Generalized Quadrature Spatial Modulation and Its Application to Vehicular Networks with NOMA

Jun Li, Member, IEEE, Shuping Dang, Member, IEEE, Yier Yan, Yuyang Peng, Member, IEEE, Saba Al-Rubaye, Senior Member, IEEE, and Antonios Tsourdos

Abstract—Quadrature spatial modulation (QSM) is recently proposed to increase the spectral efficiency (SE) of SM, which extends the transmitted symbols into in-phase and quadrature domains. In this paper, we propose a generalized QSM (GQSM) scheme to further increase the SE of QSM by activating more than one transmit antenna in in-phase or quadrature domain. A low-complexity detection scheme for GQSM is provided to mitigate the detection burden of the optimal maximum-likelihood (ML) detection method. An upper bounded bit error rate is analyzed to discover the system performance of GQSM. Moreover, by collaborating with the non-orthogonal multiple access (NOMA) technique, we investigate the practical application of GQSM to cooperative vehicular networks and propose the cooperative GQSM with OMA (C-OMA-GQSM) and cooperative GQSM with NOMA (C-NOMA-GQSM) schemes. Computer simulation results verify the reliability of the proposed low-complexity detection as well as the theoretical analysis, and show that GQSM outperforms QSM in the entire SNR region. The superior BER performance of the proposed C-NOMA-GQSM scheme make it a promising modulation candidate for next generation vehicular networks.

Index Terms—Vehicular communications, spatial modulation, index modulation, non-orthogonal multiple access, bit error rate.

I. INTRODUCTION

Index modulation (IM) technique utilizes index(es) of transmit antennas, receive antennas, time slots, subcarriers, linear block codes, or other media to transmit additional information without consuming any additional resource [1]–[3]. Due to the good balance between the spectral efficiency (SE) and energy efficiency (EE), IM technique is considered as a promising candidate for the future sixth generation (6G) networks [4], [5].

The IM concept was first applied into the frequency domain, which is incorporated in orthogonal frequency division multiplexing (OFDM) systems. Subcarrier-index modulation OFDM (SIM-OFDM) was proposed to transmit index bits by subcarriers, which suffers from error propagation influence [6]. To mitigate this problem, enhanced SIM-OFDM (ESIM-OFDM) was then proposed to use one index bit to control two consecutive subcarriers, which activates only one subcarrier in each subcarrier pair [7]. However, compared with the conventional OFDM, the SIM-OFDM and ESIM-OFDM schemes achieve a much smaller SE. Therefore, the OFDM with IM (OFDM-IM) scheme was proposed to solve this problem in [8], which allows to convey more than one index bit. By OFDM-IM, the information bits are divided into two parts, as index bits and modulation bits, where a group of subcarriers is activated to transmit the modulation bits, and index bits are carried by subcarrier activation patterns (SAPs). Note that the mapping between index bits and SAPs is decided by several existing mapping methods, e.g., the combinatorial method in [8] and the equiprobable SAP method in [9]. It is shown in [8] that OFDM-IM outperforms OFDM under the identical SE. Due to this virtue, a lot of research studies of OFDM-IM have been carried out in recent years. For instance, the bit error rate (BER) performance was improved by subcarrier interleaving strategy in [10], [11], and the achievable rate of OFDM-IM was enhanced by optimizing SAPs in [12]. Because of the mapping rule, some SAPs are discarded so as to degrade the system performance. Therefore, the full utilization method of the full set of SAPs was investigated to improve the system performance of OFDM-IM in [13]. Additionally, OFDM-IM was extended to practical applications, e.g., vehicular networks [14] and cognitive radio networks [15]. Besides, the improvement of SE of OFDM-IM has also been studied by many researchers. The OFDM with generalized index modulation (OFDM-GIM) scheme was designed to encode more index bits of OFDM-IM in [16]. In [15], layered OFDM-IM was proposed to increase the SE of OFDM-IM by extending index bits to the layer domain. Moreover, the multiple-input and multiple-output (MIMO)-OFDM-IM scheme was proposed by incorporating OFDM-IM into the conventional MIMO architecture to considerably improve SE of OFDM-IM in [18].

It is observed that the above OFDM-IM related schemes only use the indices of active subcarriers to carry index bits. However, in order to further increase the SE, the generalized dual-mode IM aided OFDM (GDM-OFDM) was designed to transmit the modulated symbols through both active and inactive subcarriers in [19]. Similarly, in [20], [21], (generalized)-multiple-mode OFDM-IM ((G-))MM-OFDM-IM) was designed to transmit the modulated symbols through all subcarriers and the additional index bits are conveyed by the permutation order of distinguished constellations. It should be noted that although the SE improves, the computational complexity for signal detection is greatly increased, especially

This work was supported in part by National Nature Science Foundation of China under Grants 61701127, and in part by The Science and Technology Development Fund, Macau SAR (0013/2019/A). (Corresponding author: Yier Yan.)

J. Li and Y. Yan are with the School of Mechanical and Electrical Engineering, Guangzhou University, Guangzhou 510006, China (e-mail: {lijum52018, yeary080}@gzhu.edu.cn).

S. Dang is with Computer, Electrical and Mathematical Science and Engineering Division, King Abdullah University of Science and Technology (KAUST), Thuwal 23955-6900, Kingdom of Saudi Arabia (e-mail: shuping.dang@kaust.edu.sa).

Y. Peng is with the Faculty of Information Technology, Macau University of Science and Technology, Taipa 999078, Macau (e-mail: yypeng@must.edu.mo).

S. Al-Rubaye and A. Tsourdos are with the School of Aerospace, Transport and Manufacturing, Cranfield University, Bedford MK43 0AL, U.K. (e-mail: {s.alrubaye, a.tsourdos}@cranfield.ac.uk).
when the number of subcarriers is large. Therefore, to combat this effect, the low-complexity detection methods for OFDM-IM and related schemes were investigated in [22]–[25].

Recently, spatial modulation (SM), which applies the IM concept into the space domain, is considered as a special case of the MIMO architecture [26]–[28]. Similar to OFDM-IM, the information bits in SM are also split into index bits and modulation bits. By SM, one single transmit antenna is activated to transmit one modulated symbol generated by the modulation bits, and the index bits are conveyed by the index of the active transmit antenna. Because of the activation of a single transmit antenna, SM has appealing features, such as the high EE and free inter-channel interference (ICI). To obtain the higher EE of SM, space shift keying (SSK), which can be considered as a special case of SM, was proposed to only convey the index of the activated transmit antenna in [29]. However, every coin has two sides. SM(SSK) obtains the high EE in exchange for sacrificing its SE. In order to figure out the SE problem of SM(SSK), lots of literature has been proposed in [30]–[32]. The generalized SSK (GSSK) and generalized SM (GSM) schemes were proposed to increase the SE of SSSK and SM by allowing to transmit multiple transmit antennas in [30], [31], respectively. Relying on expanding index bits into the in-phase/quadrature domain, quadrature SM (QSM) was proposed to achieve the doubled length of index bits than SM in [32] and [33]. Moreover, the SM/SSK scheme was extended to the space-time domain to improve its BER performance, e.g. space-time shift keying (STS) in [34], space-time block coded SM (STBC-SM) in [35], [36], and differential SM (DMS) in [37], [38]. It should be noted that all above SM related works only utilize the index(ices) of transmit antennas. Differently, (generalised) pre-coding SM ((G)PSM) was proposed to transmit the additional index bits through the index(es) of receive antennas in [39], [40]. To further explore its application, GQSM was extended to the multi-stream and in-phase/quadrature systems in [41] and [42], respectively.

Inspired by QSM, we propose a novel SM scheme, termed generalized QSM (GQSM), to increase the SE of the conventional QSM scheme. To ease signal detection, a near-maximum-likelihood (ML) low-complexity detection is proposed to relieve high detection complexity of the optimal ML detector. The upper bound on the BER performance and the diversity of GQSM are also derived. In addition, we further study the practical application of GQSM into the downlink multi-vehicle relaying network by incorporating the non-orthogonal multiple access (NOMA) technique. Note that, unlike the conventional OMA technique, which serves users via orthogonal resources, NOMA is considered as the promising technique for 6G networks, which splits multiple users in different power levels and simultaneously serve all users at the same time/frequency/space resource [43]–[46]. Specifically, cooperative QSM with OMA (C-OMA-QSM) and cooperative QSM with NOMA (C-NOMA-QSM) are proposed to investigate the combination of NOMA (OMA) and GQSM in vehicular networks. In multi-vehicle communication systems, one base station (BS) is set to serve two vehicles (vehicle 1 and vehicle 2), where vehicle 1 is located near the BS and vehicle 2 is located far from the BS, which needs one decode-and-forward (DF) relay to assist its transmission. In C-OMA-QSM, the BS separately transmits signals to two vehicles in two different phases, which results in very poor performance of vehicle 1 but excellent performance of vehicle 2. Unlike C-OMA-QSM, C-NOMA-QSM combines two vehicles’ signals with different power levels and transmits to two vehicles simultaneously. Simulation results show that the QSM scheme outperforms the conventional QSM scheme, and the proposed low-complexity detection produces the near-ML BER performance. It can also be shown that vehicle 1 (vehicle 2) in C-OMA-QSM obtains very poor (excellent) BER performance while both vehicle 1 and vehicle 2 in C-NOMA-QSM achieve comparable BER performance, which is more suitable for practical applications that consider vehicle fairness.

The rest of this paper is organized as follows. The system model, low-complexity detection method, and the upper bound on the BER of QSM are proposed in Section II. The application of GQSM to OMA/NOMA is investigated in Section III. Simulation results are discussed in Section IV. Finally, this paper is concluded in Section V.

Notations: Upper and lower case boldface letters denote matrices and column vectors, respectively. The complex number field is represented by \( \mathbb{C} \). \( (\cdot)^T \) and \( (\cdot)^H \) represent the transpose and Hermitian transpose operations, respectively. \( \mathbf{I}_M \) is an identity matrix of size \( M \times M \). \( \| \cdot \| \) and \( C(\cdot, \cdot) \) denote the Frobenius norm and binomial operations, respectively. \( \mathcal{R} \{ \cdot \} \) and \( \mathcal{I} \{ \cdot \} \) represent the real and imaginary parts of the argument, respectively. \( Q(\cdot) \) represents the Gaussian Q-function [47]. \( \text{sort}(\cdot) \) denotes the sorting function of the complex argument that sorts the elements in descending order. \( X \sim \mathcal{C}\mathcal{N}(\mu, \sigma^2) \) represents the distribution of a circularly symmetric complex Gaussian r.v. \( X \) with mean \( \mu \) and variance \( \sigma^2 \). The probability of an event is denoted by \( P(\cdot) \). \( \lfloor \cdot \rfloor \) indicates the floor operation.

II. PROPOSED GQSM

A. System Model of GQSM

We first propose the GQSM scheme in Fig. 1 in this subsection. We consider an MIMO system with \( N_t \) transmit antennas and \( N_r \) receive antennas. By GQSM, \( p \) out of \( N_t \) transmit antennas are activated for conveying the real part of the modulated signal vectors with the selected antenna activation permutation (AAP) \( I^p \). Similarly, the imaginary part of the modulated signal vectors is transmitted through another \( p \) activated transmit antennas, resulting in the corresponding AAP \( I^{q} \). Note that the total number of AAPS for the re-
TABLE I

<table>
<thead>
<tr>
<th>index bits</th>
<th>AAP</th>
<th>index bits</th>
<th>AAP</th>
</tr>
</thead>
<tbody>
<tr>
<td>00</td>
<td>{1, 2}</td>
<td>01</td>
<td>{1, 3}</td>
</tr>
<tr>
<td>10</td>
<td>{1, 4}</td>
<td>11</td>
<td>{2, 3}</td>
</tr>
</tbody>
</table>

All imaginary part is \( C(N_t, p) \), but only \( L = 2^{\lceil \log_2 C(N_t, p) \rceil} \) AAPs are adopted for modulation purposes.

As can be seen in Fig. 1, \( m \) input bits are split into three parts, \( m_1, m_2, \) and \( m_3 \), where \( m_1 \) and \( m_3 \) \((m_1 = m_3)\) bits represent the index bits for the real and imaginary parts, respectively, and \( m_2 \) bits stand for the modulating bits generating the modulated signal vector. Specifically, the modulated signal vector \( s = [s_1, s_2, \ldots, s_p] \) are first generated by \( m_2 = p \log_2 M \) bits, where \( s_\tau = s_\tau^1 + js_\tau^Q \) with \( 1 \leq \tau \leq p \), \( s_\tau \in \mathcal{X} \), and \( \mathcal{X} \) denotes the M-ary constellation set. It should be noted that \( s_\tau^1 \) and \( s_\tau^Q \) are selected from \( \mathcal{X}^I \) and \( \mathcal{X}^Q \), respectively, where \( \mathcal{X}^I \) and \( \mathcal{X}^Q \) represent the real and imaginary parts, respectively. \( m_1 = m_3 = \lceil \log_2 C(N_t, p) \rceil \) bits are then utilized to determine the AAP \( I_\alpha^I \) and \( I_\alpha^Q \) so as to transmit the real and imaginary parts \( s_\tau^I \) and \( s_\tau^Q \), respectively. Let \( s^I = [s_1^I, s_2^I, \ldots, s_p^I] \) and \( s^Q = [s_1^Q, s_2^Q, \ldots, s_p^Q] \) denote the real and imaginary parts of the modulated signal vector \( s \). Let \( I_{\alpha}^I \) and \( I_{\beta}^Q \) be the \( \alpha \)th and \( \beta \)th AAPs for the real and imaginary parts, respectively, which can be represented by

\[
I_{\alpha}^I = \{i_{\alpha}^I(1), i_{\alpha}^I(2), \ldots, i_{\alpha}^I(p)\},
\]

and

\[
I_{\beta}^Q = \{i_{\beta}^Q(1), i_{\beta}^Q(2), \ldots, i_{\beta}^Q(p)\},
\]

respectively, where \( i_{\alpha}^I(\eta), i_{\beta}^Q(\eta) \in \{1, 2, \ldots, N_t\} \) with \( 1 \leq \eta \leq p \). The resulting \( N_t \times 1 \) signal vectors for the real and imaginary parts are given by

\[
x^I = [s_1^I, 0, s_2^I, \ldots, 0, s_p^I],
\]

and

\[
x^Q = [0, s_1^Q, s_2^Q, 0, \ldots, 0, s_p^Q],
\]

respectively. It is worth noting that the non-zero elements in \( x^I \) and \( x^Q \) are decided by AAPs \( I_\alpha^I \) and \( I_\beta^Q \). The \( N_t \times 1 \) transmitted signal vector is finally obtained by

\[
x = x^I + jx^Q.
\]

For clarity and illustration purposes, we provide an example for the GQSM scheme with \( N_t = 4, p = 2 \) and 4QAM \((M = 4)\), including the AAP mapping table, in Table I. The input bits are assumed to be \([0 0 1 1 0 1 1 1]\).

Let the transmitted signal vector \( x = x^I + jx^Q \) have \( (i_2^Q(2) = 3) \) antennas, resulting in \( x^Q = [0, 1, 1, 0]^T \). Therefore, the transmitted signal vector is obtained by \( x = x^I + jx^Q = [1, j, 1 + j, 0]^T \).

Let an \( N_r \times N_t \) matrix \( H \) denote the channel matrix whose entries follow the complex Gaussian distribution with zero mean and unit variance. The \( N_r \times 1 \) vector \( n \) denotes the complex Gaussian noise with zero mean and covariance \( \sigma^2 \). The optimal ML detection can be formulated as

\[
\hat{x} = \arg\min_x \|y - Hx\|^2
\]

\[
= \arg\min_x \|y - H(x^I + jx^Q)\|^2.
\]

It can be seen from (7) that the optimal ML detection results in the tremendous computation complexity due to a tremendous search space among all possible candidates of the transmitted signal vector \( x \), especially for large \( N_t, p, \) and/or \( M \).

B. Low-Complexity Detection Method

To relieve the detection burden at the receiver employing the optimal ML detection, we aim to propose a low-complexity detection method in this subsection, which considerably reduces the computational complexity.

Let us first separately rewrite the received signal vector into the real and imaginary parts as

\[
\begin{bmatrix}
R\{y\} \\
I\{y\}
\end{bmatrix} = \begin{bmatrix}
R\{H\} & -I\{H\} \\
I\{H\} & R\{H\}
\end{bmatrix} \begin{bmatrix}
x^I \\
x^Q
\end{bmatrix} + \begin{bmatrix}
R\{n\} \\
I\{n\}
\end{bmatrix}.
\]

We then calculate the weight metric of (8) by equalization as

\[
w_\varphi = h_\varphi^H \hat{y}/(h_\varphi^H h_\varphi),
\]

where \( \varphi \in \{1, 2, \ldots, 2N_t\} \) and \( h_\varphi \) denotes the \( \varphi \)th column of \( H \). The resulting sum of weight metrics for the real and imaginary parts containing the \( \alpha \)th \((I_{\alpha}^I)\) and \( \beta \)th \((I_{\beta}^Q)\) AAPs is obtained by

\[
g_\alpha = \sum_{\mu=1}^{p} |w_{i_{\alpha}^I(\mu)}|, \quad g_\beta = \sum_{\mu=1}^{p} |w_{i_{\beta}^Q(\mu)+N_t}|,
\]

where \( 1 \leq \alpha, \beta \leq L \). Let \( g = [g_1, g_2, \ldots, g_L] \) and \( f = [f_1, f_2, \ldots, f_L] \), the sequential AAPs for the real and imaginary parts are separately estimated by sorting \( g \) and \( f \) as

\[
[a_1, a_2, \ldots, a_L] = \text{sort}\{g\},
\]

\[
b_1, b_2, \ldots, b_L = \text{sort}\{f\},
\]

where \( a_1 \) and \( b_1 \) denote the positions of maximal values among \( g \) and \( f \), respectively, and \( a_L \) and \( b_L \) denote the positions of minimal values among \( g \) and \( f \). After obtaining the sorted AAPs, the modulated signals for the real part with the \( a_\alpha \)th AAP and the imaginary part with the \( b_\beta \)th AAP can be
obtained by the zero-forcing (ZF) approach as
\[
\tilde{s}^I_{\alpha}, \tilde{s}^Q_{\beta} = \text{demod}(\mathbf{H}_{\text{new}}^H \mathbf{H}_{\text{new}}^{-1} \mathbf{H}_{\text{new}}^H \hat{y}),
\]
where \(1 \leq \alpha, \beta \leq L\); \(\mathbf{H}_{\text{new}} = [\mathbf{H}^I_{\alpha-1}, \mathbf{H}^Q_{(\beta+1)}], \mathbf{H}^I_{\alpha-1}, \mathbf{H}^Q_{(\beta+1)}\) denote the submatrices of \(\mathbf{H}\) including the \(I_{\alpha-1}\) th and \(I_{(\beta+1)}\) th columns of \(\mathbf{H}\), respectively. The resulting Euclidean distances of AAPs via \(\{a_{\alpha}\}_{\alpha=1}^L\) and \(\{b_{\beta}\}_{\beta=1}^L\) can be sequentially calculated by
\[
e_{(a, b)} = \| \hat{y} - \mathbf{H}_{\text{new}} [\tilde{s}^I_{\alpha}, \tilde{s}^Q_{\beta}] \|^2.
\]
Due to the \(L\) possible candidates for \(\tilde{s}^I_{\alpha}\) (or \(\tilde{s}^Q_{\beta}\)), it should be noted that it still leads to high detection complexity by searching all \(L^2\) possible Euclidean distances, especially for large \(L\). In order to further mitigate this problem, we preset a detection threshold \(\eta\) and stop calculating the Euclidean distances in (14) once \(e_{(a, b)} < \eta\), which rapidly results in the estimated AAPs \(\{\hat{s}^I_{\alpha}, \hat{s}^Q_{\beta}\}\) and modulated symbols \(\hat{s}^I_{\alpha}, \hat{s}^Q_{\beta}\).

After obtaining \(\hat{l}^I_{\alpha}, \hat{l}^Q_{\beta}, \hat{s}^I_{\alpha}, \hat{s}^Q_{\beta}\), the transmitted signal, including the real and imaginary parts, are directly obtained, and the corresponding input bits can be recovered by the demapping of the AAPs and the demodulation of the constellation symbols.

C. Performance Analysis

According to (7), the conditional pairwise error probability (PEP) of detecting \(\hat{x}\) when \(x\) is transmitted on \(H\) is given by
\[
\Pr\{x \rightarrow \hat{x}|H\} = \Pr\{\|y - Hx\|^2 > \|y - H\hat{x}\|^2\} = Q\left(\sqrt{\frac{\|H(x - \hat{x})\|^2}{2N_0}}\right),
\]
Applying the Q-function approximation [48]
\[
Q(x) \approx \frac{1}{2} \left[1 - \frac{x^2}{2} + \frac{1}{e^{x^2/2}}\right],
\]
the unconditional PEP can be approximated as
\[
\Pr\{x \rightarrow \hat{x}\} = E_H \{\Pr\{x \rightarrow \hat{x}|H\}\} \approx \frac{1}{12} \frac{1}{(\det(I_{N_t} + q_1 K A))^{N_r}} + \frac{1}{4} \frac{1}{(\det(I_{N_r} + q_2 K A))^{N_r}},
\]
where \(A = (x - \hat{x})(x - \hat{x})^T, K = E\{HH^H\}\) is the covariance matrix of \(H\), \(q_1 = 1/(2N_0)\) and \(q_2 = 2/(3N_0)\). After obtaining the unconditional PEP, an upper bound on BER can be derived according to the union bounding technique as [49]
\[
P_e \leq \frac{1}{m_{2m}} \sum_x \sum_{x \neq x} d(x \rightarrow \hat{x}) \Pr\{x \rightarrow \hat{x}\},
\]
where \(d(x \rightarrow \hat{x})\) measures the number of bits in error between \(x\) and \(\hat{x}\).

At high SNR (i.e., \(\{q_1, q_2\} \gg 1\)), (18) can be approximated as
\[
P_e \approx \frac{1}{m_{2m}} \sum_x \sum_{x \neq x} (x - \hat{x})^T K^{-1} (x - \hat{x}) \Pr\{x \rightarrow \hat{x}\},
\]
Due to the \(L\) possible candidates for \(s\), (or \(s^Q\)), the transmitted signal, including the real and imaginary parts, are directly obtained, and the corresponding input bits can be recovered by the demapping of the AAPs and the demodulation of the constellation symbols.

III. APPLICATION OF GQSM IN COOPERATIVE VEHICULAR NETWORKS WITH NOMA

In this section, we propose two novel schemes, namely C-OQA-GQSM and C-NOMA-GQSM, which apply the GQSM scheme to the vehicular networks utilizing the OMA and NOMA techniques.

We consider a multi-vehicle single-relay cooperative downlink vehicular communication network, comprising one BS with \(N_t\) transmit antennas, vehicle 1 and vehicle 2 both with \(N_r\) receive antennas, and one DF relay with \(N_t\) transmit and \(N_r\) receive antennas operating in the half-duplex mode, where the BS aims to transmit two GQSM signal vectors \(x_1\) and \(x_2\) generated by \(m_1\) and \(m_2\) bits to vehicle 1 and vehicle 2, respectively. For simplicity, we assume that \(m_1 = m_2 = \log_2 C(N_t, p) + p \log_2 M\). Note that the vehicle 1 can directly receive the signal from the BS due to the near location, while there is no direct transmission link between the BS and vehicle 2 because of the severe propagation attenuation (long distance). Therefore, one DF relay is assumed to be located between the BS and vehicle 2, which leads to two phases for transmission from the BS to vehicle 1 and vehicle 2. To enable this two-phase transmission, we investigate two relaying transmission schemes, C-OMA-GQSM and C-NOMA-GQSM, in this paper.

A. Proposed C-OQA-GQSM

We first propose the C-OQA-GQSM scheme in Fig. 2, where the BS transmits two \(N_t \times 1\) signal vectors \(x_1\) and \(x_2\)
to vehicle 1 and vehicle 2 in two orthogonal phases (two time slots). The $N_t \times N_r$ channel matrices between BS to Vehicle 1, BS-to-Relay, Relay-to-Vehicle 1, and Relay-to-Vehicle 2 links are defined as $H_{\text{BV}1}, H_{\text{BR}}, H_{\text{RV}1},$ and $H_{\text{RV}2}$, respectively, where all entries of $H_{\text{BV}1}, H_{\text{BR}}, H_{\text{RV}1},$ and $H_{\text{RV}2}$ follow the complex Gaussian distribution with $\mathcal{C}\mathcal{N}(0, \sigma^2_{\text{BR}})$, $\mathcal{C}\mathcal{N}(0, \sigma^2_{\text{BV}1}),$ $\mathcal{C}\mathcal{N}(0, \sigma^2_{\text{RV}1}),$ and $\mathcal{C}\mathcal{N}(0, \sigma^2_{\text{RV}2}),$ respectively.

Specifically, the BS first transmits $x_2$ to the relay in the first phase, and the $N_r \times 1$ received signal vector at the relay can be expressed as

$$y_{\text{BR}}^{\text{OMA}} = H_{\text{BR}} x_2 + n_{\text{BR}},$$

where $n_{\text{BR}}$ denotes the $N_r \times 1$ noise vector following the distribution $\mathcal{C}\mathcal{N}(0, \sigma^2_{\text{BR}})$. After receiving $y_{\text{BR}}^{\text{OMA}}$, the DF relay performs the ML detection to estimate the transmitted signal vector as

$$\hat{x}_2 = \arg \min_{x_2} \|y_{\text{BR}}^{\text{OMA}} - H_{\text{BR}} x_2\|^2.$$  

In the second phase, the DF relay broadcasts the estimated signal vector $\hat{x}_2$ to both vehicle 1 and vehicle 2. At the vehicle 2, the $N_r \times 1$ signal vector is obtained by

$$y_{\text{RV}2}^{\text{OMA}} = H_{\text{RV}2} \hat{x}_2 + n_{\text{RV}2},$$

where $n_{\text{RV}2}$ denotes the $N_r \times 1$ noise vector following the distribution $\mathcal{C}\mathcal{N}(0, \sigma^2_{\text{RV}2})$. The transmitted signal vector for vehicle 2 is finally estimated by the ML detection as

$$\hat{x}_2' = \arg \min_{x_2} \|y_{\text{RV}2}^{\text{OMA}} - H_{\text{RV}2} x_2\|^2.$$ 

Meanwhile, vehicle 1 simultaneously receives the signals transmitted from both of the BS and the DF relay, which yields the $N_r \times 1$ signal vector as

$$y_{\text{BV}1}^{\text{OMA}} = H_{\text{BV}1} x_1 + H_{\text{RV}} \hat{x}_2 + n_{\text{BV}1},$$

where $n_{\text{BV}1}$ denotes the $N_r \times 1$ noise vector following the distribution $\mathcal{C}\mathcal{N}(0, \sigma^2_{\text{BV}1})$. The transmitted signal vector for vehicle 1 is similarly estimated by the ML detection as

$$x_1 = \arg \min_{x_1} \|y_{\text{BV}1}^{\text{OMA}} - H_{\text{BV}1} x_1\|^2.$$ 

It is worth noting that although the detection on $x_2$ in (23) is sufficiently reliable, the estimation on $x_1$ is very inaccurate due to the non-erasable interference from the DF relay, which results in catastrophic BER performance of vehicle 1. Therefore, the C-OMA-NOMA scheme might not be suitable for certain practical applications, which need to guarantee the average BER performance of all vehicles. Therefore, to relieve this problem, we further propose a more practical scheme in the next subsection.

B. Proposed C-NOMA-GQSM

The system model for the C-NOMA-GQSM scheme is shown in Fig. 3, which still needs two phases to complete the whole transmission. It should be noted that vehicle 1 is near to the BS, and the DF relay is relatively far from the BS, which leads to $\sigma^2_{\text{BV}1} > \sigma^2_{\text{BR}}$. Due to the NOMA principle, the BS will combine two signal vectors by $x_{\text{ALL}} = \sqrt{\xi_1} x_1 + \sqrt{\xi_2} x_2$, where $\xi_1$ and $\xi_2$ denote the power allocation factors for vehicle 1 and vehicle 2, respectively, and $\xi_1 + \xi_2 = 1$ as the transmit power constraint. Since vehicle 1 is near to the BS and vehicle 2 as well as the relay are far from the BS, the BS will allocate more power to user 2 ($x_2$) and less power to vehicle 1 ($x_1$) by NOMA, which leads to $\xi_1 < \xi_2$.

By C-NOMA-GQSM, the BS first broadcasts the combined signal $x_{\text{ALL}}$ to both vehicle 1 and the relay in the first phase. The $N_r \times 1$ received signal vectors at vehicle 1 and the relay can be expressed as

$$y_{\text{BV}1}^{\text{NOMA}} = H_{\text{BV}1} x_{\text{ALL}} + n_{\text{BV}1} = H_{\text{BV}1} (\sqrt{\xi_1} x_1 + \sqrt{\xi_2} x_2) + n_{\text{BV}1},$$

and

$$y_{\text{BR}}^{\text{NOMA}} = H_{\text{BR}} x_{\text{ALL}} + n_{\text{BR}} = H_{\text{BR}} (\sqrt{\xi_1} x_1 + \sqrt{\xi_2} x_2) + n_{\text{BR}},$$

respectively, where $n_{\text{BV}1}$ and $n_{\text{BR}}$ denote the the $N_r \times 1$ noise vectors following the distribution $\mathcal{C}\mathcal{N}(0, \sigma^2_{\text{BV}1})$ and $\mathcal{C}\mathcal{N}(0, \sigma^2_{\text{BR}})$, respectively. Vehicle 1 retains the received signal $y_{\text{BV}1}^{\text{NOMA}}$ and waits for another upcoming signal in the second phase for detection purposes. With the help of the relay, the transmitted signal vector to vehicle 2 can be easily detected due its larger transmit power by the ML detection as

$$\hat{x}_2 = \arg \min_{x_2} \|y_{\text{BR}}^{\text{NOMA}} - \sqrt{\xi_2} H_{\text{BR}} x_2\|^2.$$ 

In the second phase, the DF relay broadcasts the estimated signal $\hat{x}_2$ to vehicle 1 and vehicle 2, where the $N_r \times 1$ received signal vectors at vehicle 1 and vehicle 2 are given by

$$y_{\text{RV}1}^{\text{NOMA}} = H_{\text{RV}1} \hat{x}_2 + n_{\text{RV}1},$$

and

$$y_{\text{RV}2}^{\text{NOMA}} = H_{\text{RV}2} \hat{x}_2 + n_{\text{RV}2},$$

respectively, where $n_{\text{RV}1}$ and $n_{\text{RV}2}$ denote the the $N_r \times 1$ noise vectors following the distribution $\mathcal{C}\mathcal{N}(0, \sigma^2_{\text{RV}1})$ and $\mathcal{C}\mathcal{N}(0, \sigma^2_{\text{RV}2})$, respectively. Vehicle 2 then performs the ML detection as

$$\hat{x}_2' = \arg \min_{x_2} \|y_{\text{RV}2}^{\text{NOMA}} - H_{\text{RV}2} x_2\|^2.$$ 

On the other hand, since the signal vector $x_2$ transmitted to vehicle 2 has larger transmit power than $x_1$ to vehicle 1 ($\xi_1 < \xi_2$), the SIC process should be implemented by incorporating the received signal in the first phase $y_{\text{BR}}^{\text{NOMA}}$ before detecting
x_1 \text{ at vehicle } 1, \text{ which is given by}
\hat{x}_2^\prime = \arg \min_{x_2} \{ \|y_{RV_1}^{\text{NOMA}} - H_{RV_1} x_2\|^2 + \|y_{RV_1}^{\text{NOMA}} - \sqrt{\zeta_2} H_{BV_1} \hat{x}_2^\prime\|^2 \}. \quad (32)

After estimating the signal vector \(x_2^\prime\) of vehicle 2, the received signal vector at vehicle 1 is updated by eliminating the interference from vehicle 2 as
\[ y_{BV_1}^{\text{Update}} = y_{RV_1}^{\text{NOMA}} - \sqrt{\zeta_2} H_{BV_1} \hat{x}_2^\prime. \quad (33) \]

The transmitted signal vector to vehicle 1 can be finally estimated by
\[ \hat{x}_1 = \arg \min_{x_1} \|y_{BV_1}^{\text{Update}} - \sqrt{\zeta_1} H_{BV_1} x_1\|^2. \quad (34) \]

Note that the BER performance of vehicle 2 is limited due to the signal detection criterion given in (28) at the DF relay, which suffers from the non-erasable multi-vehicle interference from vehicle 1, so as to bring the error floor on BER of vehicle 2. However, unlike the catastrophic BER performance of vehicle 1 in C-OMA-GQSM, vehicle 1 in C-NOMA-GQSM achieves comparable BER performance with the aid of the DF relay in (32) and the SIC process in (33), which successfully suppress the interference from vehicle 2 to some degree. Despite of the existence of error floor, the BER performance of vehicle 1 in C-NOMA-GQSM systems is significantly improved compared with that of vehicle 1 in C-OMA-NOMA systems. Therefore, it is concluded that the C-OMA-NOMA scheme has better BER performance of vehicle 2, while the C-NOMA-GQSM scheme achieves better overall average BER performance, which makes C-NOMA-GQSM more suitable for practical applications.

IV. SIMULATION RESULTS

In this section, we present computer simulation results to evaluate the BER performance of GQSM systems under the assumptions of Rayleigh fading channels and perfect channel estimation. For the sake of simplicity, we denote “GQSM (N_t, N_r, p, M\text{QAM})” as the GQSM scheme with \(N_t\) transmit antennas, \(N_r\) receive antennas, \(p\) activated transmit antennas and \(M\)-ary QAM constellation, and “QSM (N_t, N_r, M\text{QAM})” as the QSM scheme with \(N_t\) transmit antennas, \(N_r\) receive antennas, and \(M\)-ary QAM constellation. The detection threshold \(\eta\) is set to be \(2N_r \sigma^2\) for the proposed low-complexity detection.

In Fig. 4, we compare the BER performance between GQSM and QSM schemes with \(N_t = 4, N_r = 2\). It can be seen that GQSM (4, 2, 3, 4QAM) achieves a large performance gain than QSM (4, 2, 64QAM) both with 10 input bits in the entire SNR region. Specifically, compared to QSM (4, 2, 64QAM), GQSM (4, 2, 3, 4QAM) obtains almost 6 dB SNR gain at BER=10^{-2} and obtains up to 8dB SNR gain at BER=10^{-4}. This is because GQSM significantly increases the length of index bits, which improves the system performance. From Fig. 4, it can also be seen that the theoretical curves perfectly match the simulation results in the high SNR region, which validates our analysis in Section II. C. Moreover, we can see that the proposed low-complexity detection (near-ML) closely approaches the optimal ML detection for GQSM (4, 2, 3, 4QAM). To be more specific, the BER curve of the low-complexity detection is close to that of the ML detection with a trivial performance loss in the low SNR region due to the non-negligible noise impact, while the BER curves of the low-complexity and ML detection almost coincide in the high SNR region.

We further compare the BER performance between GQSM and QSM schemes with \(N_t = 8, N_r = 4\), \(m = 12\) bits in Fig. 5. Similar to Fig. 4, GQSM (8, 4, 2, 4QAM) still outperforms QSM (8, 4, 64QAM) with a large performance gain, especially in the medium-to-high SNR region. At BER=10^{-4}, GQSM (8, 4, 2, 4QAM) acquires about 10 dB SNR gain compared with QSM (8, 4, 64QAM), which attributes to the benefit of a larger number of index bits in GQSM. It can also be seen in this figure that the theoretical curve matches the simulation curve very well for GQSM (8, 4, 2, 4QAM)
Specifically, vehicle 1 with \( \zeta_1 \) region, while vehicle 1 in C-NOMA-GQSM achieves superior verification of the diversity order in Section II. B. Therefore, we can conclude that the diversity order of GQSM is only related to \( N \) and irrelevant to \( N_t, p \), and \( M \), which verifies the analysis of the diversity order in Section II. B.

To evaluate the diversity of the proposed GQSM, we compare the BER performance of GQSM (4, 2, 1, 4QAM), GQSM (4, 3, 1, 4QAM), GQSM (4, 4, 1, 4QAM), and GQSM (8, 4, 2, 8QAM) in Fig. 6. It is shown from Fig. 6 that GQSM (4, 2, 1, 4QAM) achieves the smallest diversity order as 2 \((N_r = 2)\). As only increasing \( N_r \), GQSM (4, 3, 1, 4QAM) reaches a higher diversity order as 3 \((N_r = 3)\). By keeping increasing \( N_r \) up to 4, GQSM (4, 4, 1, 4QAM) obtains the diversity order of 4 \((N_r = 4)\), which indicates that the diversity order of GQSM is continuously increased by only changing the number of receive antennas \( N_r \) from 2 to 4. To further explore the diversity property, we compare the BER performance between GQSM (4, 4, 1, 4QAM) and GQSM (8, 4, 2, 8QAM), which shows that GQSM (4, 4, 1, 4QAM) achieves the same diversity order \((N_r = 4)\) as that of GQSM (8, 4, 2, 8QAM) by only changing the number of transmit antennas \( N_t \), the number of activated transmit antennas \( p \) and the cardinality of constellation \( M \). Therefore, we can conclude that the diversity order of GQSM is only related to \( N_r \) and irrelevant to \( N_t, p \), and \( M \), which verifies the analysis of the diversity order in Section II. B.

Fig. 7 presents the BER performance comparison of vehicle 1 between C-OMA-GQSM and C-NOMA-GQSM with \( N_t = 4, N_r = 2, p = 2 \), and 4QAM. It can be seen from Fig. 7 that vehicle 1 in C-OMA-GQSM has very poor performance, which leads to the constant BER=0.4 in the entire SNR region, while vehicle 1 in C-NOMA-GQSM achieves superior BER performance with different power allocation factors \((\zeta_1)\). Specifically, vehicle 1 with \( \zeta_1 = 0.2 \) obtains the relatively poor performance and exhibits the ineradicable error floor for SNR>35dB. The performance of vehicle 1 with a smaller value of \( \zeta_1 = 0.1 \) becomes better and as low as the error floor of BER=10^{-2} for SNR>50dB. Keeping decreasing the power allocation factor to \( \zeta_1 = 0.05 \), vehicle 1 obtains much better BER performance and the error floor appears only for SNR>65dB. Setting \( \zeta_1 = 0.02 \), vehicle 1 obtains the best BER performance in Fig. 7, and the error floor appears merely in the range of SNR>80dB, which can be ignored in practical situations. Based on the above results, we can observe that allocating less power to vehicle 1 will greatly improve its BER performance and postpone the occurrence of the BER error floor in the medium-to-high SNR region. This is because of the fact that allocating less power to vehicle 1 is equivalent to allocate more power to the DF relay, which is beneficial to the SIC process of vehicle 1 in the second phase so as to improve the BER performance.

Fig. 8 shows the BER performance comparison of vehicle 2 between C-OMA-GQSM and C-NOMA-GQSM with \( N_t = 4, N_r = 2, p = 2 \), and 4QAM. From Fig. 8, we can see that
vehicle 2 by C-OMA-GQSM achieves the best BER performance and by C-NOMA-GQSM successively approximates to that in C-OMA-GQSM with increasing its power allocation factor $\zeta_2$. To be specific, vehicle 2 with $\zeta_2 = 0.8$ obtains undesirable BER performance and the error floor arises at about BER=0.4. Allocating more power with $\zeta_2 = 0.8$ yields better BER performance of vehicle 2 even if it is still worse than that of C-OMA-GQSM. Note that the error floor is delayed at BER=10^{-5} for SNR>35dB. Keeping increasing the power allocation factor up to $\zeta_2 = 0.9$, vehicle 2 obtains much better performance and the BER error floor on BER=5×10^{-4} appears for SNR>40dB. It should also be noted that the BER curve of vehicle 2 with $\zeta_2 = 0.9$ is closer to that of C-OMA-GQSM. Finally, the BER curve of vehicle 2 with a larger value of $\zeta_2 = 0.98$ obtains the best performance with the error floor of BER=10^{-5}, which approaches the BER curve of vehicle 2 by C-OMA-NOMA much closer. Therefore, despite of the error floor, we conclude that the BER performance of vehicle 2 in C-NOMA-GQSM tightly approaches that of vehicle 2 using C-OMA-GQSM in the medium-to-low SNR region. As can be seen from Figs. 7 and 8 that the C-OMA-GQSM scheme has the extreme situation of BER and brings consistently poor performance to vehicle 1 but the excellent performance to vehicle 2. However, the BER performance of vehicle 1 and vehicle 2 in C-NOMA-GQSM systems is gradually improved as $\zeta_1$ ($\zeta_2$) decreases (increases), which verifies the analysis in Section III that the C-NOMA-GQSM scheme is more suitable for practical purposes. In addition, we also find that vehicle 1 and vehicle 2 utilizing C-NOMA-GQSM finally achieve the same BER performance in the high SNR region, which guarantees the overall system performance and reliability.

To further validate our observations, we compare the BER performance of vehicle 1 and vehicle 2 in the contexts of C-OMA-GQSM and C-NOMA-GQSM with $N_t = 8$, $N_r = 4$, $p = 1$, and 4QAM in Figs. 9 and 10. Similar to Fig. 7, vehicle 1 with different values of $\zeta_1$ employing C-NOMA-GQSM achieves better performance than that of C-OMA-GQSM, where the BER of vehicle 1 employing C-NOMA-GQSM monotonously decreases when increasing $\zeta_1$. Note that the error floor of vehicle 1 comes out with a very small BER if $\zeta_1$ is small enough. For example, the error floor of vehicle 1 arises at BER=6×10^{-6} with $\zeta_1 = 0.1$, and the error floor of vehicle 1 may arise at even a lower BER (this has not been shown in Fig. 9) with $\zeta_1 = 0.05$, which indicates that vehicle 1 utilizing C-NOMA-GQSM is able to obtain sufficiently good performance as long as $\zeta_1$ is low enough. Meanwhile, we can figure out from Fig. 10 that the BER curves of vehicle 2 in C-NOMA-GQSM systems are able to closely approach those corresponding to C-OMA-GQSM with a larger value of $\zeta_2$ in the medium-to-low SNR region, even if the error floor still exists. This further corroborates that the proposed C-NOMA-GQSM scheme is more suitable for large-scale multi-vehicle communication systems with large numbers of transmit and receive antennas.

V. CONCLUSIONS

In this paper, we proposed the GQSM scheme to increase the SE of the conventional QSM scheme by allowing activating more than one active transmit antenna in in-phase and quadrature domains. By GQSM, the transmitted signals are divided into real and imaginary parts and conveyed by their corresponding active transmit antennas, which makes index bits double. A near-ML low-complexity detection was designed to greatly reduce the computational complexity of the optimal ML detection. An upper bound on BER and the diversity order were analyzed to evaluate the performance of GQSM. Moreover, we proposed the C-OMA-GQSM and C-NOMA-GQSM schemes to further investigate the practical application of GQSM by resorting to the NOMA technique. Simulation results have discovered that GQSM obtains a significant performance gain compared to QSM in the entire SNR region, and validated the effectiveness of the low-complexity detection and theoretical analysis. Besides, simulation results have also revealed that compared to C-OMA-GQSM, C-NOMA-GQSM
achieves better average BER performance and is thereby more suitable for multi-vehicle communication networks, in which vehicle fairness is of key importance.

REFERENCES


Jun Li (S’13–M’17) received the M.S. and the Ph.D. degrees from Chonbuk National University, Jeonju, South Korea, in 2011 and 2016, respectively. He is currently an Assistant Professor with Guangzhou University, Guangzhou, China. He serves as a reviewer for IEEE TRANSACTIONS ON WIRELESS COMMUNICATIONS, IEEE TRANSACTIONS ON COMMUNICATIONS, IEEE JOURNAL OF SELECTED TOPICS IN SIGNAL PROCESSING, IEEE WIRELESS COMMUNICATIONS LETTERS, IEEE COMMUNICATIONS LETTERS, and IEEE TRANSACTIONS ON VEHICULAR TECHNOLOGY. His research interests include spatial modulation and OFDM index modulation.

Shuping Dang (S’13–M’18) received B.Eng (Hons) in Electrical and Electronic Engineering from the University of Manchester (with first class honors) and B.Eng in Electrical Engineering and Automation from Beijing Jiaotong University in 2014 via a joint ‘2+2’ dual-degree program. He also received D.Phil in Engineering Science from University of Oxford in 2018. Dr. Dang joined in the R&D Center, Huanan Communication Co., Ltd. after graduating from University of Oxford and is currently working as a Postdoctoral Fellow with the Computer, Electrical and Mathematical Science and Engineering Division, King Abdullah University of Science and Technology (KAUST). He is a co-recipient of the best Paper Award for work presented at 2019 19th IEEE International Conference on Communication Technology. He serves as a reviewer for a number of key journals in communications and information science, including IEEE TRANSACTIONS ON WIRELESS COMMUNICATIONS, IEEE TRANSACTIONS ON COMMUNICATIONS, IEEE WIRELESS COMMUNICATIONS LETTERS, IEEE COMMUNICATIONS LETTERS, and IEEE TRANSACTIONS ON VEHICULAR TECHNOLOGY. Dr. Dang is recognized as the Exemplary Reviewer of IEEE Communications Letters in 2019. His current research interests include novel modulation schemes, cooperative communications, terahertz communications, and 6G wireless network design.

Yier Yan received the B.S. degree from South-Central Nationality University, Wuhan, Hubei, P. R. China, in Applied Electronics. He received the M.S. and Ph.D. degrees from Chonbuk National University, Jeonju, Korea, in Communication Engineering. He is currently working at School of Mechanical and Electrical Engineering of Guangzhou University, China. His research interests include information theory, signal processing and OFDM in MIMO system.

Yuyang Peng received M.S. and Ph.D. degrees in electrical and electronic engineering from Chonbuk National University, Jeonju, Republic of Korea in 2011 and 2014, respectively. He worked as a Postdoctoral Research Fellow in Korea Advanced Institute of Science and Technology (KAIST), Daejeon, Republic of Korea, from 2014 to 2018. He is currently an Assistant Professor with the Faculty of Information Technology at Macau University of Science and Technology (MUST), Taipa, Macau SAR. His research activities lie in the broad area of digital communications, wireless sensor networks, and computing. In particular, his current research interests include cooperative communications, spatial modulation, and energy optimization. He currently serves as an Associate Editor for the IEEE Access.

Saba Al-Rubaye received the PhD degree in Electronic and Computer Engineering from Brunel University London, United Kingdom, in 2013. She is currently a Senior Lecturer leading the connected system research group in the School of Aerospace, Transport and Manufacturing at Cranfield University, United Kingdom. She worked in industry where she participated in developing loop testbed in control and communications at Quanta’s Sustainable Technology Integration Laboratory (QT-STIL) in Toronto, Canada. She has been involved in several projects sponsored by the Natural Sciences and Engineering Research Council of Canada (NSERC), Engineering and Physical Sciences Research Council (EPSRC) for drone connectivity, Department for Transport (DfT) funded projects for Airport security, and H2020. She is a voting member of IEEE P1920.2 “Standard for Vehicle-to-Vehicle Communications for Unmanned Aircraft Systems” and IEEE P1932.1 “Standard of License/unlicensed Interoperability”. She has published many papers in IEEE journals and conferences and was the recipient of best technical paper award twice published in IEEE Vehicular Technology in 2011 and 2015. She served as a general co-chair and as a member of the technical program committee for many international conferences. Her main research interests are focusing on airport connectivity, UAV, 5G and beyond, communications networks and autonomous systems. Dr. Al-Rubaye is registered as a Chartered Engineer (CEng) by British Engineering Council, Senior Member of Institute of Electrical and Electronics Engineers (IEEE), Associate Fellow of the British Higher Education Academy (AFHEA) and a certified UAV pilot.

Antonios Tsourdos obtained an MEng on Electronic, Control and Systems Engineering from the University of Sheffield (1995), an MSc on Systems Engineering from Cardiff University (1996) and a PhD on Nonlinear Robust Flight Control Design and Analysis from Cranfield University (1999). He is a Professor of Autonomous Systems and Control with Cranfield University. He was appointed Head of the Autonomous Systems Group in 2007, Head of the Centre of Autonomous and Cyber-Physical Systems in 2012 and Director of Research - Aerospace, Transport and Manufacturing in 2015. Professor Tsourdos was member of the Team Stellar, the winning team for the UK MoD Grand Challenge (2008) and the IET Innovation Award (Category Team, 2009). Professor Tsourdos is chair of the IFAC Technical Committee on Aerospace Control, and member of the UK Autonomous Systems National Technical Committee. He is editorial board member for the IEEE Transactions on Aerospace and Electronic Systems, the Proceedings of the Institution of Mechanical Engineers, Part G: Journal of Aerospace Engineering, the Aerospace Science and Technology, the International Journal of Systems Science and the Journal of Intelligent and Robotic Systems.