Performance Analysis of NOMA-SM in Vehicle-to-Vehicle Massive MIMO Channels

Yingyang Chen, Li Wang, *Senior Member, IEEE,* Yutong Ai, Bingli Jiao, *Member, IEEE,* and Lajos Hanzo, *Fellow, IEEE*

***Abstract*—At the time of writing, vehicle-to-vehicle (V2V) communication is enjoying substantial research attention as a benefit of its compelling applications. However, the ever-increasing teletraffic is expected to result in overcrowding of the available band. As a first resort, multiple-input multiple-output (MIMO) can be utilized to enhance the attainable bandwidth efficiency or link reliability. However, in hostile V2V wireless propagation environments the achievable multiple-antenna gain is eroded by the channel correlation. As a promising MIMO technique, spatial modulation (SM) only activates a single transmit antenna (TA) in any symbol-interval and hence completely avoids the inter-antenna interference (IAI), hence showing robustness against channel correlation. As a further powerful solution, non-orthogonal multiple access (NOMA) has been proposed for improving the bandwidth efficiency. Inspired by the robustness of SM against channel correlation and the benefits of NOMA, we intrinsically amalgamate them into NOMA-SM in order to deal with the deleterious effects of wireless V2V environments as well as to support improved bandwidth efficiency. Moreover, the bandwidth efficiency of NOMA-SM is further boosted with the aid of a massive TA configuration. Specifically, a spatio-temporally correlated Rician channel is considered for a V2V scenario. We investigate the bit error ratio (BER) performance of NOMA-SM via Monte Carlo simulations, where the impact of the Rician K-factor, spatial correlation of the antenna array, time-varying effect of the V2V channel, and the power allocation factor is discussed. Furthermore, we also analyse the capacity of NOMA-SM. By analysing the capacity and deriving closed-form upper bounds on the capacity, a pair of power allocation optimization schemes are formulated. The optimal solutions are demonstrated to be achievable with the aid of our proposed algorithm. Again, instead of simply invoking a pair of popular techniques, we intrinsically amalgamate SM and NOMA to conceive a new system component exhibiting distinct benefits in the V2V scenarios considered.**

***Index Terms*—Spatial modulation (SM), non-orthogonal multiple access (NOMA), massive multiple-input multiple-output (MIMO), vehicle-to-vehicle (V2V), channel capacity, bit error ratio (BER).**

This work is supported in part by the NSFC of China (Grant No. 61571056 and 61531004), the open research fund of the National Mobile Communications Research Laboratory of Southeast University (Grant No. 2016D04), and the State Major Science and Technology Special Projects of China under Grant 2016ZX03001017-004. L. Hanzo would like to thank the European Research Council for the financial support of their Advanced Fellow Grant. (*Corresponding author: L. Wang.*)

Y. Chen and B. Jiao are with the School of Electronics Engineering and Computer Science, Peking University, Beijing 100871, China (email: {chenyingyang, jiaobl}@pku.edu.cn).

L. Wang and Y. Ai are with the Beijing Key Laboratory of Work Safety Intelligent Monitoring, School of Electronic Engineering, Beijing University of Posts and Telecommunications, Beijing 100876, China (email: {liwang, ytai}@bupt.edu.cn). L. Wang is also with the National Mobile Communications Research Laboratory, Southeast University, Nanjing, 210096, China.

L. Hanzo is with the School of Electronics and Computer Science, University of Southampton, Southampton, SO17 1BJ, U.K. (email: lh@ecs.soton.ac.uk).

I. INTRODUCTION

Multiple-input multiple-output (MIMO) schemes have found their way into operational standards for improving their performance. Traditionally, MIMO schemes have been designed either to enhance the diversity gain by combating the channel fading (e.g., Alamouti code), or for spatial multiplexing (e.g., Vertical Bell Laboratories Layered Space-Time, termed VBLAST), albeit they are amalgamated by the multi-functional MIMO concept of [1]. To accommodate the ever-increasing demands of multimedia services and applications, the massive MIMO concept emerged [2], [3]. Theoretically, massive MIMO reaps all the benefits of conventional MIMO and offers abundant degrees of freedom (DoFs). By exploiting the knowledge of the channel state information at the transmitter (CSIT), a massive antenna array becomes capable of simultaneously serving a large number of users by sharing its multiplexing gain among them, while providing higher data rates and transmission reliability. In vehicle-to-vehicle (V2V) communications, large scale MIMOs become quite attractive, since multiple antennas can be accommodated [4], [5].

However, massive MIMOs suffer from various problems, including the inter-antenna interference (IAI) and the high complexity of the receivers. It would be a particularly costly process to acquire CSIT in frequency-division duplexing (FDD) systems. Moreover, the hardware cost (e.g., a dedicated radio frequency (RF) chain associated with each antenna) becomes excessive for large antenna arrays. In vehicular wireless communications, the gravest challenge is the hostile high-Doppler propagation imposed. For example, the dominant Doppler effect aggravates the inter-subcarrier interference of orthogonal frequency division multiple (OFDM) and the strong line-of-sight (LoS) component of V2V channels would aggravate the spatial correlation between antennas. Therefore, the applications of massive antenna technologies in V2V transmissions are deemed to be problematic due to the aforementioned issues.

In recent years, spatial modulation (SM) [6] has been regarded as a promising multiple-antenna technique of improving the bandwidth efficiency. In contrast to the traditional MIMO configurations, SM only activates a single transmit antenna (TA), hence the IAI can be completely eliminated and only a single RF chain is required. Thus, a reduced implementation cost and complexity is achievable in SM systems. Moreover, the bandwidth efficiency of SM can be further improved by employing a large TA array, providing a feasible transceiver solution for massive MIMO with no CSIT [7], [8].

A recent survey of SM can be found in [9]. In [10]–[13], SM and its extensions were considered in vehicular environments. A differential SM scheme was proposed for vehicle communications in [10], exhibiting robustness against time-selective fading and Doppler effects. Fu *et al.* [11] studied the bit error ratio (BER) performance of SM under a three-dimensional V2V channel model. Peppas *et al.* [12] applied space shift keying (SSK) in inter-vehicular communications and derived a closed-form expression for the pairwise error probability. In [13], the performance of massive SM MIMO over a spatio-temporally correlated Rician channel was analysed under a high-speed railway scenario. Moreover, Cui and Fang have demonstrated that by activating a single TA, SM is capable of alleviating the channel correlation. In conclusion, SM has become increasingly appealing for V2V systems.

On the other hand, due to the explosive growth of data traffic, there are increasing demands for high bandwidth efficiency and massive connectivity in 5th generation (5G) wireless communications. To address these challenges, various novel multiple access techniques have been proposed such as sparse code multiple access (SCMA), pattern division multiple access (PDMA), and non-orthogonal multiple access (NOMA) [14]. Among these techniques, NOMA exhibits an appealing low receiver complexity, high bandwidth efficiency, and massive connectivity by allowing multiple users to share the same channel resource via power domain multiplexing. Thus, NOMA is considered to be a promising candidate for future wireless access [15]. To mitigate the multiple access interference (MAI), multi-user detection (MUD) techniques such as successive interference cancellation (SIC) [16] can be applied at the end-user receivers for detecting the desired signals. Through power domain multiplexing at the transmitter and SIC at the receivers, NOMA becomes capable of fully exploiting its capacity region hence outperforming the orthogonal multiple access (OMA) schemes [17].

The specific design aspects of the NOMA schemes have been discussed in [18]–[20]. Explicitly, in [18], the concept of basic NOMA with SIC was introduced and its performance was compared to that of the traditional orthogonal frequency division multiple access (OFDMA) scheme through a system-level evaluation. A beneficial power allocation scheme was designed in [19] for striking compelling tradeoffs between the user fairness and system throughput. Lv *et al.* [20] studied a new cooperative NOMA transmission scheme and derived the outage probability associated with fixed power allocation.

The broad objective of vehicular communications is to improve the travel-experience of users by offering improved safety, internet access, and infotainment services. IEEE 802.11p forms the standard of Wireless Access for Vehicular Environments (WAVE), providing data rates ranging from 6 to 27 Mbps for short transmission distances [21]. As an alternative to the IEEE 802.11p-based vehicular ad hoc network (VANET), Long-Term Evolution (LTE) based V2V is supported by the Third-Generation Partnership Project (3GPP) so as to provide efficient message dissemination [22]. Nevertheless, the ever-growing demands for vehicular communications increase the gravity of tele-traffic congestion.

Hence we aim for designing a novel transmission scheme, termed NOMA-SM, to intrinsically amalgamate NOMA and SM. In synergy with the inherent demand of V2V transmissions for high bandwidth efficiency, NOMA is invoked for non-orthogonally accessing all the resources combined with the single-RF benefits of SM. The bandwidth efficiency of the proposed NOMA-SM scheme is further boosted by a massive TA configuration. At the time of writing, there is a paucity of results the amalgam of NOMA and SM, especially in inter-vehicle communications.

Against this background, the main contributions of this paper are three-fold:

- Firstly, we propose the novel NOMA-SM concept conceived for V2V communications and quantify its link reliability improvement. A spatio-temporally correlated Rician channel is considered for our V2V scenario, where the effects of the Rician K-factor, adjacent antenna correlation coefficient, temporal correlation and power allocation factor are all quantified. The results demonstrate that NOMA-SM exhibits robustness against the deleterious effects of V2V environments.
- Secondly, we derive the capacity of NOMA-SM, and verify it by Monte Carlo simulations. The benefits of SIC are demonstrated both theoretically and numerically. The ergodic capacity of the collaboration-aided vehicle is also determined for a simplified V2V channel, which is shown to closely approximate that of a spatio-temporally correlated Rician channel.
- Thirdly, we formulate a pair of analytical upper bounds on the capacity of NOMA-SM in closed form and propose a pair of power allocation optimization schemes. The optimal solutions are demonstrated to be achievable with the aid of the proposed power allocation algorithms. Our numerical results verify the improved bandwidth efficiency of NOMA-SM.

Explicitly, instead of simply combining a pair of popular techniques, we intrinsically amalgamate their benefits. By investigating the BER performance of NOMA in comparison to different MIMO techniques and the bandwidth efficiency of SM combined with distinct multiple access methods, NOMA and SM are shown to cooperatively improve V2V transmissions.

The rest of this treatise is organized as follows. In Section II, the system model of NOMA-SM is presented, while Section III provides the capacity analysis and mutual information (MI) evaluation of NOMA-SM. Our capacity upper bound derivations and power allocation problem are considered in Section IV. Simulation results and discussions for BER performance are provided in Section V, together with the numerical capacity analysis and power allocation optimization. Finally, Section VI concludes by summarizing the results. For convenience, we list the most frequent notations here.

Notation: Uppercase and lowercase bold-faced letters indicate matrices and vectors, respectively. $(\cdot)^{-1}$, $(\cdot)^{H}$, $\det(\cdot)$, and $[\cdot]_{p,q}$ represent inverse, conjugate-transpose, determinant, and the entry in the p-th row and q-column of a matrix, respectively. $\mathbb{E}_X\{\cdot\}$ denotes the expectation on the random variable X. $\mathbf{A} \in \mathbb{C}^{M\times N}$ is a complex-element matrix with

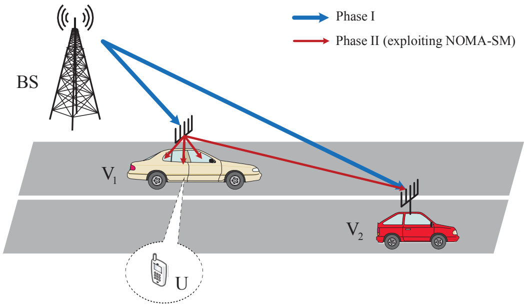

Fig. 1. An illustration of the considered vehicular communication system, where NOMA-SM is applied in Phase II.

dimensions $M \times N$, and $\mathbf{I}_N$ is an $N \times N$ identity matrix. $|\cdot|$ and $(\cdot)^*$ imply the absolute value and the conjugate of a complex scalar, while $\|\cdot\|$ denotes the Euclidean norm of a vector. Finally, $x \sim \mathcal{CN}\left(\mu, \sigma^2\right)$ indicates that the random variable x obeys a complex Gaussian distribution with mean μ and variance σ^2.

II. Proposed Scheme

We consider a generic vehicular communication system, where the vehicle-to-infrastructure (V2I), V2V, and intra-vehicle transmissions are all included. As shown in Fig. 1, a base station (BS) is located at the roadside while the vehicle V_1 and V_2 are in motion. There is a mobile user U in V_1 who requests to download a file locally cached at the BS. Vehicle V_2 also requests to download its own intended signal from BS. We assume that V_1 has also acquired the signal of V_2, as a result of the first transmission phase, during which the messages of V_1 and V_2 are transmitted simultaneously from the BS. For example, BS employs a NOMA technique to multiplex signals of V_1 and V_2 in the power domain. By involving the classical SIC, V_1 extracts the signal of V_2 in the spirit of cooperation. Another appropriate interpretation is related to the distribution of popular multimedia contents in VANET [23], using peer-to-peer protocols for exchanging popular packets through V2V channels.

Therefore, as shown in Fig. 1, cooperative inter-vehicle transmission is constructed during the second phase to enhance the reception reliability. Specifically, V_1 forwards the desired signal to V_2 for cooperatively enhancing the reception at V_2. Furthermore, the second phase scenario can be generalized to various situations. For example, user U can be a roadside unit (RSU), aiming for exchanging information with the onboard unit (OBU) of the vehicle V_1. While U may be a vehicle which is much closer to V_1 than V_2. Similar to the concept in [24], a VANET is formed among these vehicles for exchanging safety information, or for cooperatively distributing popular multimedia contents within a geographical area of interest. In general, our model is valid in a wide range of vehicular scenarios.

In the light of bandwidth scarcity, cognitive radio techniques can be exploited in the second stage to opportunistically exploit the spectrum holes in the licensed spectrum. For example, V_1 may be permitted to share the cellular uplink, for which the data traffic is typically lighter than for the downlink, hence resulting in potential spectrum wastage [25]. Basically, underlay cognitive transmission is feasible without traversing through the primary network. However, the interference imposed by V_1 on the BS in the second stage should be carefully managed, albeit this is beyond the scope of this article. Our main focus is on the second stage of the cooperative transmission in Fig. 1, since the performance in the first phase can be analysed similarly. Particularly, the NOMA-SM strategy is employed in the second stage for both V_1-V_2 and V_1-U links.

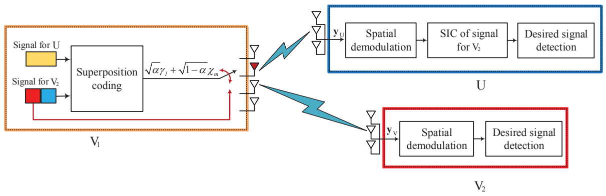

Fig. 2. The schematic diagram of the proposed NOMA-SM strategy.

The schematic diagram of NOMA-SM operated in the second stage is presented in Fig. 2, where V_1 assigns distinct transmit power to V_2 and U. The user access is based on NOMA, combined with SM. Although there is literature proposing multi-user SM schemes [26], [27], we use a classical SM designed for point-to-point transmission [6], [28] in vehicular environments. In what follows, we first elaborate on the principles of the proposed NOMA-SM scheme. Then our V2V channel model is detailed.

A. The Principles of NOMA-SM

Let us assume that N_t, N_r, and N_u omnidirectional antennas are employed at V_1, V_2, and U, respectively. As illustrated in Fig. 2, the proposed NOMA-SM strategy is applied both for the V_1-V_2 and V_1-U links. At the transmitter V_1, two independent bit streams are prepared for transmission. The bit stream for V_2 is portioned into two parts: the first $\log_2(N_t)$ bits are used for TA activation, activating a specific TA index n_t $(n_t \in \{1, \cdots, N_t\})$. The other $\log_2(M)$ bits destined for V_2 are combined with $\log_2(L)$ bits for U, employing superposition coding.

Subsequently, the modulated symbol $\sqrt{\alpha}\gamma_l + \sqrt{1-\alpha}\chi_m$ is radiated from the activated TA n_t, where γ_l and χ_m are intended for the in-car user U of V_1 and for V_2, respectively, satisfying $\mathbb{E}\{|\gamma_l|^2\} = \mathbb{E}\{|\chi_m|^2\} = E_s$, where E_s is the average energy per transmission at V_1, while α is the power allocation factor. According to the NOMA principle [19], the transmit power of the distant user in Fig. 2 must be higher than that of the close-by user, that is $(1-\alpha)E_s > \alpha E_s$. With this, $0 < \alpha < \frac{1}{2}$ should be guaranteed since the in-car user has a good channel. As a result, the block of $\log_2(N_t M L)$ bits unambiguously identify the active TA n_t and the superimposed complex symbol $\sqrt{\alpha}\gamma_l + \sqrt{1-\alpha}\chi_m$ transmitted from it. Hence a NOMA-SM super symbol can be expressed as

$$\mathbf{x} = \mathbf{e}_{n_t}\left(\sqrt{\alpha}\gamma_l + \sqrt{1-\alpha}\chi_m\right),$$

where $\mathbf{e}_{n_t}$ is the n_t-th column of the identity matrix $\mathbf{I}_{N_t}$, indicating that the n_t-th TA of V_1 is activated while the other $(N_t - 1)$ TAs are deactivated. Furthermore, χ_m is the m-th symbol in the M-ary amplitude-phase modulation (APM) used for V_1-V_2 transmission, while γ_l is the l-th symbol in the L-ary APM for V_1-U transmission.

Considering the propagation inside the vehicle V_1, we assume that the in-car user U experiences a frequency-flat Rayleigh channel. For example, the TAs of V_1 are installed on the central column of the vehicular dashboard, while the receive antennas (RAs) of U are placed behind the passenger front seat, without LoS from V_1. In [29], this scenario has been shown to be well suited to characterize diffuse scattering. Thus, we let $\mathbf{G} \in \mathbb{C}^{N_r \times N_t}$ denote the channel matrix between V_1 and U, and assume that all entries of $\mathbf{G}$ are independent identically distributed (i.i.d), obeying the distribution $\mathcal{CN}(0,1)$. The signal vector received at U and V_2 can be written as

$$\mathbf{y}_{\mathrm{U}} = \mathbf{g}_{n_t}\left(\sqrt{\alpha}\gamma_l + \sqrt{1-\alpha}\chi_m\right) + \mathbf{w}_{\mathrm{U}}, \tag{1}$$

$$\mathbf{y}_{\mathrm{V}} = \sqrt{p_0}\mathbf{h}_{n_t}\left(\sqrt{\alpha}\gamma_l + \sqrt{1-\alpha}\chi_m\right) + \mathbf{w}_{\mathrm{V}}, \tag{2}$$

respectively, where p_0 represents the average power drop between V_1 and V_2 due to the large scale fading. Furthermore, $\mathbf{g}_{n_t} \in \mathbb{C}^{N_u \times 1}$ is the n_t-th column of $\mathbf{G}$, representing the channel vector between U and the n_t-th TA of V_1, while $\mathbf{h}_{n_t} \in \mathbb{C}^{N_r \times 1}$ is the n_t-th column of the V2V channel matrix $\mathbf{H} \in \mathbb{C}^{N_r \times N_t}$, indicating the received complex fading envelope between V_2 and the n_t-th TA of V_1. Finally, $\mathbf{w}_{(\cdot)}$ denotes a complex additive white Gaussian noise (AWGN) vector with a power spectrum density of σ_0^2 per entry. For the inter-vehicle channel, the path loss is considerable in (2), while it is neglected between the in-car user and the antenna array of V_1.

In our system, the transmitter and both receivers are assumed to have perfect synchronization in both time and frequency. Full channel state information is assumed to be available at receivers (i.e., CSIR). In principle, both V_2 and U first have to detect the signal destined for V_2, i.e., the activated TA index $\hat{n}_t$ and the APM symbol $\chi_{\hat{m}}$ at each particular time instant. The corresponding optimum maximum likelihood (ML) detector is invoked at U and V_2 according to

$$\left(\hat{n}_t, \chi_{\hat{m}}\right) = \arg\min_{n_t, m} \left\|\mathbf{y}_U - \sqrt{1-\alpha}\mathbf{g}_{n_t}\chi_m\right\|^2, \tag{3}$$

$$\left(\hat{n}_t, \chi_{\hat{m}}\right) = \arg\min_{n_t, m} \left\|\mathbf{y}_V - \sqrt{p_0\left(1-\alpha\right)}\mathbf{h}_{n_t}\chi_m\right\|^2. \tag{4}$$

After eliminating the interference imposed by $\left(\hat{n}_t, \chi_{\hat{m}}\right)$ on $\mathbf{y}_{\mathrm{U}}$, U becomes capable of performing another ML detection to acquire the desired signal $\gamma_{\hat{l}}$.

B. V2V Massive MIMO Channel Model

In contrast to the conventional fixed-to-mobile cellular radio systems, in V2V systems, both the transmitter and receiver are in motion and both are equipped with low-elevation antennas, which will result in quite different propagation conditions. Hence a non-isotropic scattering V2V stochastic model was proposed in [30] for characterizing a wide variety of V2V scenarios by adjusting relevant model parameters. In [11], a novel three-dimensional V2V geometry-based stochastic channel was proposed for accurately capturing the effect of vehicular traffic density on the channel. In this article, we consider a spatio-temporally correlated Rician channel model for characterizing our narrowband V2V massive MIMO channel, which has also been exploited in [13] and [31]. We describe the underlying V2V channel as a matrix of complex fading envelopes, i.e., $\mathbf{H} \in \mathbb{C}^{N_t \times N_r}$, which can be expressed as

$$\mathbf{H} = \sqrt{\frac{K}{K+1}}\bar{\mathbf{H}} + \sqrt{\frac{1}{K+1}}\tilde{\mathbf{H}},$$

where K is the Rician factor, while $\bar{\mathbf{H}}$ is the fixed part related to the LoS component. Furthermore, $\tilde{\mathbf{H}}$ represents the variable part, whose entries are correlated complex Gaussian variables. Given $\left[\tilde{\mathbf{H}}\right]_{p,q} = \tilde{h}_{p,q}$, we assume that

$$\begin{aligned}\mathbb{E}\left\{\tilde{h}^R_{p,q}\tilde{h}^R_{\hat{p},\hat{q}}\right\} &= \mathbb{E}\left\{\tilde{h}^I_{p,q}\tilde{h}^I_{\hat{p},\hat{q}}\right\},\\ \mathbb{E}\left\{\tilde{h}^R_{p,q}\tilde{h}^I_{\hat{p},\hat{q}}\right\} &= \mathbb{E}\left\{\tilde{h}^I_{p,q}\tilde{h}^R_{\hat{p},\hat{q}}\right\} = 0,\end{aligned}$$

where $p, \hat{p} \in \{1, \cdots, N_r\}$ and $q, \hat{q} \in \{1, \cdots, N_t\}$. Explicitly, for each $\tilde{h}_{p,q}$, the auto-correlations of the real and imaginary parts are identical and the cross-correlations between real and imaginary parts are equal to zero. Hence the correlated channel matrix $\tilde{\mathbf{H}}$ can be described by the widely-used Kronecker correlation model [32], which is expressed as

$$\tilde{\mathbf{H}} = \mathbf{\Sigma}_r^{\frac{1}{2}}\hat{\mathbf{H}}\mathbf{\Sigma}_t^{\frac{1}{2}}.$$

Here $\mathbf{\Sigma}_t \in \mathbb{C}^{N_t \times N_t}$ and $\mathbf{\Sigma}_r \in \mathbb{C}^{N_r \times N_r}$ are the correlation matrices at V_1 and V_2, respectively, with the elements defined as $[\mathbf{\Sigma}_t]_{q,\hat{q}} = \sigma^t_{q,\hat{q}}$ for $q, \hat{q} \in \{1, \cdots, N_t\}$, and $[\mathbf{\Sigma}_r]_{p,\hat{p}} = \sigma^r_{p,\hat{p}}$ for $p, \hat{p} \in \{1, \cdots, N_r\}$. Furthermore, $\hat{\mathbf{H}}$ is the independent Rayleigh channel matrix whose entries are i.i.d complex Gaussian random variables, i.e., $[\hat{\mathbf{H}}]_{p,q} = \hat{h}_{p,q} \sim \mathcal{CN}(0,1)$. Specifically, the correlation matrices $\mathbf{\Sigma}_t$ and $\mathbf{\Sigma}_r$ can be determined according to a concrete model. Here the exponential model of Loyka [33] is adopted and the correlation matrix entries are formed as $\sigma^t_{q,\hat{q}} = \kappa_t^{|q-\hat{q}|}$ and $\sigma^r_{p,\hat{p}} = \kappa_r^{|p-\hat{p}|}$, where κ_t and κ_r are the adjacent antenna correlation coefficients at V_1 and V_2, respectively.

In order to mimic the influence of the V2V channel's time-varying effects, we take the temporal correlation into consideration, which is defined as

$$\delta\left(\tau\right) = \mathbb{E}\left\{\hat{\mathbf{H}}\left(t\right)\hat{\mathbf{H}}\left(t+\tau\right)\right\},$$

where τ is the sampling time. In [13], Jakes' model is used for describing the temporal correlation expressed as $\delta\left(\tau\right) = J_0\left(2\pi f_D \tau\right)$, where f_D is the maximum Doppler frequency related to both the carrier frequency and the velocity of the terminal. For simplicity of analysis, in the following we omit the index τ. Observe that $\delta = 1$ indicates that the underlying V2V channel is quasi-static, while $\delta < 1$ is related to a time-varying channel due to mobility. Naturally, both the spatial and temporal correlations would affect the performance of the receivers.

III. Capacity Analysis of the NOMA-SM System

Recall that the proposed NOMA-SM transmission scheme relies on a pair of independent spaces: the classical signal-domain, pertaining to the radiated superimposed symbol $\sqrt{\alpha}\gamma_l + \sqrt{1-\alpha}\chi_m$, and the TA-domain, pertaining to the activated TA index n_t. More specifically, the message intended for V_2 is conveyed by both of the two streams. While the message destined for U is only mapped to the classical signal-domain, superimposed with part of V_2's signal in the power domain. In what follows, we investigate the capacity of the collaboration-aided vehicle V_2 and the in-car user U. Monte Carlo estimates are also provided for MI evaluation, followed by an illustrative example to augment the theoretical analysis.

A. Capacity Analysis of the Collaboration-Aided Vehicle V_2

In the NOMA protocol, the transmit power assigned by V_1 to the distant user V_2 has to be higher than that to the close-by user U. Then the distant user directly detects its signal, since the interference induced by the close-by user is smaller and can thus be treated as background noise. Considering that all TAs of V_1 are activated with the same probability for NOMA-SM, the instantaneous capacity pertaining to the classical signal-domain of V2V transmission is given by

$$\begin{aligned} C_V^{sig} &= \max_{f_\chi} I\left(\chi; \mathbf{y}_V \left| n_t \right.\right) \\ &= \tfrac{1}{N_t} \sum_{i=1}^{N_t} \log_2 \left(\tfrac{E_s p_0 \|\mathbf{h}_i\|^2 + \sigma_0^2}{\alpha E_s p_0 \|\mathbf{h}_i\|^2 + \sigma_0^2} \right). \end{aligned} \tag{5}$$

Observe that no practical modulation constellation is assumed, when performing these capacity calculations. Since the channel capacity relates to the highest rate in bits per channel use at which information can be sent with arbitrarily low probability of error, in (5), we substitute χ_m by χ, which denotes a random input signal alphabet with a distribution of f_χ. On the other hand, the MI conveyed by the spatial-domain TA-constellations can be written as

$$I\left(n_t; \mathbf{y}_V\right) = \frac{1}{N_t} \sum_{i=1}^{N_t} \int \Pr\left(\mathbf{y}_V \middle| \mathbf{h}_i\right) \log_2 \frac{\Pr\left(\mathbf{y}_V \middle| \mathbf{h}_i\right)}{\Pr\left(\mathbf{y}_V\right)} \mathrm{d}\mathbf{y}_V, \tag{6}$$

where $\Pr\left(\mathbf{y}_V \left| \mathbf{h}_i \right.\right)$ denotes the probability density function (PDF) of the channel output $\mathbf{y}_V$ received over the i-th channel vector of $\mathbf{H}$, given by

$$\Pr\left(\mathbf{y}_V \left| \mathbf{h}_i \right.\right) = \frac{1}{\pi^{N_r} \det\left(\mathbf{\Sigma}_i\right)} \exp\left\{ -\mathbf{y}_V^H \mathbf{\Sigma}_i^{-1} \mathbf{y}_V \right\},$$

where $\mathbf{\Sigma}_i = \sigma_0^2 \mathbf{I} + p E_s \mathbf{h}_i {\mathbf{h}_i}^H$. As a result, the instantaneous capacity of V_2 in the NOMA-SM system is formulated as

$$C_V = C_V^{sig} + I\left(n_t; \mathbf{y}_V\right). \tag{7}$$

Remark: It is worth noting that in (5), C_V^{sig} is achievable where the optimum input distribution for χ is Gaussian. In fact, this optimum input distribution is also regarded as the optimum input distribution for a conventional SM system. This is a common assumption in the majority of SM capacity-related contributions [6], [34]–[36], effectively simplifying the analysis. Nevertheless, a fundamental weakness of the Gaussian input assumption is that f_χ affects both $I\left(\chi; \mathbf{y}_V \left| n_t \right.\right)$ and $I\left(n_t; \mathbf{y}_V\right)$. Clearly, the Gaussian input distribution maximizes $I\left(\chi; \mathbf{y}_V \left| n_t \right.\right)$, but it is unclear whether it maximizes $I\left(n_t; \mathbf{y}_V\right)$. In addition, the equiprobable activation of antennas is a widely accepted assumption for SM-enabled systems, albeit this activation regime cannot guarantee the optimal spatial design capable of achieving the capacity in the TA-domain. Actually, Liu *et al.* in [34] studied the optimal antenna activation required for TA-domain capacity maximization. Moreover, Basnayaka *et al.* [8] have demonstrated that the Gaussian input does not achieve the upper limit of the MI provided by an SM-aided system. As a further insight, although the MI conveyed by the TA-domain cannot be formulated as an analytical expression, we are inspired to derive the capacity upper bound and to conceive the associated power allocation optimization schemes, which will be addressed in Section V.

Below we will derive the ergodic capacity of V_2. To calculate the specific part pertaining to the signal-domain, i.e. $\mathbb{E}_{\mathbf{H}}\{C_V^{sig}\}$, we first introduce the notation $\psi = \|\mathbf{h}_i\|^2$. Since it is non-trivial to explicitly formulate the PDF $f_\Psi\left(\psi\right)$ with $\mathbf{H}$ being a spatio-temporally correlated Rician fading channel, we set out to simplify the channel model. Explicitly, $\mathbf{H}$ is temporarily approximated by an uncorrelated Rician channel matrix, yielding $\mathbf{\Sigma}_t = \mathbf{\Sigma}_r = \mathbf{I}$. Hence ψ obeys the non-central Chi-square distribution with a degree of freedom $2N_r$. Then the PDF of ψ may be expressed as [37]

$$f_\Psi\left(\psi\right) = e^{-\left(\psi+\lambda\right)} \left(\frac{\psi}{\lambda}\right)^{\left(N_r-1\right)/2} I_{N_r-1}\left(\sqrt{\lambda\psi}\right),$$

where $\lambda = N_r K$ is termed as the non-centrality parameter and $I_v\left(y\right)$ is the modified Bessel function of the first kind, which is a built-in function in popular mathematical software packages, such as MATLAB or Mathematica. Therefore, the ergodic capacity of V_2 pertaining to the signal-domain is given in analytical form as

$$\mathbb{E}_{\mathbf{H}}\{C_V^{sig}\} = \int_0^\infty \log_2 \left(\frac{E_s p_0 \psi + \sigma_0^2}{\alpha E_s p_0 \psi + \sigma_0^2} \right) f_\Psi\left(\psi\right) \mathrm{d}\psi,$$

where the integral can be evaluated via numerical integration.

On the other hand, since there is no closed-form expression for the mutual information of SM systems [34], it is a challenge to derive the ergodic capacity related to the spatial domain in closed form. As a remedy, we can resort to a Monte Carlo estimate of (6), which is given by

$$I\left(n_t; \mathbf{y}_V\right) \approx \frac{1}{N_t S} \sum_{i=1}^{N_t} \sum_{s=1}^{S} \log_2 \frac{\Pr\left(\mathbf{y}_V^s \middle| \mathbf{h}_i\right)}{\sum\limits_{j=1}^{N_t} \Pr\left(n_t = j\right) \Pr\left(\mathbf{y}_V^s \middle| \mathbf{h}_j\right)}, \tag{8}$$

where $\mathbf{y}_V^s$ associated with $s = 1, \cdots, S$ represents i.i.d random samples drawn from $\mathbf{y}_V$. The value of S should be sufficiently high to guarantee the statistically relevant evaluation of $I\left(n_t; \mathbf{y}_V\right)$. Subsequently, by averaging $I\left(n_t; \mathbf{y}_V\right)$ over multiple channel realizations and adding it to $\mathbb{E}_{\mathbf{H}}\{C_V^{sig}\}$, the ergodic capacity of V_2 for the uncorrelated Rician fading channel is obtained.

B. Capacity Analysis of the In-Car User U

In contrast to the receiver of V_2, the receiver of U can detect its own signal after removing the interference imposed by V_2, as seen in Fig. 2. To demonstrate the feasibility of this SIC procedure, we first deduce the maximum rate of which U can detect the message of V_2. Specifically, the maximum rate for U detecting the message related to the classical signal-domain of V_2 is given by

$$C_U^{V,sig} = \frac{1}{N_t}\sum_{i=1}^{N_t}\log_2\left(\frac{E_s\|\mathbf{g}_i\|^2+\sigma_0^2}{\alpha E_s\|\mathbf{g}_i\|^2+\sigma_0^2}\right). \tag{9}$$

The MI associated with U detecting the information embedded in the TA-constellation of V_2 can be written as

$$I\left(n_t;\mathbf{y}_U\right) = \frac{1}{N_t}\sum_{i=1}^{N_t}\int \Pr\left(\mathbf{y}_U|\,\mathbf{g}_i\right)\log_2\frac{\Pr\left(\mathbf{y}_U|\,\mathbf{g}_i\right)}{\Pr\left(\mathbf{y}_U\right)}\mathrm{d}\mathbf{y}_U, \tag{10}$$

where $\Pr\left(\mathbf{y}_U\,|\mathbf{g}_i\right)$ denotes the PDF of the channel output $\mathbf{y}_U$ received over the i-th channel vector of $\mathbf{G}$ given by

$$\Pr\left(\mathbf{y}_U\,|\mathbf{g}_i\right) = \frac{1}{\pi^{N_u}\det\left(\mathbf{\Omega}_i\right)}\exp\left\{-\mathbf{y}_U^H\mathbf{\Omega}_i^{-1}\mathbf{y}_U\right\},$$

where $\mathbf{\Omega}_i = \sigma_0^2\mathbf{I}+E_s\mathbf{g}_i{\mathbf{g}_i}^H$. Note that a Monte Carlo estimate to (10) can be performed similarly as in (8), though we do not explicitly present here due to the space limitation.

As a result, the instantaneous capacity for U detecting the signal of V_2 can be expressed as

$$C_U^V = C_U^{V,sig} + I\left(n_t;\mathbf{y}_U\right). \tag{11}$$

It may be readily seen that $C_U^V > C_V$ is always satisfied, since $\|\mathbf{g}_i\|^2 > p_0\|\mathbf{h}_i\|^2$, guaranteeing the success of SIC. Hence the capacity of U detecting its own desired signal is written as

$$\begin{aligned} C_U &= \max_{f_\gamma} I\left(\gamma;\mathbf{y}_U|\,n_t,\chi,\mathbf{G}\right) \\ &= \tfrac{1}{N_t}\sum_{i=1}^{N_t}\log_2\left(1+\tfrac{\alpha E_s}{\sigma_0^2}\|\mathbf{g}_i\|^2\right), \end{aligned} \tag{12}$$

where γ denotes the random input signal variable related to the desired message of U, with a distribution of f_γ. The capacity for U detecting γ indeed becomes achievable when the channel's input distribution f_γ is Gaussian.

To formulate the ergodic capacity of U, we temporarily introduce the notation $\varphi = \|\mathbf{g}_i\|^2$. Based on the assumption that each entry of $\mathbf{G}$ obeys an i.i.d zero-mean unit-variance Gaussian distribution, φ obeys the central Chi-square distribution with the degree of $2N_u$. The PDF $f_\Phi\left(\varphi\right)$ is given by [38]

$$f_\Phi\left(\varphi\right) = \frac{1}{\left(N_u-1\right)!}\varphi^{N_u-1}e^{-\varphi}.$$

Therefore, the ergodic capacity of U is given in analytical form at the top of the next page, where $\mathrm{Ei}\left(x\right) = \int_{-\infty}^{x}\frac{e^u}{u}\mathrm{d}u, x<0$ is the exponential integral function.

C. Mutual Information

To appreciate the above theoretical analysis in terms of its relevance, next we characterize the bandwidth efficiency of the proposed NOMA-SM. Assuming perfect knowledge of the instantaneous channel state information at both receivers, the MI achieved by V_2 and U with the aid of practical APM constellations is evaluated by the classical Monte Carlo method. For the collaboration-aided vehicle V_2, the MI between a discrete signal input (n_t,χ_m) and the received signal $\mathbf{y}_V$ can be formulated as

$$\begin{aligned} &I\left(n_t,\chi_m;\mathbf{y}_V|\,\mathbf{H}\right) = \mathbb{E}_{n_t,\chi_m,\mathbf{y}_V}\left\{\log_2\tfrac{\Pr\left(\mathbf{y}_V|n_t,\chi_m,\mathbf{H}\right)}{\Pr\left(\mathbf{y}_V|\mathbf{H}\right)}\right\} \\ &= \tfrac{1}{N_tM}\times\int\Pr\left(\mathbf{y}_V|\,\chi_m,\mathbf{h}_i\right)\log_2\tfrac{\Pr\left(\mathbf{y}_V|\chi_m,\mathbf{h}_i\right)}{\Pr\left(\mathbf{y}_V|\mathbf{H}\right)}\mathrm{d}\mathbf{y}_V, \end{aligned} \tag{13}$$

where the conditional probability $\Pr\left(\mathbf{y}_V|\,\chi_m,\mathbf{h}_i\right)$ is expressed as

$$\begin{aligned} \Pr(\mathbf{y}_V|\,\chi_m,\mathbf{h}_i) = &\tfrac{1}{\pi^{N_r}\det(\mathbf{\Psi}_i)}\exp\Big\{-\left(\mathbf{y}_V-\sqrt{p_0\left(1-\alpha\right)}\mathbf{h}_i\chi_m\right)^H \\ &\times\mathbf{\Psi}_i^{-1}\left(\mathbf{y}_V-\sqrt{p_0\left(1-\alpha\right)}\mathbf{h}_i\chi_m\right)\Big\}, \end{aligned}$$

with $\mathbf{\Psi}_i = \sigma_0^2\mathbf{I}+\alpha p_0E_s\mathbf{h}_i\mathbf{h}_i^H$. With regard to the in-car user U performing SIC first, the MI between the information input (n_t,χ_m) and the received signal $\mathbf{y}_U$ is given by

$$\begin{aligned} &I\left(n_t,\chi_m;\mathbf{y}_U|\,\mathbf{G}\right) = \tfrac{1}{N_tM}\times \\ &\int\Pr\left(\mathbf{y}_U|\,\chi_m,\mathbf{g}_i\right)\log_2\tfrac{\Pr\left(\mathbf{y}_U|\chi_m,\mathbf{g}_i\right)}{\Pr\left(\mathbf{y}_U|\mathbf{G}\right)}\mathrm{d}\mathbf{y}_U, \end{aligned} \tag{14}$$

where the conditional probability $\Pr\left(\mathbf{y}_U|\,\chi_m,\mathbf{g}_i\right)$ is expressed as

$$\begin{aligned} &\Pr\left(\mathbf{y}_U|\,\chi_m,\mathbf{g}_i\right) = \tfrac{1}{\pi^{N_u}\det(\mathbf{\Phi}_i)}\times \\ &\exp\Big\{-\left(\mathbf{y}_U-\sqrt{1-\alpha}\mathbf{g}_i\chi_m\right)^H\mathbf{\Phi}_i^{-1}\left(\mathbf{y}_U-\sqrt{1-\alpha}\mathbf{g}_i\chi_m\right)\Big\}, \end{aligned}$$

with $\mathbf{\Phi}_i = \sigma_0^2\mathbf{I}+\alpha E_s\mathbf{g}_i\mathbf{g}_i^H$.

Subsequently, the MI between the information input γ_l and the received signal $\mathbf{y}_U$ after perfect SIC is expressed as

$$\begin{aligned} &I\left(\gamma_l;\tilde{\mathbf{y}}_U|\,\mathbf{g}_{n_t}\right) = \tfrac{1}{N_tL}\times \\ &\int\Pr\left(\tilde{\mathbf{y}}_U|\,\gamma_l,\mathbf{g}_i\right)\log_2\tfrac{\Pr\left(\tilde{\mathbf{y}}_U|\gamma_l,\mathbf{g}_i\right)}{\frac{1}{N_tL}\sum\limits_{k,j}\Pr\left(\tilde{\mathbf{y}}_U|\gamma_k,\mathbf{g}_j\right)}\mathrm{d}\tilde{\mathbf{y}}_U, \end{aligned} \tag{15}$$

where $\tilde{\mathbf{y}}_U = \mathbf{y}_U-\sqrt{1-\alpha}\mathbf{g}_i\chi_m$ with $i\in\{1,\cdots,N_t\}$ and $m\in\{1,\cdots,M\}$ denotes the received vector after SIC. The conditional probability $\Pr\left(\tilde{\mathbf{y}}_U|\,\gamma_l,\mathbf{g}_i\right)$ is given by

$$\Pr(\tilde{\mathbf{y}}_U|\,\gamma_l,\mathbf{g}_i) = \frac{1}{\left(\pi\sigma_0^2\right)^{N_u}}\exp\left\{-\frac{\|\tilde{\mathbf{y}}_U-\sqrt{\alpha}\mathbf{g}_i\gamma_l\|^2}{\sigma_0^2}\right\}.$$

D. An Illustration

In this part, a simulation based study of our theoretical expressions is provided with the aid of the MI attained by practical APM constellations. We set $N_t = 64$, $N_r = N_u = 2$ for our MIMO configurations in conjunction with $\alpha = 0.1$, $E_s = 1$ and $p_0 = 10^{-3}$ are given. The channel matrix $\mathbf{H}$ is generated according to Section II-B, where $K = 0.2$, $\kappa_t = \kappa_r = 0.5$, and $\delta = 1$ are used. Each entry of $\mathbf{G}$ is identically and independently generated according to a complex Gaussian distribution $\mathcal{CN}\left(0,1\right)$. In our Monte Carlo evaluations, the 16PSK signal constellation is chosen as the

$$
\begin{aligned}
&\mathbb{E}_{\mathbf{G}}\{C_U\} = \int_0^\infty \log_2\left(1+\frac{\alpha E_s}{\sigma_0^2}\varphi\right) f_\Phi(\varphi)\,\mathrm{d}\varphi \\
&= \begin{cases} -\frac{1}{\ln 2}\mathrm{Ei}\left(-\frac{\sigma_0^2}{\alpha E_s}\right)\exp\left(\frac{\sigma_0^2}{\alpha E_s}\right), N_u = 1 \\ \frac{1}{\ln 2}\sum\limits_{n=0}^{N_u-1}\frac{1}{(N_u-n-1)!}\left[(-1)^{N_u-n}\left(\frac{\sigma_0^2}{\alpha E_s}\right)^{N_u-n-1}\mathrm{Ei}\left(-\frac{\sigma_0^2}{\alpha E_s}\right)\exp\left(\frac{\sigma_0^2}{\alpha E_s}\right) + \sum\limits_{m=1}^{N_u-n-1}(m-1)!\left(-\frac{\sigma_0^2}{\alpha E_s}\right)^{N_u-n-m-1}\right], N_u > 1 \end{cases}
\end{aligned}
$$

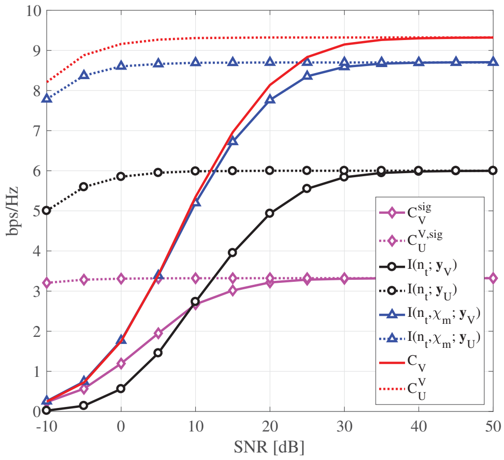

Fig. 3. Capacity and MI performance for $N_t = 64$, $N_r = N_u = 2$, $M = L = 16$, and $\alpha = 0.1$. Specifically, C_V^{sig}, $C_U^{V,sig}$, $I(n_t;\mathbf{y}_V)$, and $I(n_t;\mathbf{y}_U)$ are obtained from (5), (9), (6), and (10), respectively. While $I(n_t,\chi_m;\mathbf{y}_V)$ and $I(n_t,\chi_m;\mathbf{y}_U)$ are generated from (13) and (14), respectively, after averaging over multiple channel realizations. Finally, C_V and C_U^V are evaluated from (7) and (11), respectively.

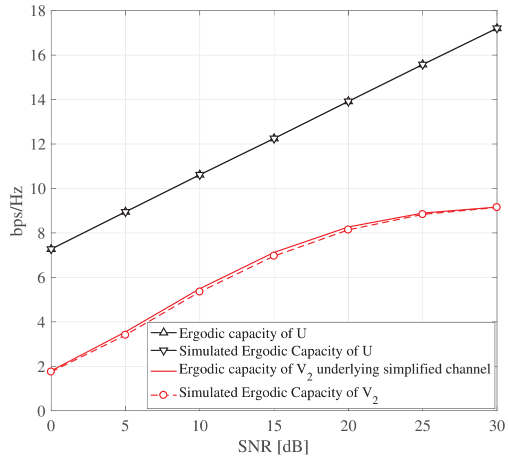

Fig. 4. Ergodic capacity for $N_t = 64$, $N_r = N_u = 2$, $M = L = 16$, and $\alpha = 0.1$. Specifically, the simulated ergodic capacity for both V_2 and U are obtained by averaging the instantaneous capacity over multiple channel realizations.

APM for χ_m and γ_l, hence we have $M = L = 16$. The effective transmit signal-to-noise ratio (SNR) at V_1 is given by $p_0 E_s/\sigma_0^2$ as the horizontal axis of Fig. 3. Notice that, the transmit-SNR cannot be readily interpreted physically, because it relates the transmitter power to the noise power at the receiver, but its notion is convenient to use in NOMA-aided scenarios. Given the effective transmit-SNR at V_1 as $\mathrm{SNR} = p_0 E_s/\sigma_0^2$, the average receive-SNR at V_2 can be computed as

$$\mathrm{SNR}_r^{V_2} = \frac{(1-\alpha)\,\mathrm{SNR}}{1+\alpha\mathrm{SNR}}.$$

Similarly, the average receive-SNR at U for detecting the signal of V_2 and that of itself are respectively expressed as

$$\mathrm{SNR}_r^{U,V_2} = \frac{(1-\alpha)\mathrm{SNR}}{p_0+\alpha\mathrm{SNR}},$$
$$\mathrm{SNR}_r^{U} = \frac{\alpha\mathrm{SNR}}{p_0}.$$

Hence the effective transmit SNR at V_1 is unambiguously related to the SNRs at each receiver. Furthermore, we use $\mathrm{SNR} = p_0 E_s/\sigma_0^2$ in all of the subsequent performance analyses. The relevant results of Fig. 3 are discussed as follows.

- The capacity of V_2 gleaned from the signal-domain, that is C_V^{sig} obtained from (5), increases steadily upto a saturation point as the SNR increases. By contrast, the capacity for U detecting the signal-domain destined for V_2, i.e., $C_U^{V,sig}$ obtained from (9) is higher than C_V^{sig} in the low and moderate SNR domain. Clearly, a successful detection of the signal-domain of V_2 can be performed by U.
- The MI $I(n_t;\mathbf{y}_V)$ generated using (6) increases with the SNR and saturates at 6 bps/Hz, since the input entropy of the TA-domain space is $\log_2(N_t)$. By contrast, the MI $I(n_t;\mathbf{y}_U)$ attained by (10) is as high as 6 bps/Hz across almost the entire SNR range, since the channel quality of U is much higher than that of V_2, implying that U can successfully detect the signal of V_2 embedded in the TA-domain.
- The capacity of V_2, i.e., C_V grows steadily as the SNR increases upto its saturation at high SNRs, but it remains lower than C_U^V. Since C_V is obtained by the summation of C_V^{sig} and $I(n_t;\mathbf{y}_V)$, and C_U^V equals to the sum of $C_U^{V,sig}$ and $I(n_t;\mathbf{y}_U)$. Naturally, $C_U^V > C_V$ is satisfied, as $C_U^{V,sig}$ and $I(n_t;\mathbf{y}_U)$ are higher than C_V^{sig} and $I(n_t;\mathbf{y}_V)$, respectively. Therefore, U can always perform successful SIC.
- The MI curves $I(n_t,\chi_m;\mathbf{y}_V)$ and $I(n_t,\chi_m;\mathbf{y}_U)$ are generated from (13) and (14), respectively, after averaging over multiple channel realizations. It may be observed that the simulated curve $I(n_t,\chi_m;\mathbf{y}_V)$ matches the analytical capacity C_V quite closely upto an SNR of 5 dB, but beyond that $I(n_t,\chi_m;\mathbf{y}_V)$ starts to drift away from C_V. By contrast, the drift of $I(n_t,\chi_m;\mathbf{y}_U)$ from C_U^V remains nearly unchanged. Both drifts are due to the fact that the MI attained with the aid of practical APM modulation is upper bounded by the capacity, namely by the maximum data rate related to the optimal input distribution.

Moreover, we depict the ergodic capacity of both receivers in Fig. 4. Clearly, the simulated ergodic capacity of U obtained from (12) after averaging over multiple channel realizations is perfectly matched with the exact one. As for V_2, the ergodic capacity gap between the simplified channel model and the original one are shown to be modest within the low- and high-SNR regimes. Therefore, the ergodic capacity derived for V_2 gives a good approximation of that for a spatio-temporally correlated Rician channel.

IV. Power Allocation Algorithms

It has been demonstrated that the MI conveyed by the TA-domain cannot be readily formulated as a closed-form expression, only by resorting to simulations. Thus, it is very hard to perform an optimal power allocation for NOMA-SM. To circumvent this problem, we first derive an upper bound of the NOMA-SM capacity. Then the power allocation, which is capable of maximizing the capacity bound is considered, leading to the optimal solution.

A. Problem Formulation

Theoretically, the instantaneous capacity of V_2 in the NOMA-SM system can be expressed as

$$C_V = \max_{f_\chi} I\left(n_t, \chi; \mathbf{y}_V\right) = \max_{f_\chi} h\left(\mathbf{y}_V\right) - h\left(\mathbf{y}_V \left| n_t, \chi\right.\right), \tag{16}$$

where $h\left(\cdot\right)$ denotes the differential entropy. The conditional differential entropy $h\left(\mathbf{y}_V \left| n_t, \chi\right.\right)$ in (16) is explicitly given by

$$h\left(\mathbf{y}_V \left| n_t, \chi\right.\right) = \frac{1}{N_t}\sum_{i=1}^{N_t} \log_2 \det\left[\pi e\left(p_0 \alpha E_s \mathbf{h}_i \mathbf{h}_i^H + \sigma_0^2 \mathbf{I}\right)\right].$$

To determine C_V, we have to evaluate $h\left(\mathbf{y}_V\right)$, which requires the knowledge of the distribution of $\mathbf{y}_V$. It may be readily seen that the MI $I\left(n_t, \chi; \mathbf{y}_V\right)$ is maximized if the vector variable $\mathbf{y}_V$ has a Gaussian distribution. Thus, we assume that the received vector $\mathbf{y}_V$ has a Gaussian distribution, which is a zero-mean vector having a covariance matrix presented as

$$\begin{aligned}
&\mathbb{E}\left\{\mathbf{y}_V \mathbf{y}_V^H\right\} = \mathbf{H}\mathbb{E}_{n_t}\left\{\mathbf{e}_{n_t}\mathbb{E}_\chi\left\{p_0\left(1-\alpha\right)\chi\chi^*\right\}\mathbf{e}_{n_t}^H\right\}\mathbf{H}^H \\
&+\mathbf{H}\mathbb{E}_{n_t}\left\{\mathbf{e}_{n_t}\mathbb{E}_\gamma\left\{p_0\alpha\gamma\gamma^*\right\}\mathbf{e}_{n_t}^H\right\}\mathbf{H}^H + \sigma_0^2\mathbf{I} \\
&= \mathbf{H}\left\{\tfrac{1}{N_t}\sum_{i=1}^{N_t}\mathbf{e}_i\mathbf{e}_i^H p_0\left(1-\alpha\right)E_s\right\}\mathbf{H}^H \\
&+\mathbf{H}\left\{\tfrac{1}{N_t}\sum_{n_t=1}^{N_t}\mathbf{e}_i\mathbf{e}_i^H p_0\alpha E_s\right\}\mathbf{H}^H + \sigma_0^2\mathbf{I} \\
&= \tfrac{p_0 E_s}{N_t}\mathbf{H}\mathbf{H}^H + \sigma_0^2\mathbf{I}.
\end{aligned}$$

An upper bound of $h\left(\mathbf{y}_V\right)$ can be formulated as

$$h\left(\mathbf{y}_V\right) \le \log_2 \det\left(\pi e\left(\frac{p_0 E_s}{N_t}\mathbf{H}\mathbf{H}^H + \sigma_0^2\mathbf{I}\right)\right).$$

Hence we obtain an upper bound of C_V which is written as

$$\begin{aligned}
C_V &\le \log_2 \det\left(\pi e\left(\tfrac{p_0 E_s}{N_t}\mathbf{H}\mathbf{H}^H + \sigma_0^2\mathbf{I}\right)\right) \\
&\quad -\tfrac{1}{N_t}\sum_{i=1}^{N_t}\log_2\det\left(\pi e\left(p_0\alpha E_s\mathbf{h}_i\mathbf{h}_i^H + \sigma_0^2\mathbf{I}\right)\right) \\
&= \sum_{j=1}^{N_r}\log_2\left(\tfrac{p_0 E_s}{N_t}\lambda_j^2 + \sigma_0^2\right) \\
&\quad -\tfrac{1}{N_t}\sum_{i=1}^{N_t}\log_2\left(p_0\alpha E_s\left\|\mathbf{h}_i\right\|^2 + \sigma_0^2\right) \triangleq C_V^{B_1},
\end{aligned} \tag{17}$$

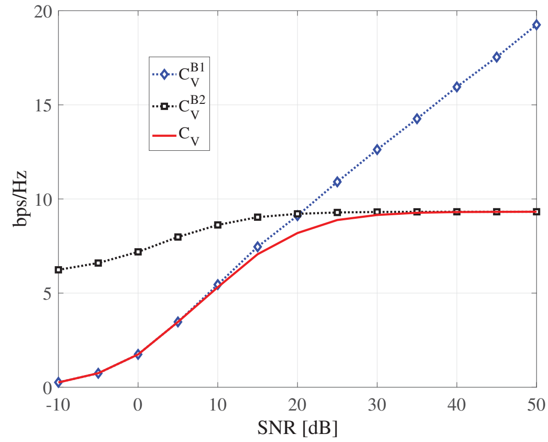

Fig. 5. Capacity and two upper bounds of the V2V transmission link with $N_t = 64$, $N_r = 2$, $M = 16$, and $\alpha = 0.1$. Specifically, C_V, $C_V^{B_1}$, and $C_V^{B_2}$ are evaluated from (7), (17), and (18), respectively.

where λ_j is the j-th singular value of $\mathbf{H}$ with $j \in \{1, \cdots, N_r\}$. Clearly, $C_V^{B_1}$ has N_r DoFs and it is the same as the capacity of an $(N_t \times N_r)$-element spatially multiplexed MIMO system, subject to inter-user interference.

On the other hand, the MI of the TA-domain has a natural upper bound written as

$$I\left(n_t; \mathbf{y}_V\right) \le \log_2\left(N_t\right),$$

which corresponds to the maximum MI that can be conveyed by the TA-domain of the V2V transmission link. Now, another upper bound of C_V may also be formulated as

$$\begin{aligned}
C_V &\le C_V^{sig} + \log_2\left(N_t\right) \\
&= \tfrac{1}{N_t}\sum_{i=1}^{N_t}\log_2\left(\tfrac{E_s p_0\|\mathbf{h}_i\|^2+\sigma_0^2}{\alpha E_s p_0\|\mathbf{h}_i\|^2+\sigma_0^2}\right) + \log_2\left(N_t\right) \triangleq C_V^{B_2}.
\end{aligned} \tag{18}$$

Before proceeding, we provide a numerical illustration in order to evaluate both of the upper bounds on the capacity of V_2. Figure 5 depicts C_V and both upper bounds of the NOMA-SM system in conjunction with $N_t = 64$, $N_r = 2$, $M = 16$, and $\alpha = 0.1$, which exhibit distinct approximations of C_V within certain SNR regions. The upper bound $C_V^{B_1}$ gives a tight bound of C_V at low SNRs, indicating that the NOMA-SM capacity at V_2 is almost the same as that of a spatially multiplexed MIMO system of the same configuration in the presence of inter-user interference. However, the MI embedded in the TA-domain saturates as the SNR increases, which is due to the fact that N_t is finite. Hence, at high SNRs, $C_V^{B_2}$ is much tighter.

Based on the above observations, a refined upper bound on the capacity of V_2 in the NOMA-SM system is represented as

$$C_V^B \triangleq \min\left(C_V^{B_1}, C_V^{B_2}\right). \tag{19}$$

Considering the QoS of the two receivers from a practical perspective, we define the minimum rate requirement of V_2 and U as $\tilde{C}_V$ and $\tilde{C}_U$, respectively. The optimization problem constructed for maximizing the sum capacity with a power allocation factor of α can be formulated as

$$\mathcal{P}:\max_{\alpha}\quad C_U + C_V^B \\ s.t. \begin{cases} C_U \geq \tilde{C}_U, & (a) \\ C_V^B \geq \tilde{C}_V, & (b) \\ 0 < \alpha < \frac{1}{2}. & (c) \end{cases} \tag{20}$$

B. The Proposed Power Allocation Algorithm

To solve the proposed optimization problem, we first express the derivatives of C_U, $C_V^{B_1}$, and $C_V^{B_2}$ with respect to α as

$$\begin{aligned} \frac{\mathrm{d}C_U}{\mathrm{d}\alpha} &= \frac{1}{N_t}\sum_{i=1}^{N_t}\frac{E_s\|\mathbf{g}_i\|^2}{\alpha E_s\|\mathbf{g}_i\|^2+\sigma_0^2}, \\ \frac{\mathrm{d}C_V^{B_1}}{\mathrm{d}\alpha} &= -\frac{1}{N_t}\sum_{n_t=1}^{N_t}\frac{E_s p_0\|\mathbf{h}_i\|^2}{\alpha E_s p_0\|\mathbf{h}_i\|^2+\sigma_0^2}, \\ \frac{\mathrm{d}C_V^{B_2}}{\mathrm{d}\alpha} &= -\frac{1}{N_t}\sum_{n_t=1}^{N_t}\frac{E_s p_0\|\mathbf{h}_i\|^2}{\alpha E_s p_0\|\mathbf{h}_i\|^2+\sigma_0^2}, \end{aligned} \tag{21}$$

respectively. Observe from (21) that, C_U is a monotonically increasing function of α, given its positive derivative, while both $C_V^{B_1}$ and $C_V^{B_2}$ are decreasing ones. Thus, when the constraint (c) of (20) is taken into account, there exist both minimum and maximum capacities that V_2 and U can achieve. Furthermore, to satisfy the constraint (a) and (b), we have the following conditions for $\tilde{C}_U$ and $\tilde{C}_V$, respectively

$$0 < \tilde{C}_U < C_U\left(\alpha = \frac{1}{2}\right),$$

$$C_V^B\left(\alpha = \frac{1}{2}\right) < \tilde{C}_V < C_V^B\left(\alpha = 0\right).$$

Given the above conditions, we can rewrite the constraints of problem $\mathcal{P}$ in a compact form as

$$g^{-1}\left(\tilde{C}_U\right) < \alpha < f^{-1}\left(\tilde{C}_V\right),$$

where $g^{-1}(\cdot)$ and $f^{-1}(\cdot)$ indicate the inverse function of C_U and C_V^B, respectively. To guarantee that the feasible set of problem $\mathcal{P}$ is non-empty, a further refined condition for setting $\tilde{C}_V$ is given by

$$C_V^B\left(\alpha = \frac{1}{2}\right) < \tilde{C}_V < C_V^B\left[\alpha = g^{-1}\left(\tilde{C}_U\right)\right].$$

Moreover, since $\|\mathbf{g}_{n_t}\|^2 > p_0\|\mathbf{h}_{n_t}\|^2$ is always satisfied, the derivative of $\left(C_U + C_V^B\right)$ can be guaranteed to have a positive value. Accordingly, the objective function of problem $\mathcal{P}$ is a monotonically increasing function and can be maximized, when α reaches the upper bound of its feasible set. With $\tilde{C}_U$ and $\tilde{C}_V$ being appropriately set, we find that the upper bound of α's feasible set is related to the constraint (b) of (20), and the lower bound corresponds to the constraint (a) of (20). Thus, the optimal solution of problem $\mathcal{P}$ is

$$\alpha_{opt}^{\mathcal{P}} = f^{-1}\left(\tilde{C}_V\right). \tag{22}$$

This optimal solution implies that the amount of power allocated to V_2 is 'just' sufficient to meet the minimum rate requirement $\tilde{C}_V$, while the remaining power is used for U, aiming for maximizing its capacity. Nevertheless, we should notice that there may exist some practical considerations, which require us to give high priority to the V2V transmission link, such as those of safety applications, which have to be served reliably. By contrast, the transmissions for in-car users are typically related to infotainment applications, for example, peer-to-peer video sharing and multimedia advertisements [39]. Hence it may be desirable to maximize the data rate of the V2V link, while guaranteeing the minimum rate requirement of the in-car user. To this end, we develop an alternative optimization problem formulated as

$$\mathcal{O}:\max_{\alpha}\quad C_V^B \\ s.t. \begin{cases} C_U \geq \tilde{C}_U, & (a) \\ C_V^B \geq \tilde{C}_V, & (b) \\ 0 < \alpha < \frac{1}{2}. & (c) \end{cases} \tag{23}$$

Clearly, the objective function of (23) is a monotonically decreasing function of α and it is maximized, when the constraint (a) is inactive. Therefore, the optimal solution of problem $\mathcal{O}$ can be written as

$$\alpha_{opt}^{\mathcal{O}} = g^{-1}\left(\tilde{C}_U\right). \tag{24}$$

So far, we have proposed a pair of power allocation schemes and analysed the solvability of the optimization problems considered. Explicitly, we provided an algorithm for finding the optimal solution of each problem, which are summarized in Table I. The proposed algorithm essentially performs bounding through with the aid of a bisection procedure, yielding globally optimal solutions at linearly increasing computational complexity [40]. In specific, the minimum rate requirements of V_2 and U are respectively set as

$$\begin{aligned} \tilde{C}_U &= \frac{C_U\left(\alpha=\frac{1}{2}\right)}{2}, \\ \tilde{C}_V &= \frac{C_V^B\left(\alpha=\frac{1}{2}\right)+C_V^B\left[\alpha=g^{-1}\left(\tilde{C}_U\right)\right]}{2}, \end{aligned} \tag{25}$$

for simplicity. Basically, both of the two power allocation optimization problems satisfy realistic practical considerations and the suitable one can be flexibly selected based on the specific data priority of the distinct transmission links.

V. Simulations and Discussions

In this section, simulation results are provided for evaluating the performance of the proposed NOMA-SM scheme. The system parameters are summarized as follows. The MIMO configurations for the NOMA-SM system are set as $N_t = 64$, $N_r = N_u = 2$. We fix $p_0 = 10^{-3}$, or, equivalently, the path loss exponential is set 3 and the distance between V_1 and V_2 is assumed to be 10 meters, which is typical for urban environments, especially during rush hours.

A. BER Results

In this subsection, the BER performance of the NOMA-SM scheme is compared to NOMA relying on the popular VBLAST technique, where NOMA-VBLAST is used as a reference. Specifically, we focus on the receiver performance of V_2. The effects of the Rician K-factor, adjacent antenna correlation coefficient, temporal correlation, and power allocation factor are all taken into consideration. The Rician K-factors are configured as $K = 2.186$ and $K = 0.2$ for

TABLE I
POWER ALLOCATION ALGORITHM

Power Allocation Algorithm for Problem $\mathcal{P}$ and Problem $\mathcal{O}$

1. Initialization

Set tolerance $0 < \varepsilon \ll 1$. Calculate $C_U\left(\alpha = \frac{1}{2}\right)$ and set $\tilde{C}_U = C_U\left(\alpha = \frac{1}{2}\right)/2$.

2. Determine the lower bound of α and find the optimal solution of problem $\mathcal{O}$

Set $\alpha_L = 0$ and $\alpha_U = \frac{1}{2}$.
While $\alpha_L - \alpha_U > \varepsilon$
  Set $\alpha = \frac{\alpha_L + \alpha_U}{2}$. Calculate $C_U(\alpha)$.
  If $C_U(\alpha) - \tilde{C}_U > 0$
    $\alpha_U = \alpha$
  Else
    $\alpha_L = \alpha$.
  End
End
Set $\tilde{C}_V = \left[C_V^B\left(\alpha = \frac{1}{2}\right) + C_V^B\left(\alpha = \frac{\alpha_L + \alpha_U}{2}\right)\right]/2$.
The optimal solution to the problem $\mathcal{O}$ is obtained as $\alpha_{opt}^{\mathcal{O}} = \frac{\alpha_L + \alpha_U}{2}$. Calculate $C_U\left(\alpha_{opt}^{\mathcal{O}}\right)$ and $C_V^B\left(\alpha_{opt}^{\mathcal{O}}\right)$.

3. Determine the upper bound of α and find the optimal solution of problem $\mathcal{P}$

Set $\alpha_{\min} = \frac{\alpha_L + \alpha_U}{2}$ and $\alpha_{\max} = \frac{1}{2}$.
While $\alpha_{\max} - \alpha_{\min} > \varepsilon$
  Set $\alpha = \frac{\alpha_{\min} + \alpha_{\max}}{2}$. Calculate $C_V^B(\alpha)$.
  If $C_V^B(\alpha) - \tilde{C}_V > 0$
    $\alpha_{\min} = \alpha$
  Else
    $\alpha_{\max} = \alpha$.
  End
End
The optimal solution of the problem $\mathcal{P}$ is obtained as $\alpha_{opt}^{\mathcal{P}} = \frac{\alpha_{\min} + \alpha_{\max}}{2}$. Calculate $C_U\left(\alpha_{opt}^{\mathcal{P}}\right)$ and $C_V^B\left(\alpha_{opt}^{\mathcal{P}}\right)$.

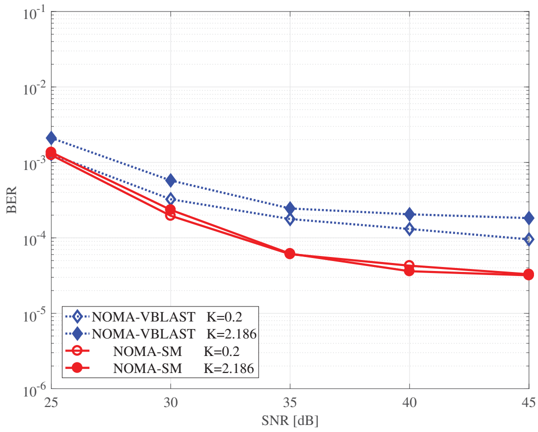

Fig. 6. BER comparisons with different Rician K-factor when $\kappa_t = \kappa_r = 0.2$ and $\delta = 1$ are given, and the power allocation factor is fixed at $\alpha = 0.001$, as evaluated by the Monte Carlo simulation with 10^6 channel realizations.

low and high vehicular traffic density, respectively (see [30] for more details). More specifically, QPSK and 16QAM are applied for NOMA-SM and NOMA-VBLAST, respectively. The MIMO configuration of the reference is the same as that of NOMA-SM except for using $N_t = 2$. Thus, the following BER comparisons are carried out for the same bandwidth efficiency of 8 bits per channel use (bpcu). The optimum ML detector described in (4) is employed at V_2 in both schemes. All simulation results of this subsection are obtained through a Monte Carlo method.

In Fig. 6, we show the BER performance for different Rician K-factor. It is observed that NOMA-SM significantly outperforms the benchmark. Additionally, the increase of K imposes

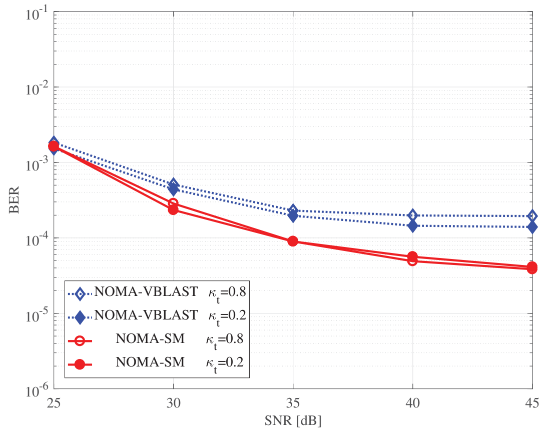

Fig. 7. BER comparisons with different adjacent antenna correlation coefficient at V_1, i.e., κ_t, when $K = 0.2$, $\kappa_r = 0.5$, and $\delta = 1$ are given, and the power allocation factor is fixed at $\alpha = 0.001$, as evaluated by the Monte Carlo simulation with 10^6 channel realizations.

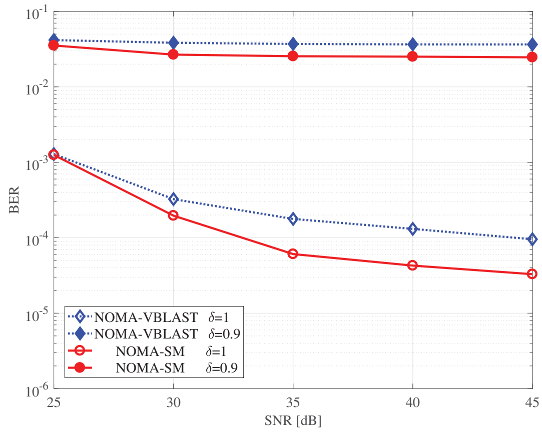

Fig. 8. BER comparisons with different temporal correlation coefficient δ when $K = 0.2$ and $\kappa_t = \kappa_r = 0.5$ are given, and the power allocation factor is fixed at $\alpha = 0.001$, as evaluated by the Monte Carlo simulation with 10^6 channel realizations.

a more dominant degradation on NOMA-VBLAST, which relies more vitally on the presence of rich non-LoS scattering. This phenomenon can be explained as follows. The higher Rician factor K represents a stronger LoS component, which increases the spatial correlation among the adjacent channel paths. For NOMA-VBLAST, the multiple-stream information is conveyed with the aid of multiple DoFs. By contrast, for NOMA-SM, although the more severe spatial correlation of the LoS scenario makes it difficult to determine the index of the activated TA, the remaining information related to the APM signal-domain is transmitted over a single DoF, hence it is less susceptible to spatial correlation.

Figure 7 investigates the BER results associated with different adjacent TA-correlation coefficients at V_1. Compared to $\kappa_t = 0.8$, $\kappa_t = 0.2$ represents an insignificant spatial correlation. Again, observe from Fig. 7 that NOMA-SM is less susceptible to spatial correlation. This phenomenon can be interpreted similarly to the trend of Fig. 6.

Below we investigate the impact of the V2V channel's time-varying nature. Observe from Fig. 8 that compared to the

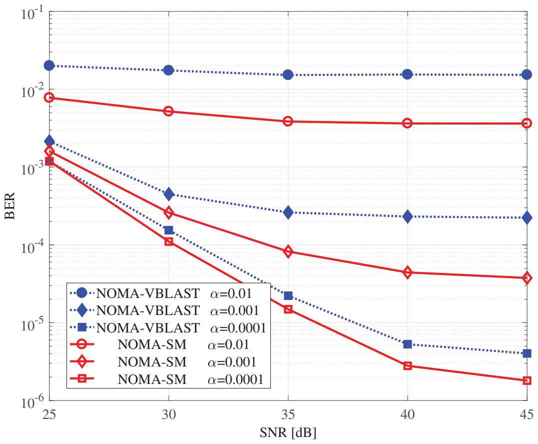

Fig. 9. BER comparisons with different power allocation factor α when $K = 0.2$, $\kappa_t = \kappa_r = 0.5$, and $\delta = 1$ are given, as evaluated by the Monte Carlo simulation with 10^6 channel realizations.

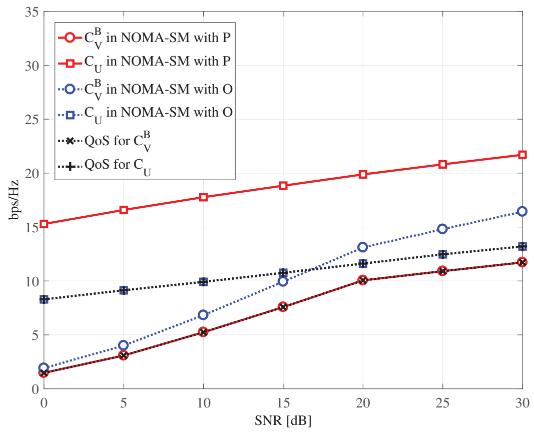

Fig. 11. Capacity of V_2 and U, or the respective QoS versus SNR for NOMA-SM with power allocation optimization $\mathcal{P}$ or $\mathcal{O}$. Specifically, C_V^B and C_U in NOMA-SM with $\mathcal{P}$ or $\mathcal{O}$ are evaluated with the aid of the algorithm in Table I. The QoS for C_V^B and C_U, i.e., $\tilde{C}_V$ and $\tilde{C}_U$ are set according to (25).

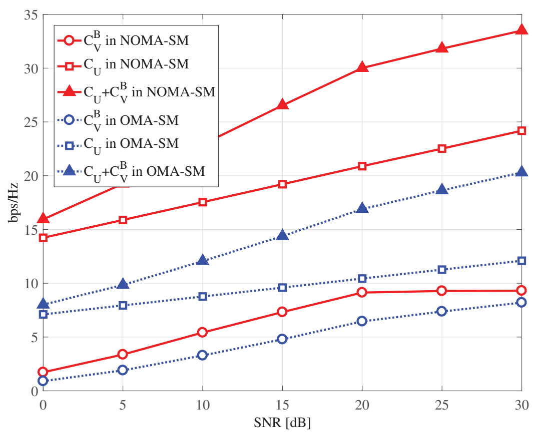

Fig. 10. Capacity of V_2 and U, or the sum capacity versus SNR for the NOMA-SM and OMA-SM scheme with a fixed power allocation factor, i.e., $\alpha = 0.1$. Specifically, C_V^B and C_U in NOMA-SM are evaluated from (19) and (12), while C_V^B and C_U in OMA-SM are obtained from (26).

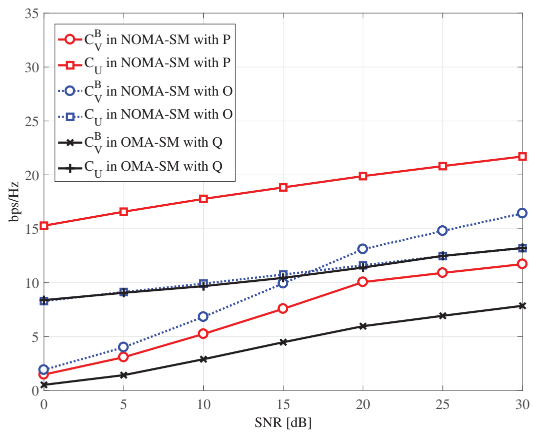

Fig. 12. Capacity of V_2 and U versus SNR for NOMA-SM with power allocation optimization $\mathcal{P}$ or $\mathcal{O}$, and OMA-SM with power allocation optimization $\mathcal{Q}$, respectively. Specifically, C_V^B and C_U in NOMA-SM with $\mathcal{P}$ or $\mathcal{O}$ are evaluated with the aid of the algorithm in Table I. While C_V^B and C_U in OMA-SM with $\mathcal{Q}$ are obtained from a full-search algorithm.

performance of no time-varying effect associated with $\delta = 1$, the BER has been substantially degraded in both schemes for $\delta = 0.9$. Although a perfect channel estimation procedure is assumed for the receivers, the estimated channel coefficients used for ML detection becomes partially outdated due to the channel's time-varying nature, hence resulting in a degraded BER performance. Nevertheless, the proposed NOMA-SM scheme maintains its advantage over the reference, regardless of the grade of temporal correlation.

Figure 9 shows the BER performance associated with different α values. For both schemes, the lower α values exhibit a better detection performance, since less power is allocated to U and hence V_2 suffers from a lower inter-user interference. More importantly, we observe that NOMA-SM consistently outperforms NOMA-VBLAST. By jointly considering the above observations, we conclude that NOMA-SM constitutes a potent amalgam.

B. Capacity Results and Discussions

Below we evaluate the capacity of the NOMA-SM system associated with different power allocation strategies. All results presented in this subsection are obtained by averaging the instantaneous capacities over multiple channel realizations. For convenience, 16PSK is applied to both V_2 and U. In particular, we fix $K = 0.2$, $\kappa_t = \kappa_r = 0.5$, and $\delta = 1$ unless otherwise stated. For benchmarking, we use an OMA-SM system, where V_1 transmits messages to V_2 using SM in the first slot. Then V_1 sends messages through the previous activated antenna to U, without activating another antenna. This OMA-SM model constitutes a fair reference for the NOMA-SM system, since the signal intended for V_2 is conveyed by both the APM signal- and TA-domain, whereas the signal destined for U is only embedded in the classical signal-domain. The distinctive feature of OMA-SM is that data transmissions destined for V_1-V_2 and V_1-U are operated in an orthogonal time-division way within the classical APM signal-domain. Accordingly, the

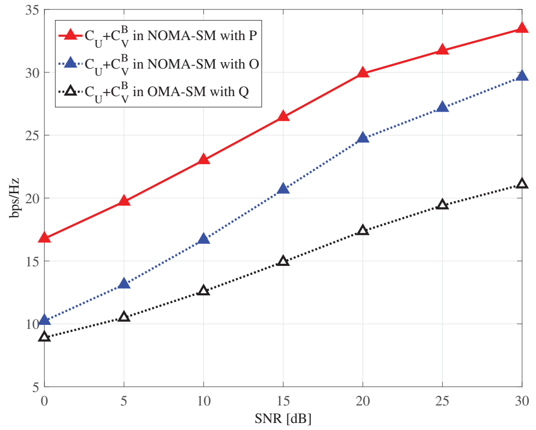

Fig. 13. Sum capacity versus SNR for NOMA-SM with power allocation optimization $\mathcal{P}$ or $\mathcal{O}$, and OMA-SM with power allocation optimization $\mathcal{Q}$, respectively. Specifically, C_V^B and C_U in NOMA-SM with $\mathcal{P}$ or $\mathcal{O}$ are evaluated with the aid of the algorithm in Table I. While C_V^B and C_U in OMA-SM with $\mathcal{Q}$ are obtained from a full-search algorithm.

capacity upper bound for V_2 and the capacity for U in the OMA-SM system are expressed as

$$\begin{aligned} \mathcal{C}_V^B &= \min\left\{\mathcal{C}_V^{B_1}, \mathcal{C}_V^{B_2}\right\}, \\ \mathcal{C}_U &= \frac{1}{2N_t}\sum_{i=1}^{N_t}\log_2\left(1+\frac{\alpha E_s}{\sigma_0^2}\|\mathbf{g}_i\|^2\right), \end{aligned} \tag{26}$$

respectively, where

$$\begin{aligned} \mathcal{C}_V^{B_1} &= \frac{1}{2N_t}\sum_{i=1}^{N_t}\log_2\left(1+\frac{(1-\alpha)E_s p_0}{\sigma_0^2}\|\mathbf{h}_i\|^2\right)+\frac{1}{2}\log_2(N_t), \\ \mathcal{C}_V^{B_2} &= \frac{1}{2}\log_2\det\left[\mathbf{I}+\frac{(1-\alpha)E_s p_0}{\sigma_0^2 N_t}\mathbf{H}\mathbf{H}^H\right]. \end{aligned} \tag{27}$$

Let us first check the capacity associated with a fixed power allocation, that is $\alpha = 0.1$. Figure 10 depicts the capacity of V_2 and U, as well as the sum capacity versus SNR for both NOMA-SM and OMA-SM. Compared to OMA-SM, NOMA-SM provides substantial capacity gains both for the collaboration-aided vehicle V_2 and for the in-car user, and accordingly obtains a significant sum capacity enhancement. Specifically, the capacity $\mathcal{C}_U$ of the in-car user has been beneficially boosted by the proposed scheme, about twice as high as that of OMA-SM. Since the APM signal-domain of the proposed scheme is combined with a NOMA strategy, each user accesses the channel resources via power domain multiplexing.

Subsequently, we investigate the efficiency of the proposed power allocation optimization. Specifically, the power allocation optimization denoted by $\mathcal{Q}$ is considered for OMA-SM, which is formulated as

$$\begin{aligned} \mathcal{Q}: \max_{\alpha} \quad & \mathcal{C}_U + \mathcal{C}_V^B \\ s.t. & \begin{cases} \mathcal{C}_U \geq \tilde{\mathcal{C}}_U, \\ \mathcal{C}_V^B \geq \tilde{\mathcal{C}}_V^B, \\ 0 < \alpha < 1. \end{cases} \end{aligned} \tag{28}$$

For simplicity, the minimum rate requirements of V_2 and U are set to $\tilde{\mathcal{C}}_U = \frac{\mathcal{C}_U(\alpha=1)}{2}$ and $\tilde{\mathcal{C}}_V^B = \frac{\mathcal{C}_V^B(\alpha=0)}{2}$, which respectively correspond to the lower bound and upper bound of α's feasible set. Then a full-search algorithm is applied for OMA-SM within the feasible set.

Figure 11 illustrates the capacity of V_2 and U for NOMA-SM with optimization $\mathcal{P}$ or $\mathcal{O}$, where the QoS of the collaboration-aided vehicle V_2 and the in-car user U, i.e., $\tilde{C}_V^B$ and $\tilde{C}_U$, are also plotted for reference. It can be observed that C_V^B always meets the requirement of $\tilde{C}_V^B$ with the aid of the optimization $\mathcal{P}$, and C_U associated with the optimization $\mathcal{O}$ exactly meets the QoS $\tilde{C}_U$. This observation is in accordance with the foregoing analysis, which indicates that the optimization $\mathcal{P}$ intends to maximize C_U, while maintaining the QoS $\tilde{C}_V^B$ for V2V transmission, whereas the optimization $\mathcal{O}$ aims for maximizing C_V^B while guaranteeing the minimum rate requirement $\tilde{C}_U$ for the in-car user. Thus, we find that the optimized C_U of $\mathcal{P}$ is higher than that of $\mathcal{O}$, whereas the optimized C_V^B of $\mathcal{O}$ outperforms that of $\mathcal{P}$. Accordingly, the more appropriate optimization scheme can be readily selected based on the data priority of distinct transmission links.

Figure 12 compares the results of the optimization $\mathcal{Q}$ to that of $\mathcal{P}$ and $\mathcal{O}$. Let us contrast $\mathcal{P}$ and $\mathcal{Q}$ first. Clearly, both C_V^B and C_U in $\mathcal{P}$ have been remarkably improved, demonstrating that the NOMA strategy offers a bandwidth efficiency improvement. By considering the results of $\mathcal{O}$ and $\mathcal{Q}$ in Fig. 12, we find that C_U of NOMA-SM associated with optimization $\mathcal{O}$ is tightly lower bounded by that of OMA-SM associated with optimization $\mathcal{Q}$, and C_V^B with $\mathcal{O}$ provides a substantial gain, achieving more than twice that of $\mathcal{Q}$. Clearly, the NOMA-SM system associated with optimization $\mathcal{O}$ is capable of offering better user fairness than that of optimization $\mathcal{P}$.

Furthermore, it can be observed from Fig. 13 that the NOMA-SM systems achieve higher sum capacity than OMA-SM. Specifically, optimization $\mathcal{P}$ provides higher capacity gain than $\mathcal{O}$, since $\mathcal{P}$ aims for maximizing the data rate of the in-car user U, which experiences a much better channel than the collaboration-aided vehicle V_2.

VI. Conclusions

The new NOMA-SM transmission strategy has been proposed in this treatise. Its BER performance has been investigated with the impact of the Rician K-factor, spatial correlation of antenna array, time-varying effect of the V2V channel, and the power allocation factor being discussed. Compared to NOMA relying on VBLAST, NOMA-SM has been demonstrated to exhibit improved robustness against the spatial and temporal effects of the V2V channel. By analysing the capacity and deriving analytical upper bounds in closed form, a pair of power allocation optimization schemes have been formulated for NOMA-SM. The optimal solutions have also been shown to be achievable with the aid of the proposed power allocation algorithm. Our numerical results have verified that with the aid of an appropriate power allocation, NOMA-SM is capable of satisfying the QoS support of a low priority flow, whilst maximizing the throughput of the high priority flow. In summary, NOMA-SM has been demonstrated to cooperatively improve the link reliability and bandwidth efficiency of V2V transmissions.

REFERENCES

[1] L. Hanzo, O. Alamri, M. El-Hajjar, and N. Wu, *Near-capacity multi-functional MIMO systems: Sphere-packing, iterative detection and cooperation*. New York, NY, USA: Wiley, May 2009.

[2] E. G. Larsson, O. Edfors, F. Tufvesson, and T. L. Marzetta, "Massive MIMO for next generation wireless systems," *IEEE Communications Magazine*, vol. 52, no. 2, pp. 186–195, Feb. 2014.

[3] Y. Wu, R. Schober, D. W. K. Ng, C. Xiao, and G. Caire, "Secure massive MIMO transmission with an active eavesdropper," *IEEE Transactions on Information Theory*, vol. 62, no. 7, pp. 3880–3900, Jul. 2016.

[4] R. Zhang, Z. Zhong, J. Zhao, B. Li, and K. Wang, "Channel measurement and packet-level modeling for V2I spatial multiplexing uplinks using massive MIMO," *IEEE Transactions on Vehicular Technology*, vol. 65, no. 10, pp. 7831–7843, Oct. 2016.

[5] P. Harris *et al.*, "Performance characterization of a real-time massive MIMO system with LOS mobile channels," *IEEE Journal on Selected Areas in Communications*, April 2017, to appear.

[6] Y. Yang and B. Jiao, "Information-guided channel-hopping for high data rate wireless communication," *IEEE Communications Letters*, vol. 12, no. 4, pp. 225–227, Apr. 2008.

[7] L. He, J. Wang, J. Song, and L. Hanzo, "On the multi-user, multi-cell massive spatial modulation uplink: How many antennas for each user?" *IEEE Transactions on Wireless Communications*, Dec. 2016, to appear.

[8] D. A. Basnayaka, M. Di Renzo, and H. Haas, "Massive but few active MIMO," *IEEE Transactions on Vehicular Technology*, vol. 65, no. 9, pp. 6861–6877, Sept. 2016.

[9] M. Di Renzo, H. Haas, A. Ghrayeb, S. Sugiura, and L. Hanzo, "Spatial modulation for generalized MIMO: Challenges, opportunities, and implementation," *Proceedings of the IEEE*, vol. 102, no. 1, pp. 56–103, Jan. 2014

[10] M. Zhang, X. Cheng, and L. Q. Yang, "Differential spatial modulation in V2X," in *Proc. 2015 9th European Conference on Antennas and Propagation (EuCAP)*, Lisbon, Apr. 2015, pp. 1–5.

[11] Y. Fu *et al.*, "BER performance of spatial modulation systems under 3-D V2V MIMO channel models," *IEEE Transactions on Vehicular Technology*, vol. 65, no. 7, pp. 5725–5730, Jul. 2016.

[12] K. P. Peppas, P. S. Bithas, G. P. Efthymoglou, and A. G. Kanatas, "Space shift keying transmission for intervehicular communications," *IEEE Transactions on Intelligent Transportation Systems*, May 2016, to appear.

[13] Y. Cui and X. Fang, "Performance analysis of massive spatial modulation MIMO in high-speed railway," *IEEE Transactions on Vehicular Technology*, vol. 65, no. 11, pp. 8925–8932, Nov. 2016.

[14] Y. Saito, Y. Kishiyama, A. Benjebbour, T. Nakamura, A. Li, and K. Higuchi, "Non-orthogonal multiple access (NOMA) for cellular future radio access," in *Proc. IEEE 77th Vehicular Technology Conference*, Dresden, Germany, Jun. 2013, pp. 1–5.

[15] L. Dai, B. Wang, Y. Yuan, S. Han, C. I, and Z. Wang, "Non-orthogonal multiple access for 5G: Solutions, challenges, opportunities, and future research trends," *IEEE Communications Magazine*, vol. 53, no. 9, pp. 74–81, Sept. 2015.

[16] J. Liberti, S. Moshavi, and P. Zablocky, "Successive interference cancellation", U.S. Patent 8670418 B2, Mar. 11th, 2014.

[17] Q. Sun, S. Han, C. L. I, and Z. Pan, "On the ergodic capacity of MIMO NOMA systems," *IEEE Wireless Communications Letters*, vol. 4, no. 4, pp. 405–408, Aug. 2015.

[18] Y. Saito, A. Benjebbour, Y. Kishiyama, and T. Nakamura, "System-level performance evaluation of downlink Non-orthogonal Multiple Access (NOMA)," in *Proc. IEEE 24th Int. Symp. Personal Indoor and Mobile Radio Communications (PIMRC)*, London, UK, Sept. 2013, pp. 611–615.

[19] Z. Yang, Z. Ding, P. Fan, and N. Al-Dhahir, "A general power allocation scheme to guarantee quality of service in downlink and uplink NOMA systems," *IEEE Transactions on Wireless Communications*, vol. 15, no. 11, pp. 7244–7257, Nov. 2016.

[20] L. Lv, J. Chen, and Q. Ni, "Cooperative non-orthogonal multiple access in cognitive radio," *IEEE Communications Letters*, vol. 20, no. 10, pp. 2059–2062, Oct. 2016.

[21] D. Jiang and L. Delgrossi, "IEEE 802.11p: Towards an international standard for wireless access in vehicular environments," in *Proc. VTC Spring 2008 - IEEE Vehicular Technology Conference*, Singapore, May 2008, pp. 2036–2040.

[22] S. Chen, J. Hu, Y. Shi, and L. Zhao, "LTE-V: A TD-LTE-based V2X solution for future vehicular network," *IEEE Internet of Things Journal*, vol. 3, no. 6, pp. 997–1005, Dec. 2016.

[23] T. Wang, L. Song, and Z. Han, "Coalitional graph games for popular content distribution in cognitive radio VANETs," *IEEE Transactions on Vehicular Technology*, vol. 62, no. 8, pp. 4010–4019, Oct. 2013.

[24] H. Ilhan, I. Altunbas, and M. Uysal, "Optimized amplify-and-forward relaying for vehicular ad-hoc networks," in *Proc. IEEE Vehicular Technology Conference*, Calgary, BC, Sept. 2008, pp. 1–5.

[25] K. Huang, V. K. N. Lau, and Y. Chen, "Spectrum sharing between cellular and mobile ad hoc networks: transmission-capacity trade-off," *IEEE Journal on Selected Areas in Communications*, vol. 27, no. 7, pp. 1256–1267, Sept. 2009.

[26] T. Lakshmi Narasimhan, P. Raviteja, and A. Chockalingam, "Generalized spatial modulation in large-scale multiuser MIMO systems," *IEEE Transactions on Wireless Communications*, vol. 14, no. 7, pp. 3764–3779, Jul. 2015.

[27] Y. Chen, L. Wang, Z. Zhao, M. Ma, and B. Jiao, "Secure multiuser MIMO downlink transmission via precoding-aided spatial modulation," *IEEE Communications Letters*, vol. 20, no. 6, pp. 1116–1119, Jun. 2016.

[28] L. Wang, S. Bashar, Y. Wei, and R. Li, "Secrecy enhancement analysis against unknown eavesdropping in spatial modulation," *IEEE Communications Letters*, vol. 19, no. 8, pp. 1351–1354, Aug. 2015.

[29] O. Delangre, S. Van Roy, P. De Doncker, M. Lienard, and P. Degauque, "Modeling in-vehicle wideband wireless channels using reverberation chamber theory," in *Proc. 2007 IEEE 66th Vehicular Technology Conference*, Baltimore, MD, Oct. 2007, pp. 2149–2153.

[30] X. Cheng, C. X. Wang, D. I. Laurenson, S. Salous, and A. V. Vasilakos, "An adaptive geometry-based stochastic model for non-isotropic MIMO mobile-to-mobile channels," *IEEE Transactions on Wireless Communications*, vol. 8, no. 9, pp. 4824–4835, Sept. 2009.

[31] M. Koca and H. Sari, "Performance analysis of spatial modulation over correlated fading channels," in *Proc.2012 IEEE Vehicular Technology Conference (VTC Fall)*, Quebec City, QC, Sept. 2012, pp. 1–5.

[32] J. P. Kermoal, L. Schumacher, K. I. Pedersen, P. E. Mogensen, and F. Frederiksen, "A stochastic MIMO radio channel model with experimental validation," *IEEE Journal on Selected Areas in Communications*, vol. 20, no. 6, pp. 1211–1226, Aug. 2002.

[33] S. L. Loyka, "Channel capacity of MIMO architecture using the exponential correlation matrix," *IEEE Communications Letters*, vol. 5, no. 9, pp. 369–371, Sept. 2001.

[34] C. Liu, M. Ma, Y. Yang, and B. Jiao, "Optimal spatial-domain design for spatial modulation capacity maximization," *IEEE Communications Letters*, vol. 20, no. 6, pp. 1092–1095, Jun. 2016.

[35] X. Guan, Y. Cai, and W. Yang, "On the mutual information and precoding for spatial modulation with finite alphabet," *IEEE Wireless Communications Letters*, vol. 2, no. 4, pp. 383–386, Aug. 2013.

[36] Z. An, J. Wang, J. Wang, S. Huang, and J. Song, "Mutual information analysis on spatial modulation multiple antenna system," *IEEE Transactions on Communications*, vol. 63, no. 3, pp. 826–843, Mar. 2015.

[37] R J. Muirhead, *Aspects of Multivariate Statistical Theory*. Wiley, 1982: 191-192.

[38] S. Wu *et al.*, "A novel method for ergodic sum rate analysis of spatial modulation systems with maximum likelihood receiver," *2015 International Wireless Communications and Mobile Computing Conference (IWCMC)*, Dubrovnik, Aug. 2015, pp. 32–36.

[39] Y. Toor, P. Muhlethaler, A. Laouiti, and A. D. La Fortelle, "Vehicle Ad Hoc networks: applications and related technical issues," *IEEE Communications Surveys and Tutorials*, vol. 10, no. 3, pp. 74–88, Third Quarter 2008.

[40] S. Boyd and L. Vandenberghe, *Convex Optimization*. Cambridge, U.K.: Cambridge Univ. Press, 2004.

Yingyang Chen received the B. Eng. degree in electronic engineering from University of Science and Technology of China (UESTC), Chengdu, China, in 2014. She is currently pursuing the Ph. D. degree with the Modern Communications Research Institute at Peking University, Beijing, China. Her research interests include full-duplex communications, physical layer security, and signal processing in wireless communications.

Li Wang (S'08–M'14–SM'16) received her Ph.D. degree in 2009 from Beijing University of Post and Telecommunications (BUPT), Beijing, China. She is now a professor in the School of Electronic Engineering, BUPT, where she leads the Lab of High Performance Computing and Networking, and serves as the special assistant to the dean on international relations. She held a visiting professor position in the School of Electrical and Computer Engineering at Georgia Tech, Atlanta, USA from December 2013 to January 2015 and in the Department of Signals and Systems at Chalmers University of Technology, Gothenburg, Sweden, from August 2015 to November 2015.

Her research interests include wireless communications, cognitive radio, secure communications, distributed storage systems, device-to-device communication systems, and social networking. Due to her outstanding research achievement in these areas, she received the 2013 Beijing Young Elite Faculty for Higher Education Award, the best paper award at ICCTA 2011, and best paper runner up from WASA 2015, respectively. She has been serving as a SWAT Team Member since January 2016, an Editor for IEEE Transactions on Vehicular Technology since January 2017, and an Associate Editor for IEEE Access since June 2016. She also served the Technical Program Committee of several IEEE conferences, including IEEE GLOBECOM, ICC, WCNC, and VTC.

Yutong Ai received the B.S. degree from Beijing University of Posts and Telecommunications (BUPT), Beijing, China, in 2016, where he is currently pursuing the M.S. degree in electronic science and technology. His research interests include wireless networking, secure communication, and computation offloading in wireless communications.

Bingli Jiao (M'05) received the B.S. and M.S. degrees from Peking University, China, in 1983 and 1988, respectively, and the Ph.D. degree from Saarland University, Germany, in 1995. Then, he became an associate professor and professor with Peking University in 1995 and 2000, respectively. His current interests include communication theory and techniques and sensor design.

Lajos Hanzo (FREng, FIEEE, FIET, Fellow of EURASIP, DSc) received his degree in electronics in 1976 and his doctorate in 1983. In 2009 he was awarded an honorary doctorate by the Technical University of Budapest and in 2015 by the University of Edinburgh. In 2016 he was admitted to the Hungarian Academy of Science. During his 40-year career in telecommunications he has held various research and academic posts in Hungary, Germany and the UK. Since 1986 he has been with the School of Electronics and Computer Science, University of Southampton, UK, where he holds the chair in telecommunications. He has successfully supervised 111 PhD students, co-authored 18 John Wiley/IEEE Press books on mobile radio communications totalling in excess of 10 000 pages, published 1600+ research contributions at IEEE Xplore, acted both as TPC and General Chair of IEEE conferences, presented keynote lectures and has been awarded a number of distinctions. Currently he is directing a 60-strong academic research team, working on a range of research projects in the field of wireless multimedia communications sponsored by industry, the Engineering and Physical Sciences Research Council (EPSRC) UK, the European Research Council's Advanced Fellow Grant and the Royal Society's Wolfson Research Merit Award. He is an enthusiastic supporter of industrial and academic liaison and he offers a range of industrial courses. He is also a Governor of the IEEE VTS. During 2008 - 2012 he was the Editor-in-Chief of the IEEE Press and a Chaired Professor also at Tsinghua University, Beijing. For further information on research in progress and associated publications please refer to http://www-mobile.ecs.soton.ac.uk. Dr. Hanzo has 30 000+ citations and an H-index of 68.