d_i :$$
$$a \in \{-\sqrt{A_i}/2 + 1, \ \dots, \ \sqrt{A_i}/2\}\}\},$$

where d_i is half the distance between adjacent QAM constellation symbols of User i. Next, we define the *unreliable* zone with respect to this QAM constellation as

$$\mathcal{V}_i(T) = \{u + vj \ | u, v \in \mathcal{U}_i(T)\}, \tag{4}$$

where

$$\mathcal{U}_i(T) = \bigcup_{a=-\sqrt{A_i}/2+1}^{\sqrt{A_i}/2-1} \mathcal{U}_{i,a}(T) \tag{5}$$

and

$$\mathcal{U}_{i,a}(T) = \{u \ | \ 2ad_i - T/2 < u < 2ad_i + T/2\},$$

where T is a pre-defined threshold which determines the size of the unreliable zone. To demonstrate, the shaded areas in Fig. 1 shows an illustration of the unreliable zone $\mathcal{V}_i(T)$ for a 16-QAM constellation.

In the detection process, decisions are made in a symbol-by-symbol manner. If a symbol $x_i[n]$ is outside $\mathcal{V}_i(T_i)$, then it is deemed reliable and can be quantized to the nearest symbol in $\mathcal{A}_i$. If $x_i[n]$ is inside $\mathcal{V}_i(T_i)$ then it is deemed unreliable and no quantization takes place. The detected reliable symbols can then be used for interference cancellation.

C. Proposed equalization and detection algorithm

In this subsection, we describe the proposed algorithm for equalization and detection of the OTFS-NOMA signal at User i's receiver - this is detailed in Algorithm 2. Each iteration of the proposed method begins by using the LSQR algorithm to equalize the channel and obtain a new estimate of the superimposed transmitted symbol vector via

$$\tilde{\mathbf{x}}_{\text{sup}} = \text{LSQR}(\mathbf{G}_i, \mathbf{y}^{(k)}), \tag{6}$$

which can be seen on line 5 of Algorithm 2. Next, two sub-vectors are formed from $\tilde{\mathbf{x}}_{\text{sup}}$, which can be seen on lines 6 and 7. The vector $\tilde{\mathbf{x}}_1$ containing the elements of $\tilde{\mathbf{x}}_{\text{sup}}$ whose indices are in $\mathcal{N}_1$, which is the set of undetected User 1 symbols, is given by

$$\tilde{\mathbf{x}}_1 = ((\tilde{\mathbf{x}}_{\text{sup}}[m])_{m \in \mathcal{N}_1})/\sqrt{p_1}. \tag{7}$$

Since the RZ detector can only make decisions on User 2 symbols once the corresponding User 1 symbols have been detected on a previous iteration, the vector $\tilde{\mathbf{x}}_2$ contains the elements of $\tilde{\mathbf{x}}_{\text{sup}}$ whose indices are in $\mathcal{N}_2$, the set of undetected User 2 symbols, *and* $\mathcal{D}_1$, the set of detected User 1 symbols. Therefore, $\tilde{\mathbf{x}}_2$ is given by

$$\tilde{\mathbf{x}}_2 = ((\tilde{\mathbf{x}}_{\text{sup}}[m])_{m \in \{\mathcal{N}_2 \cup \mathcal{D}_1\}})/\sqrt{p_2}. \tag{8}$$

The RZ detector identifies the reliable symbols of both users, i.e., those on which we will make a decision in this iteration. The index sets of reliable symbols for each users, $\mathcal{R}_1$ and $\mathcal{R}_2$, are given by

$$\mathcal{R}_j = \{n \in \mathcal{N}_j : \tilde{\mathbf{x}}_j[n] \notin \mathcal{V}_j(T_j^{(k)})\}, \forall j \in \{1, 2\}, \tag{9}$$

which can be seen on line 8. After the reliable symbols are identified, they are quantized to the nearest QAM symbol and are stored in the empty vectors $\tilde{\mathbf{x}}_{1,\text{q}}$ and $\tilde{\mathbf{x}}_{2,\text{q}}$ via

$$\tilde{\mathbf{x}}_{j,\text{q}}[r] = Q_j(\tilde{\mathbf{x}}_j[r]), \ \forall r \in \mathcal{R}_j, \forall j \in \{1, 2\}, \tag{10}$$

where the function $Q_j(\tilde{\mathbf{x}}_j[r])$ quantizes the reliable symbol $x_i[n]$ to the nearest symbol in $\mathcal{A}_j$ (line 9). No decision is made on the unreliable symbols and they are not quantized.

The quantized reliable symbols are used to remove interference from the received signal vector via

$$\mathbf{y}^{(k+1)} = \mathbf{y}^{(k)} - \mathbf{G}_i(\sqrt{p_1}\tilde{\mathbf{x}}_{1,\text{q}} + \sqrt{p_2}\tilde{\mathbf{x}}_{2,\text{q}}). \tag{11}$$

The quantized symbols are also stored in the estimated symbol vectors $\hat{\mathbf{x}}_1$ and $\hat{\mathbf{x}}_2$ (line 11). After canceling the interference from the detected symbols of both users, the algorithm updates the sets $\mathcal{N}_1$ and $\mathcal{N}_2$ of undetected symbols, and the set $\mathcal{D}_1$ of detected User 1 symbols based on the state of the output vectors $\hat{\mathbf{x}}_1$ and $\hat{\mathbf{x}}_2$. The thresholds for the RZ detector for

---

Algorithm 2 Proposed Algorithm for symbol detection at User i receiver

1: **Input**: User index i, Channel matrix $\mathbf{G}_i$, received symbol vector $\mathbf{y}_i$, power allocation fractions p_1 and p_2, starting thresholds $T_1^{(1)}$ and $T_2^{(1)}$
2: **Initialize**: $\mathbf{y}^{(1)} = \mathbf{y}_i$, $\hat{\mathbf{x}}_1 = \hat{\mathbf{x}}_2 = \tilde{\mathbf{x}}_{1,\text{q}} = \tilde{\mathbf{x}}_{2,\text{q}} = \mathbf{0}_{MN \times 1}$
3: Define $\mathcal{N} = \{0, \ldots, MN - 1\}$, $\mathcal{N}_1 = \mathcal{N}_2 = \mathcal{N}$, $\mathcal{D}_1 = \emptyset$
4: **for** $k = 1$ to K **do**
5:     $\tilde{\mathbf{x}}_{\text{sup}} = \text{LSQR}(\mathbf{G}_i, \mathbf{y}^{(k)})$
6:     $\hat{\mathbf{x}}_1 = ((\tilde{\mathbf{x}}_{\text{sup}}[m])_{m \in \mathcal{N}_1})/\sqrt{p_1}$
7:     $\hat{\mathbf{x}}_2 = ((\tilde{\mathbf{x}}_{\text{sup}}[m])_{m \in \{\mathcal{N}_2 \cup \mathcal{D}_1\}})/\sqrt{p_2}$
8:     Update users' reliable symbol index sets via

    $\mathcal{R}_j = \{n \in \mathcal{N}_j : \tilde{\mathbf{x}}_j[n] \notin \mathcal{V}_j(T_j^{(k)})\}, \forall j \in \{1, 2\}$

9:     Quantize reliable symbols:

    $\tilde{\mathbf{x}}_{j,\text{q}}[r] = Q_j(\tilde{\mathbf{x}}_j[r]), \ \forall r \in \mathcal{R}_j, \forall j \in \{1, 2\}$

10:    Remove interference:

    $\mathbf{y}^{(k+1)} = \mathbf{y}^{(k)} - \mathbf{G}_i(\sqrt{p_1}\tilde{\mathbf{x}}_{1,\text{q}} + \sqrt{p_2}\tilde{\mathbf{x}}_{2,\text{q}})$

11:    Store detected symbols in output vectors:

    $\hat{\mathbf{x}}_j = \tilde{\mathbf{x}}_{j,\text{q}}[r], \ \forall r \in \mathcal{R}_j, \forall j \in \{1, 2\}$

12:    Reset: $\tilde{\mathbf{x}}_{1,\text{q}} = \mathbf{0}$ and $\tilde{\mathbf{x}}_{2,\text{q}} = \mathbf{0}$
13:    Update: $\mathcal{N}_1 = \{n \in \mathcal{N} : \hat{x}_1[n] = 0\}$, $\mathcal{N}_2 = \{n \in \mathcal{N} : \hat{x}_2[n] = 0\}$, $\mathcal{D}_1 = \{n \in \mathcal{N} : n \notin \mathcal{N}_1\}$
14:    Update thresholds: $T_1^{(k+1)} = T_1^{(1)}(1 - \frac{k+1}{K-1})$, and $T_2^{(k+1)} = T_2^{(1)}(1 - \frac{k+1}{K})$
15:    **if** $\mathcal{N}_i = \emptyset$, **break**
16: **end for**
17: **Output**: $\hat{\mathbf{x}}_i$

---

each user are then reduced geometrically within each iteration in line 14. This ensures that the algorithm converges within K iterations and that all User i symbols are detected. Since this is at User i's receiver, the algorithm stops when all of the User i symbols are detected, i.e., User 1 will detect all of its symbols before it detects all the User 2 symbols and can therefore stop once $\hat{\mathbf{x}}_1$ has no entries equal to zero.

D. Computational complexity

In this subsection, the computational complexity of the proposed method is compared to that of the MMSE-SIC benchmark, in terms of the number of complex multiplications. Direct implementation of MMSE equalization involves the inversion of an $MN \times MN$ matrix and hence has a computational complexity of $\mathcal{O}(M^3 N^3)$. Each iteration of the LSQR algorithm has a computational complexity of $\mathcal{O}(MN \log_2(MN))$ [12]. In the worst-case scenario Algorithm 1 performs LSQR K times, each with U LSQR iterations. Therefore, the computational complexity of Algorithm 1 is $\mathcal{O}(UKMN \log_2(MN))$. In practice, the typical values of K and U are in the order of tens and the typical value of MN is in the order of thousands. Thus, $UK << M^2 N^2$ and
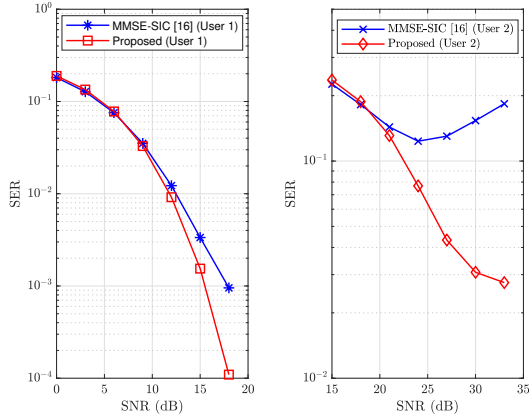
Fig. 2. SER performance comparison between the proposed detector and a detector using MMSE-SIC, with different SNR levels, for the case where each user is allocated a 4-QAM constellation.
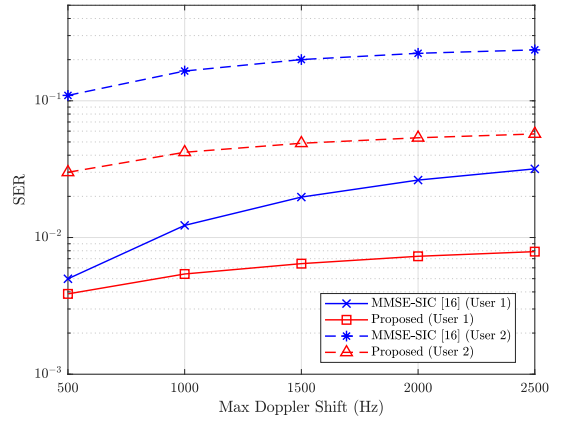

Fig. 4. SER performance comparison between the proposed detector and a detector using MMSE-SIC, with different maximum Doppler shifts, for the case where each user is allocated a 16-QAM constellation.
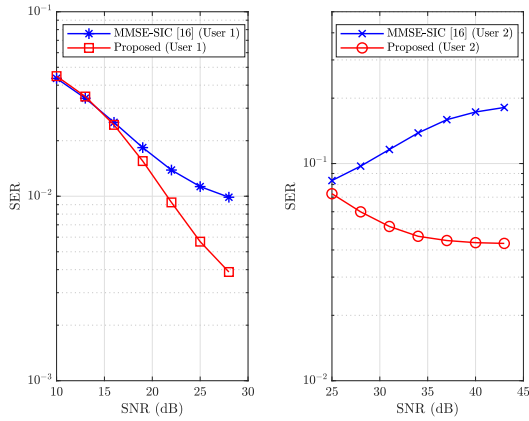

Fig. 3. SER performance comparison between the proposed detector and a detector using MMSE-SIC, with different SNR levels, for the case where each user is allocated a 16-QAM constellation.

Fig. 5. SER for User 2 using Algorithm 1 for different values of the User 1 starting threshold.

our method can achieve orders of magnitude computational complexity improvement over MMSE-SIC for OTFS-NOMA.

IV. NUMERICAL RESULTS AND DISCUSSION

This section presents numerical results to showcase the effectiveness of the proposed OTFS-NOMA equalization and detection algorithm. As a benchmark, an OTFS-NOMA system using MMSE equalization and SIC for detection is considered, which is referred to as MMSE-SIC. Monte Carlo simulation is used to average the results over 10^5 random channel realizations. A carrier frequency of $f_c = 5.9$ GHz, a transmission bandwidth of 4.95 MHz and a delay-Doppler grid size of $M = 64$ and $N = 16$ are considered. For power allocation, we use the average-SNR-based fractional transmit power allocation (FTPA) scheme outlined in [9]. Additionally, we consider a fixed SNR difference of 15 dB between the users, i.e., User 2 has an average SNR that is 15 dB higher than that of User 1. For each user, we consider a starting threshold
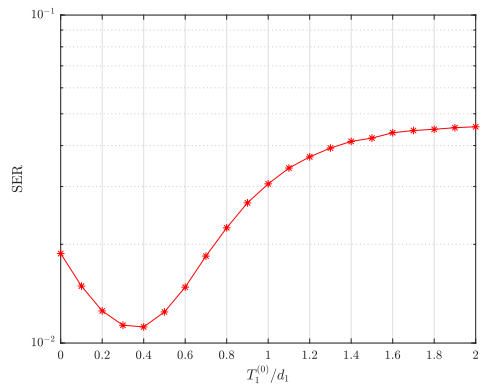
of $T_i^{(1)} = 2d_i$. The Tapped Delay Line C (TDL-C) model with a delay spread of 300 ns [14] is used for the channel model. We consider a range of maximum Doppler shifts from 500 Hz to 2500 Hz, which corresponds to velocities of approximately 90 km/h to 450 km/h at a carrier frequency of 5.9 GHz. The Doppler shifts are generated using Jakes' model [15]. For the LSQR algorithm, a maximum number of iterations of $U = 15$ and a tolerance of $\epsilon = 10^{-2}$ are considered, which are commonly used values for LSQR in related literature [11], [16]. For Algorithm 1, we consider a maximum number of iterations of $K = 10$ to limit the computational complexity.

Fig. 2 shows the SER of both users under the proposed equalization and detection method, compared to the benchmark MMSE-SIC scheme, for different signal-to-noise ratio (SNR) conditions. For these simulations each user's symbols are taken from a 4-QAM constellation, i.e., $\mathcal{A}_1 = \mathcal{A}_2 = 4$ and the user velocity is fixed at 200 km/h, which equates to a maximum Doppler shift of approximately 1000 Hz. It can be seen from Fig. 2 that for both users, the proposed method outperforms the MMSE-SIC method, especially at high SNR.

For User 1, the proposed method provides an SNR gain of 3 dB at an SER of 10^{-3} over MMSE-SIC. For User 2, the proposed method provides significant performance gains as MMSE-SIC performs very poorly, especially at high SNRs. This is due to the fact that, as the SNR increases, MMSE equalization becomes closer to zero-forcing equalization and the interference is amplified by the matrix inversion.

Fig. 3 shows the SER of both users under the proposed equalization and detection method, compared to the benchmark MMSE-SIC scheme, for different SNR conditions, for the case where each user's symbols are taken from a 16-QAM constellation. As with the 4-QAM results, the user velocity is fixed at 200 km/h. It can be seen from Fig. 3 that for both users the proposed method outperforms the MMSE-SIC method also for the 16-QAM scenario. For User 1, the proposed method provides an SNR gain of 6 dB at an SER of 10^{-2} over MMSE-SIC. For User 2, the proposed method provides significant performance gains over MMSE-SIC, which once again performs poorly due to the noise amplification caused by the matrix inversion.

Fig. 4 shows the SER of both users under the proposed equalization and detection method, compared to the benchmark MMSE-SIC scheme, for different values of maximum Doppler shift, for the case where each user's symbols are taken from a 16-QAM constellation. Fixed SNRs of 20 dB and 35 dB are considered for User 1 and 2, respectively. It can be seen that the performance gains of the proposed method over MMSE-SIC actually improves in high Doppler environments, as the performance of MMSE-SIC deteriorates significantly at higher maximum Doppler shifts. This is because as the Doppler spread increases, the channel matrix is more likely to be ill-conditioned [11]; hence, the matrix inversion involved in MMSE equalization may not be robust and can introduce significant equalization error.

The performance of the proposed algorithm is dependent on the effectiveness of the interference cancellation process which is in turn dependent on the thresholds of the RZ detector. With this in mind, we have studied the effect of the choice of starting threshold on the performance of our proposed method. To demonstrate this, Fig. 5 shows the SER of User 2 as a function of the starting User 1 threshold $T_1^{(1)}$, normalized by d_1. The results in this figure are for a fixed SNR of 35 dB and a velocity of 200 km/h. It can be seen from Fig. 5 that the User 1 threshold has an effect on the detection performance of User 2 detection, and that, somewhat counter-intuitively, a tighter (conservative) User 1 threshold degrades the performance of User 2. This is because using tight starting thresholds means that fewer User 1 symbols are detected during early iterations and their MUI is still present in the system when the User 2 symbols are being detected. This highlights that the overall detection performance for both users' symbols is sensitive to the choice of the initial RZ thresholds, and that these should be chosen carefully to obtain the best performance.

V. Conclusion

This paper has presented a novel low-complexity receiver for downlink OTFS-NOMA. The proposed method uses an iterative process which deploys the LSQR algorithm to equalize the channel, RZ detection to detect symbols from all users within the iterations and interference cancellation to remove MUI as well as IDI and ISI. Numerical results demonstrate the superiority of the proposed method, in terms of SER performance, with respect to an MMSE-SIC benchmark scheme. Furthermore, it has been demonstrated that the performance of the proposed algorithm is sensitive to the choice of starting thresholds, which need to be designed carefully. This motivates future work on developing a systematic approach to choosing these starting thresholds.

References

[1] H. Tataria, M. Shafi, A. F. Molisch, M. Dohler, H. Sjöland, and F. Tufvesson, "6G Wireless Systems: Vision, Requirements, Challenges, Insights, and Opportunities," *Proceedings of the IEEE*, vol. 109, no. 7, pp. 1166–1199, 2021.

[2] Z. Wei, W. Yuan, S. Li, J. Yuan, G. Bharatula, R. Hadani, and L. Hanzo, "Orthogonal Time-Frequency Space Modulation: A Promising Next-Generation Waveform," *IEEE Wireless Communications*, vol. 28, no. 4, pp. 136–144, 2021.

[3] R. Hadani, S. Rakib, M. Tsatsanis, A. Monk, A. J. Goldsmith, A. F. Molisch, and R. Calderbank, "Orthogonal Time Frequency Space Modulation," in *2017 IEEE Wireless Communications and Networking Conference (WCNC)*, 2017, pp. 1–6.

[4] R. Chong, S. Li, J. Yuan, and D. W. K. Ng, "Achievable Rate Upper-Bounds of Uplink Multiuser OTFS Transmissions," *IEEE Wireless Communications Letters*, vol. 11, no. 4, pp. 791–795, 2022.

[5] H. R. Rakib Shlomo, "Multiple access in wireless telecommunications system for high-mobility applications," August 2017.

[6] G. D. Surabhi, R. M. Augustine, and A. Chockalingam, "Multiple Access in the Delay-Doppler Domain using OTFS modulation," 2019.

[7] L. Dai, B. Wang, Z. Ding, Z. Wang, S. Chen, and L. Hanzo, "A Survey of Non-Orthogonal Multiple Access for 5G," *IEEE Communications Surveys Tutorials*, vol. 20, no. 3, pp. 2294–2323, 2018.

[8] Z. Ding, R. Schober, P. Fan, and H. Vincent Poor, "OTFS-NOMA: An Efficient Approach for Exploiting Heterogenous User Mobility Profiles," *IEEE Transactions on Communications*, vol. 67, no. 11, pp. 7950–7965, 2019.

[9] A. Chatterjee, V. Rangamgari, S. Tiwari, and S. S. Das, "Nonorthogonal Multiple Access With Orthogonal Time–Frequency Space Signal Transmission," *IEEE Systems Journal*, vol. 15, no. 1, pp. 383–394, 2021.

[10] K. Deka, A. Thomas, and S. Sharma, "OTFS-SCMA: A Code-Domain NOMA Approach for Orthogonal Time Frequency Space Modulation," *IEEE Transactions on Communications*, vol. 69, no. 8, pp. 5043–5058, 2021.

[11] H. Qu, G. Liu, L. Zhang, S. Wen, and M. A. Imran, "Low-Complexity Symbol Detection and Interference Cancellation for OTFS System," *IEEE Transactions on Communications*, vol. 69, no. 3, pp. 1524–1537, 2021.

[12] T. Hrycak, S. Das, G. Matz, and H. G. Feichtinger, "Low Complexity Equalization for Doubly Selective Channels Modeled by a Basis Expansion," *IEEE Transactions on Signal Processing*, vol. 58, no. 11, pp. 5706–5719, 2010.

[13] C. C. Paige and M. A. Saunders, "LSQR: An Algorithm for Sparse Linear Equations and Sparse Least Squares," *ACM Trans. Math. Softw.*, vol. 8, no. 1, p. 43–71, mar 1982.

[14] *3GPP TS 38.901*, 3rd Generation Partnership Project (3GPP), June 2018, v15.0.0.

[15] C. Xiao, Y. Zheng, and N. Beaulieu, "Second-order statistical properties of the WSS Jakes' fading channe simulator," *Communications, IEEE Transactions on*, vol. 50, pp. 888 – 891, 07 2002.

[16] H. Qu, G. Liu, Y. Wang, Q. Chen, C. Yi, and J. Peng, "A Time-Domain Approach to Channel Estimation and Equalization for the SC-FDM System," *IEEE Transactions on Broadcasting*, vol. 65, no. 4, pp. 713–726, 2019.