

Space-Time Super-Modulation: Concept, Design Rules, and its Application to Joint Medium Access and Rateless Transmission

Konstantinos Nikitopoulos, *Member, IEEE*, Farhad Mehran, *Member, IEEE*,
and Hamid Jafarkhani, *Fellow, IEEE*

Abstract

We introduce the concept of Space-Time Super-Modulation according to which additional low-rate and highly reliable information can be transmitted on top of traditionally modulated and space-time encoded information, without increasing the transmitted block length or degrading their error-rate performance. This is achieved by exploiting the temporal redundancy introduced by the space-time block codes and, specifically, by efficiently mapping transmission patterns to specific information content. We show that Space-Time Super-Modulation can be efficiently used in the context of machine-type communications to enable “one-shot”, “grant-free” joint medium access and rateless data transmission while reducing or even eliminating the need for transmitting preamble sequences. As a result, compared with traditional approaches that use correlatable preamble sequences or encoded preambles to transmit the signature information of transmitted packets, Space-Time Super-Modulation can achieve significant throughput gains. For example, we show up to 35% throughput gains from the second best examined preamble-based scheme when transmitting blocks of 200 bits.

Index Terms

Multiple-input multiple-output (MIMO), Space-time block coding (STBC), Machine-type communications (MTCs), Multilevel codes.

Part of this work has been presented at IEEE GLOBECOM 2016.

Konstantinos Nikitopoulos and Farhad Mehran are with 5G Innovation Center, Institute for Communication Systems (ICS), University of Surrey, Guildford, GU2 7XH, UK (e-mail:{k.nikitopoulos,f.mehran}@surrey.ac.uk).

Hamid Jafarkhani is with Center for Pervasive Communications and Computing, University of California, Irvine, CA, 92697, USA (e-mail:hamidj@uci.edu). This work was supported in part by the NSF Award CCF-1526780.

I. INTRODUCTION

As telecommunication technologies and applications evolve, a continuously increasing number of devices require to be connected wirelessly. Such machine-type communications (MTC) have diverse requirements depending on the service, the application, and the type of devices that need to communicate [1]-[3]. These diverse requirements, together with the expected number of devices to be connected during the coming years, introduce new challenges and trigger a need to revisit the current medium access and data transmission strategies [4]-[8].

One of the main MTC challenges relates to the *sporadic* wireless traffic which is expected to dramatically increase in the near future [6][9][10]. In sporadic data transmission, a small amount of information is typically transmitted. Then, the signaling overhead required to connect (and synchronize) a machine, together with the signaling required for reliable transmission, can result in severe network underutilization. For example, for a Random Access Channel (RACH) as used in LTE/LTE-A, to transmit 100 bytes of data from a user to the Base Station (BS), the access procedure requires approximately 59 and 136 bytes of overhead in the uplink and downlink, respectively [2]. To avoid this overhead, as well as the delays induced from such an information exchange, recent research focuses on finding solutions able to simultaneously handle medium access and data transmission [11]. These methods are referred to as “one-shot” or “grant-free” transmission [12] or “joint medium access and data transmission techniques”.

Ideally, a future MTC protocol should enable one-shot, asynchronous, and highly-reliable transmission, with very low (or no) signaling overhead. However, reliability and low signaling overhead are, in principle, competing requirements. For example, for recovering the transmitted information of a specific user, it is necessary for the receiver to reliably identify its identification (ID) information, and, therefore, the ID information should be protected with very strong codes, or long preamble transmissions, that involves heavier ID signaling [13].

In addition to ID transmission, in order to efficiently transmit information close to the capabilities of the transmission channel (i.e., close to channel capacity [14]), efficient rate adaptation that takes place at the transmitter side is required [15]-[17]. Current rate adaptation schemes that are based on adaptive modulation and coding require instantaneous knowledge of the channel condition and add undesirable signaling overhead [6]. This overhead can become significantly higher if information is transmitted over different coherence times. Applying Rateless codes to the physical (PHY) layer is a very promising way to alleviate the need for this overhead [18][19].

However, for the decoding of rateless codes, the receiver needs to know not only the ID of the machine, but also the ordered position of the received packet among the entire rateless-coded packets, which would typically require additional signaling. In order to avoid long packet ID transmission, the idea of jointly coding the machine header and payload has been highlighted for future wireless networks [2]. However, to the best of the authors' knowledge, no practical solution has been proposed so far able to identify machines that transmit information in an asynchronous, *ad hoc* and sporadic manner.

This work introduces a *Space-Time Super-Modulation* (STSM) scheme that enables highly-reliable joint medium access and rateless transmission, without requiring the transmission of preambles for delivering the *signature information* (SI) of a transmitted packet. In particular, with STSM, an additional low-rate and highly reliable information stream (or subchannel) can be transmitted by further super modulating (SM) on top of space-time encoded [20][21][22] sequences. The STSM is performed by altering the pattern of the transmitted space-time encoded packet in a way that the Euclidean distance is increased between possible codewords of the highly reliable information stream. As a result, STSM can be used for joint medium access and data transmission where useful information is encoded by means of “traditional” (e.g., rateless) binary codes and SI is encoded by altering the pattern of transmitted space-time-encoded packet. To the best of the authors' knowledge, STSM is the first approach that allows the transmission of additional flexible rate and highly reliable information by “encoding” on top of a space-time encoded sequence and by exploiting its temporal redundancy. Then, as shown in Section V, STSM can provide more reliable SI identification compared to the traditional preamble-based techniques, even in the case of colliding users, but by obviating the need for machine header preambles. In addition, it is the first time that such an approach is used to enable “rateless” coding of payload in MTC communications and reliable machine header transmission concurrently with the useful data, resulting in throughput gains of up to 35% compared to conventional preamble-based approaches, when a SI of 9 bits, and a packet size of 200 bits are assumed. To enable “one-shot” or “grant free” access, prior techniques like [28] or [29], require synchronous user transmission and unique per-user access patterns with specific properties (i.e., sparsity) in order to be efficiently identifiable and decodable, while the technique in [29] further requires a temporal correlation of the active user sets. To the best of our knowledge, STSM is the first approach that can enable user identification for both synchronous and asynchronous user transmission, and *ad hoc* (temporal) patterns that are unknown to the receiver. In contrast to traditional,

coordinated/synchronous approaches that may require from the signal of a specific user to be received at a specific time instant in order to be identifiable, STSM-aided transmission, does not require any “time stamps” and therefore, it also obviates any need for user delay estimation from the access point. The adaptation of the proposed super-modulation is not limited to space-time coded systems only. SM can be extended to any scheme that imposes spatial or frequency redundancy e.g., when repetition coding is employed.

Since STSM tries to exploit the increase of the Euclidean distance in order to transmit additional, highly reliable information sequences of very low rate, it can be assumed to be a member of the greater family of *multilevel codes (MLCs)* [23][24]. Therefore, superficial similarities exist between STSM and other members of MLC family. Still, there are fundamental differences between them. In particular, MLCs, including *Trellis-Coded-Modulation (TCM)* [25], [26] and their Space-Time versions [30], aim the joint optimization of coding and modulation for minimizing error-rate and enhancing transmission quality. In particular, traditional MLC schemes partition information sequence into component sequences and encode each part by using an individual encoder. Transmission symbols are constructed by combining codewords created by each encoder. The individual codes are co-optimized for maximizing the minimum Euclidean distance of the codewords. Then, computationally intensive joint decoding schemes are required. STSM, on the other hand, targets the concurrent transmission of two information streams, with one stream being of much smaller rate that can be flexible, without accounting for the particular coding scheme that can be further applied to these streams. STSM “encodes” the additional information by exploiting the temporal redundancy introduced by the space-time codes and without increasing the transmission length. STSM does not necessitate any channel coding scheme on top of the sequences. Still, the two sequences can be further channel encoded by any known code at any rate. In such a case, the two streams can be (channel) decoded independently since the detection process (taking place before decoding), presented in Section II, can demultiplex the two jointly transmitted information streams into two independent ones. In Section V, where the application of STSM to joint medium access is examined, the SI is supposed to be uncoded, and the conventionally transmitted information is ratelessly encoded by means of Raptor codes.

In the same family of MLCs, *Trellis-Coded-Modulation (TCM)* [25], [26] and their Space-Time extensions (e.g., Super-Orthogonal Space-Time Trellis Codes [30]) also aim the joint optimization of coding and modulation. On the other hand, STSM allows the transmission of an additional

information stream of a flexible transmission rate. In addition, while traditional TCM schemes are based on convolutional codes, STSM can support any type of channel coding.

Spatial modulation [27] is an alternative, but fundamentally different approach to transmit additional information to the conventionally modulated one. In particular, while STSM transmits the additional information by exploiting the temporal redundancy of the space-time code, spatial modulation exploits the spatial dimension (i.e., it selects transmit antennas). When spatial modulation is applied to systems with small antenna numbers, in contrast to STSM, the transmit antenna identification (and therefore detection of the additional information) becomes less reliable and the diversity gains are compromised, resulting in significantly degraded error-rate performance compared with systems exploiting space-time-coding approaches.

The rest of the paper is organized as follows. In Section II, the concept of STSM is presented. Section III presents the design of efficient STSM codewords. Section IV discusses how STSM can be used for joint data medium access and data transmission, and in Section V the evaluation of the proposed approach follows.

II. SPACE-TIME SUPER-MODULATION (STSM)

Typically, the transmission pattern of the conventionally modulated symbols after space-time block coding (i.e., the phase, the amplitude and the relative position of the actual and redundant information in the space/time/phase grid) is unique, predetermined, and *a priori* known to both the transmitter and the receiver [20][21]. Instead of having such a unique pattern, and as we have first discussed in [31], STSM allows the employment of multiple but still predefined sets of Super Modulation Patterns (SMPs). Which pattern will be transmitted is finally dictated by the additional information to be transmitted after appropriate *bit-to-pattern* (similar to the traditional bit-to-symbol) mapping that targets the maximization of the corresponding minimum Euclidean distance between those patterns. Then, if the transmitted pattern can be reliably identified at the receiver side, the corresponding information content can be recovered and, therefore, a throughput increase can be achieved.

Various approaches can be used to super-modulate the conventionally modulated symbols as a function of the corresponding SMP as long as the corresponding pattern can be uniquely identified (i.e., demodulated) at the receiver side. For example, the SMP can modulate the phase and/or the amplitude of conventionally modulated symbols, the relative position of the actual and redundant information (in the case of space-time block codes), or even a combination of those

parameters. This paper focuses on the case of phase STSM due to its simplicity and because in contrast to amplitude modulation methods, it avoids increasing the peak to average power ratio which makes the detection efficiency very sensitive to the nonlinear devices of the processing loop (e.g., digital to analog converter, high power amplifier).

In the rest of this section, the encoding and decoding processes of STSM are presented. While the proposed approach is applicable to any type of space-time block code, the practical 2×2 Alamouti space-time block code (STBC) [20] is examined, especially since for MTCs a low number of antennas is expected. The discussion is focused on low order constant amplitude constellations (e.g., BPSK), since as previously discussed the Super Modulation (SM) scheme primarily targets “unfavorable” transmission scenarios.

A. STSM Encoding

The STSM scheme requires transmission in blocks. The size of the block is assumed to be equal to L channel uses, such that the corresponding transmission channel can be assumed static for the block duration. The proposed STSM scheme for the case of a 2×2 Alamouti scheme is depicted in Fig. 1. For each transmitted STSM block, the bits to be transmitted are split into two subsets: (a) *The Conventionally Modulated Bits (CMB)* and (b) *The Super-Modulated Bits (SMB)*. The CMB subset consists of the bits which would typically be transmitted without STSM. These bits are mapped onto conventional complex information symbols \mathbf{S} . The SMB subchannel is of lower rate, and therefore of higher reliability than the CMB, and consists of the additional bits to be transmitted via the proposed mapping technique. In MTCs, this subchannel is used to transmit each packet’s signature bits. The SMBs are mapped onto patterns (SMPs) via an appropriate SMB-to-SMP mapping. Then, the selected SMP c , which is characterized by its *characteristic SMP vector* $\mathbf{\Pi}_c$, determines the way that the conventionally produced symbols (from CMBs) will be further modulated via SM. After SM, the produced symbols are space-time encoded to produce \mathbf{S}_c which will be finally transmitted.

1) *CMB to Conventional Symbol Mapping*: Since for a 2×2 scheme with Alamouti space-time block code $B = L/2$ channel uses deliver actual information and B channel uses are related to the same information, $2B \log_2 |\mathcal{S}|$ bits can be mapped onto conventional complex information symbols $s_{i,b}$ drawn from a PSK or QAM constellation \mathcal{S} of cardinality $|\mathcal{S}|$, with $i = 1, 2$ denoting

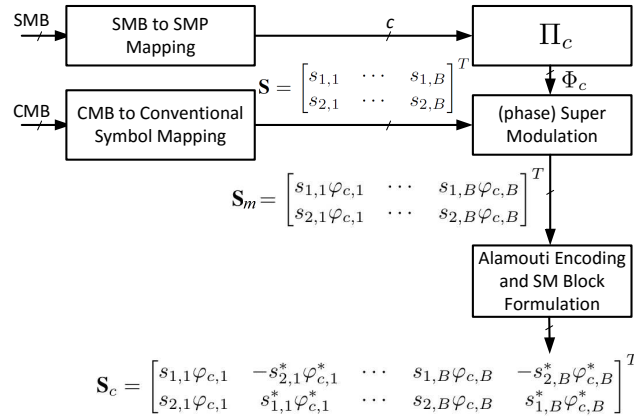


Fig. 1: Phase STSM scheme for 2×2 Alamouti space-time block code.

the antenna index and $b = 1, \dots, B$. Then, the conventionally modulated word is

$$\mathbf{S} = \begin{bmatrix} s_{1,1} & \dots & s_{1,B} \\ s_{2,1} & \dots & s_{2,B} \end{bmatrix}^T. \quad (1)$$

2) *SMB to Super-Modulation-Pattern (SMP) Mapping*: If C SMPs are available, $\log_2(\lfloor C \rfloor_{2^N})$ SMBs are transmitted per block with an appropriate SMB-to-SMP mapping, with $\lfloor C \rfloor_{2^N}$ being the maximum power of 2 not exceeding C .

3) *Phase Super-Modulation*: Each SMP c is related to a unique *characteristic SMP vector* $\mathbf{\Pi}_c$ of length B . This vector is introduced to describe how the produced complex information symbols will be super-modulated. To produce the super-modulated symbols, it is assumed that each symbol can be further modulated by using one of the M_{SM} , predefined *super-modulation states*. For phase STSM, these states are pre-defined distinct phase rotations. Then, if c is the SMP to be transmitted, the symbols $s_{i,b}$, with $i = 1, 2$ and $b = 1, \dots, B$ will be super-modulated using the SM state (e.g., phase rotations) given by the b -th element of $\mathbf{\Pi}_c$. For example, if a phase super-modulation scheme with $M_{SM} = 2$ (i.e., available phase rotations) is employed and if $\mathbf{\Pi}_c(4) = 2$, the symbols $s_{1,4}$ and $s_{2,4}$ will be phase super-modulated by using the second available phase rotation. More specifically, with phase super-modulation, the resulting symbol is

$$s_{i,b}^{(c)} = s_{i,b} \varphi_{c,b}. \quad (2)$$

For symmetric M -PSK modulations with the minimum phase distance between symbol constellations being $\phi_{\min} = 2\pi/M$, the phase rotation can be

$$\varphi_{c,b} = \exp\{j\phi_{c,b}\} = \exp\left\{j \frac{2\pi}{M_{SM}M} [\mathbf{\Pi}_c(b) - 1]\right\}. \quad (3)$$

It is noted that the phase rotations are such that the phase modulated symbols over different $b = 1, \dots, B$ do not coincide for any possible conventionally transmitted symbol. This attribute makes the different SMPs distinguishable at the receiver side.

It can be easily observed that the maximum number of the available SMPs is a function of the available modulation states. In particular, the number of candidate SMPs for M_{SM} available modulation states cannot be larger than M_{SM}^B . However, and as described in detail in Section III, increasing M_{SM} for a fixed B will result in a larger number of SMPs but of smaller “effective distance” and therefore of reduced identifiability (i.e., detection quality) at the receiver side. Therefore, even if a very large number of SMPs is available, only a subset of them will be finally employed, such that their “effective distance” is large, and therefore their decoding quality is high. In other words, the number of bits which can be efficiently super-modulated, is not determined by the number of the available SMPs, but by the “effective distances” between the finally selected SMPs which need to be efficiently chosen so that the detection quality is high. In Section III, we describe in detail how such an efficient SMPs selection and mapping is achieved.

B. Alamouti Encoding and SM Block Formulation

Eventually, according to the Alamouti space-time block code, the corresponding redundant information for each pair of the $s_{i,b}^{(c)}$ symbols over different b indices is calculated as an orthogonal transformation of these symbols. Then, without their exact positioning affecting the performance of the proposed scheme, it is assumed that the actual information of the b -th pair of symbols is transmitted over the $t = 2b - 1$ channel use and the corresponding redundant information is transmitted over the $t = 2b$ channel use. Therefore, the transmitted super-modulated word employing the c -th pattern is

$$\begin{aligned} \mathbf{S}_c &= \begin{bmatrix} s_{1,1}^{(c)} & -s_{2,1}^{(c)*} & \dots & s_{1,B}^{(c)} & -s_{2,B}^{(c)*} \\ s_{2,1}^{(c)} & s_{1,1}^{(c)*} & \dots & s_{2,B}^{(c)} & s_{1,B}^{(c)*} \end{bmatrix}^T \\ &= \begin{bmatrix} s_{1,1}\varphi_{c,1} & -s_{2,1}^*\varphi_{c,1}^* & \dots & s_{1,B}\varphi_{c,B} & -s_{2,B}^*\varphi_{c,B}^* \\ s_{2,1}\varphi_{c,1} & s_{1,1}^*\varphi_{c,1}^* & \dots & s_{2,B}\varphi_{c,B} & s_{1,B}^*\varphi_{c,B}^* \end{bmatrix}^T \end{aligned} \quad (4)$$

where $\varphi_{c,b}$ is given by (3). Then, it can be easily verified that the proposed scheme preserves the structure of Alamouti space-time block code and therefore the corresponding diversity gain.

Example 1: A phase STSM scheme with $L = 2B = 8$, $n_{sm} = 2$, $M_{SM} = 2$, and BPSK conventionally modulated symbols is considered. The $L = 8$ CMBs are mapped onto eight

BPSK symbols per block. Let us assume that SMBs and CMBs are “01” and “01101001”, respectively. Hence, the conventional modulated symbols are given by

$$\mathbf{S} = \begin{bmatrix} 1 & -1 & -1 & 1 \\ -1 & 1 & 1 & -1 \end{bmatrix}^T$$

Since $M_{SM} = 2$, $2^2 = 4$ SMPs (and therefore characteristic SM vectors) are available, allowing the transmission of at most two SMBs via STSM, or a maximum of a 0.25 bits per channel use (or 25%) throughput increase. Each pair of SMBs is then mapped onto a pattern, which is then mapped onto a characteristic SM vector. For this example, the following mapping can take place: “00” $\rightarrow c = 1 \rightarrow \mathbf{\Pi}_1 = [1 \ 1 \ 1 \ 1]^T$, “01” $\rightarrow c = 2 \rightarrow \mathbf{\Pi}_2 = [1 \ 2 \ 1 \ 2]^T$, “11” $\rightarrow c = 3 \rightarrow \mathbf{\Pi}_3 = [2 \ 2 \ 2 \ 2]^T$, “10” $\rightarrow c = 4 \rightarrow \mathbf{\Pi}_4 = [2 \ 1 \ 2 \ 1]^T$. The selected $\mathbf{\Pi}_c$ will be used to phase SM the conventionally modulated symbols according to (3). The exact mapping rule is later described in Section III. Since SMBs are “01”, the $\mathbf{\Pi}_2$ and therefore (according to (3)) the phase rotations $\varphi_{2,1} = 1$, $\varphi_{2,2} = j$, $\varphi_{2,3} = 1$, and $\varphi_{2,4} = j$ are chosen. Hence, the SM word is given by

$$\mathbf{S}_m = \begin{bmatrix} 1 & -j & -1 & j \\ -1 & j & 1 & -j \end{bmatrix}^T$$

The transmitted STSM codeword, using (4), is obtained as follows

$$\mathbf{S}_c = \begin{bmatrix} 1 & 1 & -j & j & -1 & -1 & j & -j \\ -1 & 1 & j & j & 1 & -1 & -j & -j \end{bmatrix}^T.$$

C. STSM Receiver Processing

The transmission channel \mathbf{H} consisting of the subchannels $H_{m,n}$, from Tx antenna m to Rx antenna n , is assumed static for the duration of a block transmission. The received $2B \times 2$ signal \mathbf{Y} can be described as

$$\mathbf{Y} = \mathbf{S}_c \mathbf{H} + \mathbf{N} \quad (5)$$

where \mathbf{N} is the $2B \times 2$ noise matrix consisting of independent and identically distributed (i.i.d.), zero-mean, complex Gaussian samples with variance $2\sigma_n^2$. Then, the maximum-likelihood (ML) detector of the transmitted word is given by

$$\hat{\mathbf{S}}_c = \arg \min_{\mathbf{S}_c \in \mathcal{W}} \{\mathcal{M}(\mathbf{S}_c)\}$$

with

$$\mathcal{M}(\mathbf{S}_c) = \|\mathbf{Y} - \mathbf{S}_c \mathbf{H}\|^2. \quad (6)$$

and \mathcal{W} being the set of all possible super-modulated words. The above minimization problem typically involves exhaustive calculation over all possible words, namely, over all possible transmitted symbols and SMPs, which is typically of prohibitive complexity. In order to reduce Rx complexity, it can be easily shown after some algebraic manipulations that for a specific SMP c , the corresponding ML metric $\mathcal{M}(\mathbf{S}_c)$ can be expressed as

$$\mathcal{M}(\mathbf{S}_c) = \sum_{b=1}^B \|\tilde{\mathbf{Y}}_b - \varphi_{c,b} \tilde{\mathbf{H}} \tilde{\mathbf{S}}_b\|^2 \quad (7)$$

where $\tilde{\mathbf{Y}}_b = [\mathbf{Y}[2b-1, 1] \quad \mathbf{Y}^*[2b, 1] \quad \mathbf{Y}[2b-1, 2] \quad \mathbf{Y}^*[2b, 2]]^T$, $\tilde{\mathbf{S}}_b = [s_{1,b} \quad s_{2,b}]^T$ and

$$\tilde{\mathbf{H}} = \begin{bmatrix} \tilde{\mathbf{h}}_1 & \tilde{\mathbf{h}}_2 \end{bmatrix} = \begin{bmatrix} H_{1,1} & H_{2,1} \\ H_{2,1}^* & -H_{1,1}^* \\ H_{1,2} & H_{2,2} \\ H_{2,2}^* & -H_{1,2}^* \end{bmatrix}.$$

Then, since the terms summed in (7) are independent of each other, the corresponding minimization can be achieved through the minimization of each term. Therefore, the conventionally modulated symbols which minimize $\mathcal{M}(\mathbf{S}_c)$ for a given c can be calculated as

$$\begin{aligned} \hat{\mathbf{S}}_b &= \arg \min_{\tilde{\mathbf{S}}_b \in \mathcal{S}^2} \mathcal{M}_{c,b} \\ &= \arg \min_{\tilde{\mathbf{S}}_b \in \mathcal{S}^2} \|\tilde{\mathbf{Y}}_b - \varphi_{c,b} \tilde{\mathbf{H}} \tilde{\mathbf{S}}_b\|^2, \forall b = 1, \dots, B. \end{aligned} \quad (8)$$

The corresponding minimum metric value for the specific SMP c is hence calculated as

$$\begin{aligned} \mathcal{M}_{min}(c) &= \min \left\{ \sum_{b=1}^B \|\tilde{\mathbf{Y}}_b - \varphi_{c,b} \tilde{\mathbf{H}} \tilde{\mathbf{S}}_b\|^2 \right\} \\ &= \sum_{b=1}^B \|\tilde{\mathbf{Y}}_b - \varphi_{c,b} \tilde{\mathbf{H}} \hat{\mathbf{S}}_b\|^2. \end{aligned} \quad (9)$$

The exhaustive search over all possible constellation symbols in (8) can be avoided by QR decomposition of the channel $\tilde{\mathbf{H}}$ as

$$\tilde{\mathbf{H}} = \mathbf{Q} \begin{bmatrix} \mathbf{R} \\ \mathbf{0}_{2 \times 2} \end{bmatrix} \quad (10)$$

where $\mathbf{Q} = \begin{bmatrix} \mathbf{Q}_1 & \mathbf{Q}_2 \end{bmatrix}$ is a unitary 4×4 matrix consisting of two 4×2 sub-matrices \mathbf{Q}_1 and \mathbf{Q}_2 , \mathbf{R} is a 2×2 upper triangular matrix with real-valued positive diagonal entries, and $\mathbf{0}_{2 \times 2}$ is a 2×2 zero matrix. Then,

$$\begin{aligned} \|\tilde{\mathbf{Y}}_b - \varphi_{c,b} \tilde{\mathbf{H}} \tilde{\mathbf{S}}_b\|^2 &= \left\| \tilde{\mathbf{Y}}_b - \varphi_{c,b} \begin{bmatrix} \mathbf{Q}_1 & \mathbf{Q}_2 \end{bmatrix} \begin{bmatrix} \mathbf{R} \\ \mathbf{0} \end{bmatrix} \tilde{\mathbf{S}}_b \right\|^2 \\ &= \left\| \begin{bmatrix} \mathbf{Q}_1^* \\ \mathbf{Q}_2^* \end{bmatrix} \tilde{\mathbf{Y}}_b - \varphi_{c,b} \begin{bmatrix} \mathbf{R} \\ \mathbf{0} \end{bmatrix} \tilde{\mathbf{S}}_b \right\|^2 \\ &= \|\mathbf{Q}_1^* \tilde{\mathbf{Y}}_b - \varphi_{c,b} \mathbf{R} \tilde{\mathbf{S}}_b\|^2 + \|\mathbf{Q}_2^* \tilde{\mathbf{Y}}_b\|^2. \end{aligned} \quad (11)$$

The second term in (11) is not a function of the symbols that need to be decoded. Due to the orthogonality of the code and using the Gram-Schmidt method to calculate \mathbf{Q}_1 and \mathbf{R} , the conventionally modulated symbols for the specific SMP can then be decoded as

$$\hat{\tilde{\mathbf{S}}}_b = \arg \min_{\mathbf{s}_{i,b} \in \mathcal{S}} \sum_{i=1}^2 \|\tilde{\mathbf{h}}_i^H \tilde{\mathbf{Y}}_b - \varphi_{c,b} \mathcal{E}_H \mathbf{s}_{i,b}\|^2 \quad (12)$$

where

$$\mathcal{E}_H = \sum_{k=1}^2 \sum_{l=1}^2 |H_{k,l}|^2 \quad (13)$$

is the energy of the multiple-input multiple-output (MIMO) channel. Since the corresponding symbols in each of the sums in (12) are independent

$$\hat{s}_{i,b} = \text{demod} \left\{ \frac{\tilde{\mathbf{h}}_i^H \tilde{\mathbf{Y}}_b}{\varphi_{c,b} \mathcal{E}_H} \right\}; i = 1, 2; b = 1, \dots, B. \quad (14)$$

where $\text{demod}\{R\}$ represents the typical constellation demodulator (i.e., slicer) which exploits the geometrical properties of the constellation to find the symbol closest to the point R and thus avoids performing exhaustive search over all possible symbols. Consequently, after estimating the corresponding symbols using (14), the $\mathcal{M}_{\min}(c)$ can be calculated using (9) for each SMP. Finally, denoting the set of all possible SMPs by \mathcal{C} , the ML solution will appear as

$$\hat{c} = \arg \min_{c \in \mathcal{C}} \mathcal{M}_{\min}(c). \quad (15)$$

For the decoding of the rateless coded information, soft-information-based sum-product rateless decoder is employed. From (12) the Log-Likelihood Ratios (LLRs) of the CMBs can be calculated as

$$L_{(s_{i,b})} = \ln \left(\frac{\sum_{s_{i,b} \in S_b^0} \exp \left[-\frac{\|\tilde{\mathbf{h}}_i^H \tilde{\mathbf{Y}}_b - \hat{\varphi}_{c,b} \mathcal{E}_H s_{i,b}\|^2}{2\mathcal{E}_H \sigma_n^2} \right]}{\sum_{s_{i,b} \in S_b^1} \exp \left[-\frac{\|\tilde{\mathbf{h}}_i^H \tilde{\mathbf{Y}}_b - \hat{\varphi}_{c,b} \mathcal{E}_H s_{i,b}\|^2}{2\mathcal{E}_H \sigma_n^2} \right]} \right) \quad (16)$$

where $\hat{\varphi}_{c,b}$ is the phase rotation obtained from SMP \hat{c} (given by (15)), and S_b^0 and S_b^1 are the subsets of possible symbols that have the b^{th} bit equal to 0 and 1, respectively.

D. Blind CMB Detection

As discussed in Section I, the proposed approach allows to multiplex two logical (information) subchannels, namely the SMB and CMB subchannels. Then, the joint optimal detection of the two subchannels requires the knowledge of the block size L and the exact SMB-to-SMP mapping function. However, while the detection of the SMB subchannel is not feasible without this knowledge, the detection of the CMB subchannel is still feasible with a performance loss. In particular, it can be (sub-optimally) assumed that all possible SMPs (and not only a subset) are employed for STSM. Then, following the aforementioned detection approach, the candidate vectors of conventionally modulated information, $\hat{\mathbf{S}}^{(m)} = [\hat{s}_1^{(m)} \ \hat{s}_2^{(m)}]^T$ for each modulation state $m = 1, \dots, M_{SM}$, with a phase rotation of φ_m , can be detected independently for each b by (14). Then, the transmitted modulation state is

$$\hat{m} = \arg \min_m \left\| \tilde{\mathbf{Y}} - \varphi_m \tilde{\mathbf{H}} \hat{\mathbf{S}}^{(m)} \right\|^2 \quad (17)$$

and therefore $\hat{s}_i = \hat{s}_i^{(\hat{m})}$ (with $i = 1, 2$). This ability to blindly decode the CMBs can be explored in various ways. For example, it allows SMB detection from only the receivers which are aware of the CMB block size and the SMB-to-SMP function, without preventing the CMB detection from all users. In addition, it allows the detection of the conventionally modulated information, even for those users where the initial assumption of static channel per block does not hold.

E. Complexity requirements

Typically, the ML detection via exhaustive calculations of (6) requires $8L$ complex multiplications to calculate the Frobenius metric. Therefore, since for n_{sm} bits transmitted via STSM and a constellation cardinality of $|\mathcal{S}|$, $2^{n_{sm}} |\mathcal{S}|^L$ metric calculations are required, the complexity would be

$$J_{ex} = 2^{n_{sm}+3} L |\mathcal{S}|^L \quad (18)$$

complex multiplications. For example for BPSK modulation, $L = 16$, $n_{sm} = 4$ (i.e., throughput increase of 25%), $1.3 \cdot 10^8$ complex multiplications are required, which makes such a decoding approach of prohibitive complexity.

The calculation of (14) for each $b = 1, \dots, B = L/2$ requires 12 complex multiplications/divisions. In addition, the norm calculation in (9) requires 14 complex multiplications (if $\varphi_{c,b}$ is first multiplied with $\hat{\mathbf{S}}_b$). Therefore, the complexity is $13L2^{n_{sm}}$ complex multiplications, where n_{sm} is the number of SMBs. However, independent of the number of available SMPs and for each b value, M_{\min} can take only as many values as number of phase modulation states (M_{SM}). Therefore, the complexity can be calculated to be $13LM_{SM}$ complex multiplications. Then, the complexity of the proposed scheme can be calculated as

$$J_p = \min\{13L2^{n_{sm}}, 13LM_{SM}\}. \quad (19)$$

However, as it is later discussed, the M_{SM} value can be kept low (e.g., $M_{SM} = 2$), so the overall complexity is manageable. For the previous example, the complexity of the proposed scheme is 416 complex multiplications, which in contrast to the exhaustive search makes its implementation feasible. Finally, the corresponding complexity for a conventional Alamouti space-time scheme, over the same block, can be calculated by (14) as $J_{conv} = 5L$ complex multiplications. Therefore, for $M_{SM} = 2$ the complexity of the proposed STSM detection scheme can be reduced to only 5.2 times the conventional one, independent of the number of SMBs. On the other hand, the need to store the SMP patterns results in increased memory requirements, which, however, can be kept small since the patterns consist of integer (and also binary in the case of $M_{SM} = 2$) values. For the simulation evaluations of Section V, $M_{SM} = 2$ is assumed.

III. SMP SET SELECTION

A. Effective Distance Criterion

In order to efficiently design rules capable of providing low (uncoded) BER, the determinant design criterion of [32] is employed. According to [32], the probability of erroneously detecting the word $\mathbf{S}_n^{(u)}$ (consisting of the u -th conventionally modulated word \mathbf{S} and the n -th SMP, see (4)) when $\mathbf{S}_m^{(v)}$ has been transmitted over a Rician channel, is a function of their “effective word distance”, defined as

$$d^2(\mathbf{S}_m^{(v)}, \mathbf{S}_n^{(u)}) = \det \left\{ \left(\Delta \mathbf{S}_{m,n}^{(v,u)} \right)^T \left(\Delta \mathbf{S}_{m,n}^{(v,u)} \right)^* \right\} \quad (20)$$

where $\det\{\cdot\}$ denotes the matrix determinant and

$$\Delta \mathbf{S}_{m,n}^{(v,u)} = \mathbf{S}_m^{(v)} - \mathbf{S}_n^{(u)} \quad (21)$$

Therefore, according to the determinant criterion, the minimum effective distance over all word pairs should be maximized. Then, (4) results in

$$\Delta \mathbf{S}_{m,n}^{(v,u)} = \begin{bmatrix} s_{1,1}^{(v)}\varphi_{m,1} - s_{1,1}^{(u)}\varphi_{n,1} & -\left(s_{2,1}^{(v)}\varphi_{m,1} - s_{2,1}^{(u)}\varphi_{n,1}\right)^* & \cdots \\ s_{2,1}^{(v)}\varphi_{m,1} - s_{2,1}^{(u)}\varphi_{n,1} & \left(s_{1,1}^{(v)}\varphi_{m,1} - s_{1,1}^{(u)}\varphi_{n,1}\right)^* & \cdots \end{bmatrix}^T \quad (22)$$

which can easily be verified to preserve the Alamouti STBC structure and diversity gain. Therefore the matrix $\left(\Delta \mathbf{S}_{m,n}^{(v,u)}\right)^T \left(\Delta \mathbf{S}_{m,n}^{(v,u)}\right)^*$ is diagonal, and the effective distance can be easily calculated as

$$d^2 \left(\mathbf{S}_m^{(v)}, \mathbf{S}_n^{(u)} \right) = \left(\sum_{b=1}^B \sum_{i=1}^2 \left| s_{i,b}^{(v)}\varphi_{m,b} - s_{i,b}^{(u)}\varphi_{n,b} \right|^2 \right)^2 \quad (23)$$

which is a function of both the conventionally transmitted symbols and the corresponding SMP. Due to (3), for constant-amplitude, symmetric constellations of M symbols (e.g., M -PSK) and symbol energy E_S , it can be easily shown that

$$d^2 \left(\mathbf{S}_m^{(v)}, \mathbf{S}_n^{(u)} \right) = \left(E_S \sum_{b=1}^B \sum_{i=1}^2 \left| 1 - e^{j(k_i(b)\phi_{\min} + \Delta\phi_b)} \right|^2 \right)^2 \quad (24)$$

where $\Delta\phi_b = \phi_{n,b} - \phi_{m,b}$ and $k_i(b)$ is the phase difference of the $s_{i,b}^{(u)}$ and $s_{i,b}^{(v)}$ symbols in integer multiples of the minimum distance between constellation symbols ϕ_{\min} , or

$$k_i(b) = \frac{\angle \left\{ s_{i,b}^{(u)} / s_{i,b}^{(v)} \right\}}{\phi_{\min}} \quad (25)$$

which is a function of the corresponding traditionally modulated symbols. The target is to find the set of SMPs which maximize the minimum $d^2 \left(\mathbf{S}_m^{(v)}, \mathbf{S}_n^{(u)} \right)$, over any word pair belonging into the set, independently of the conventionally modulated symbols. In this direction, it can be easily verified that

$$I_{(m,n)}(b) = \min_{k(b)=0,\dots,M-1} \left\{ \left| 1 - e^{j(k(b)\phi_{\min} + \Delta\phi_b)} \right|^2 \right\} \quad (26)$$

$$= \begin{cases} \left| 1 - e^{j\Delta\phi_b} \right|^2, & |\Delta\phi_b| \leq \frac{\phi_{\min}}{2} = \frac{\pi}{M} \\ \left| 1 - e^{j(\phi_{\min} - |\Delta\phi_b|)} \right|^2, & \text{else} \end{cases} \quad (27)$$

Therefore, the minimum distance between the m -th and n -th pattern is

$$d_{\min}^2(m, n) = d_{\min}^2(n, m) = \left(2E_S \sum_{b=1}^B I_{(m,n)}(b) \right)^2 \quad (28)$$

Then, according to the determinant criterion, for specific number of SMBs n_{sm} and block size L , the employed subset $C_{n_{sm}}$ of SMPs of size $2^{n_{sm}}$ should be the one maximizing the minimum $d_{\min}^2(m, n)$ over SMP pairs. Equivalently, the selected set of SMPs should be the one maximizing

$$D^2 = \left(\min_{m, n \in C_{n_{sm}}, m \neq n} \left\{ 2E_S \sum_{b=1}^B I_{(m,n)}(b) \right\} \right)^2 \quad (29)$$

From the above equations, it becomes apparent that increasing the size of a block, while keeping all the other parameters fixed, can result in increased minimum effective distance and therefore improved STSM codeword detection performance. Also, from (29), it becomes apparent that in order to efficiently utilize the available block length, each $b = 1, \dots, B$ should have at least two states. Otherwise the corresponding $I_{(m,n)}(b)$ values will be always zero, resulting in smaller D^2 .

Reducing the word detection error rate does not necessarily result in lower bit error rate (BER). To achieve this, and as discussed later in detail, an efficient SMB-to-SMP mapping is necessary requiring the SMPs with the smaller effective distances (and therefore of larger probability of appearance) to differ in as less bits as possible, similar to Gray coding.

B. SMP Set Selection and SMB-to-SMP mapping

Finding the SMP function which maximizes D^2 is a non-linear optimization problem, involving B values as well as the number of available states M_{SM} (and therefore the corresponding available phases) per characteristic pattern vector element b . Solving this optimization problem is a very tedious task not only analytically but also numerically. In particular, for a block length of $L = 2B$, M_{SM} available states and n_{sm} bits to be transmitted via phase STSM, there are

$$\binom{M_{SM}^{\frac{L}{2}}}{2^{n_{sm}}} = \frac{\left(M_{SM}^{\frac{L}{2}}\right)!}{\left(M_{SM}^{\frac{L}{2}} - 2^{n_{sm}}\right)! (2^{n_{sm}})!} \quad (30)$$

candidate SMP subsets. For example, even for very small block sizes e.g., $L = 16$, with $M_{SM} = 2$ and $n_{sm} = 3$, there are 4.09×10^{14} possible subsets. Since, as discussed, the optimal set of SMPs is difficult to find, a practical SMP selection and an efficient SMB-to-SMP mapping approach is herein proposed which can always guarantee a high effective distance (however not necessarily optimal). For M_{SM} available modulation states and a block size of $L = 2B$, a number of $\log_2 M_{SM}$ bits can be mapped onto each SMP element. Then, if the transmission of n_{sm} via phase STSM is targeted, each of the n_{sm} bits can be redundantly appear $B \log_2 M_{SM} / n_{sm}$ times in each pattern, which can increase its identifiability at the receiver side. However, increasing M_{SM} , reduces

the minimum non-zero $\Delta\varphi_b$ (see (3)) and therefore the minimum non-zero $I_{(m,n)}$ which affects negatively on the identifiability. However, it can be easily verified that the negative effect (on the D^2) when reducing $\Delta\varphi_b$ tends to be larger than the resulting gain after repetitively combining the sub-patterns.

For example, for BPSK modulation, if $M_{SM} = 2$, $B = 4$, and $n_{sm} = 2$, only one bit can be mapped onto each of the B SMP elements, and each bit can be redundantly appear $B/n_{sm} = 2$ times to increase its detection reliability at the receiver side. In order to maximize the non-common elements for the transmission of two SMB bits, and therefore minimize the number of zero $I_{m,n}$ terms, the following mapping can be used: “00” $\rightarrow \mathbf{\Pi}_1 = [1 \ 1 \ 1 \ 1]^T$, “01” $\rightarrow \mathbf{\Pi}_2 = [1 \ 2 \ 1 \ 2]^T$, “11” $\rightarrow \mathbf{\Pi}_3 = [2 \ 2 \ 2 \ 2]^T$, “10” $\rightarrow \mathbf{\Pi}_4 = [2 \ 1 \ 2 \ 1]^T$, where each element of SMP represents one of the $M_{SM} = 2$ possible modulation states. It can be observed that, unavoidably, two of the elements will be equal resulting in zero $I_{(m,n)}$. Hence, for $E_S=1$,

$$\begin{aligned} d_{\min}^2(m, n) &= \left(2 \sum_{b=1}^4 I_{(m,n)}(b)\right)^2 \geq \left(2 \left(2 \cdot 0 + 2 \min_{m,n} \{I_{m,n}\}\right)\right)^2 \\ &= 16 \left(\min_{m,n} \{I_{m,n}\}\right)^2 = 16 \left(\left|1 - e^{j\frac{\pi}{2}}\right|^2\right)^2 = 64. \end{aligned} \quad (31)$$

For $M_{SM} = 4$, two bits can be mapped onto each of the SMP elements and an example SMB-to-SMP mapping can be “00” $\rightarrow \mathbf{\Pi}_1 = [1 \ 1 \ 1 \ 1]^T$, “01” $\rightarrow \mathbf{\Pi}_2 = [2 \ 2 \ 2 \ 2]^T$, “11” $\rightarrow \mathbf{\Pi}_3 = [3 \ 3 \ 3 \ 3]^T$, “10” $\rightarrow \mathbf{\Pi}_4 = [4 \ 4 \ 4 \ 4]^T$, allowing for all the $I_{m,n}$ terms to be non-zero. Thus, similar to the $M_{SM} = 2$ case,

$$\begin{aligned} d_{\min}^2(m, n) &= \left(2 \sum_{b=1}^4 I'_{(m,n)}(b)\right)^2 \geq \left(2 \left(4 \min_{m,n} \{I'_{m,n}\}\right)\right)^2 \\ &= 64 \left(\min_{m,n} \{I'_{m,n}\}\right)^2 = 64 \left(\left|1 - e^{j\frac{\pi}{4}}\right|^2\right)^2 \approx 21.96. \end{aligned} \quad (32)$$

Consequently, we can come up with the practical guideline that for efficient SM transmission, the number of modulation states employed by SMPs should be kept minimum, but not less than two as discussed before. Based on this practical assumption, the case of $M_{SM} = 2$ is considered in the rest of this paper. For $M_{SM} = 2$ available states, two $\varphi_{c,b}$ values exist (0 and π/M) and therefore $I_{(m,n)}(b)$ can only take the values

$$I_{(m,n)}(b) = \begin{cases} I_{\min} = 0, & \Delta\varphi_b = 0 \\ I_{\max} = \left|1 - e^{j\frac{\pi}{M}}\right|^2, & \Delta\varphi_b = \frac{\pi}{M} \end{cases} \quad (33)$$

Therefore, increasing the Hamming distance between the possible SMPs, is equivalent to increasing their effective distance. Based on this observation, a simple and efficient SMB-to-SMP mapping is proposed which is an extension of the typical Gray coding for M -PSK schemes to the phase STSM case. In particular, in order to map n_{sm} bits onto SMPs, a Gray coding approach similar to that of the n_{sm} -PSK case. Therefore, consequent symbols are allowed to differ only in one bit. The mapping function between the i -th symbol and the vector of bits (\mathbf{b}) to be mapped onto this symbol is defined by $G_{(M_{SM})}(i) = \mathbf{b}^T$. If m is the modulus and r the remainder after the division B/n_{sm} (i.e., $m = \text{mod}[B, n_{sm}]$ and $r = \text{rem}[B, n_{sm}]$), any $\mathbf{\Pi}_i$ with $i = 1, \dots, 2^{n_{sm}}$ is constructed as follows

$$\mathbf{\Pi}_i = \left[\underbrace{G_{(M_{SM})}(i), \dots, G_{(M_{SM})}(i)}_{m \text{ times}} \quad G_{(r)}(\text{mod}[i-1, 2^r] + 1) \right]^T + 1. \quad (34)$$

This kind of mapping not only allows reaching large D^2 values, but also results in low BER performance since each of the most possible word errors (over consequent SMPs) results in only one bit error. In addition, increasing the block size, while keeping n_{sm} fixed, increases the D^2 and therefore, reduced the SMP error-rate.

Example 2: A phase STSM scheme with $L = 2B = 10$, $n_{sm} = 2$, $M_{SM} = 2$, and BPSK conventionally modulated symbols is considered. For modulating two bits (according to Gray coding for BPSK), $G_{(2)}(1) = [0 \ 0]$, $G_{(2)}(2) = [0 \ 1]$, $G_{(2)}(3) = [1 \ 1]$, $G_{(2)}(4) = [1 \ 0]$, while for modulating one bit $G_{(1)}(1) = 0$ and $G_{(1)}(2) = 1$. Hence, the following SMB-to-SMP mapping takes place “00” $\rightarrow \mathbf{\Pi}_1 = [1 \ 1 \ 1 \ 1 \ 1]^T$, “01” $\rightarrow \mathbf{\Pi}_2 = [1 \ 2 \ 1 \ 2 \ 2]^T$, “11” $\rightarrow \mathbf{\Pi}_3 = [2 \ 2 \ 2 \ 2 \ 1]^T$, “10” $\rightarrow \mathbf{\Pi}_4 = [2 \ 1 \ 2 \ 1 \ 2]^T$. If SMBs and CMBs are given by “01” and “0110100101”, respectively, the conventional modulated symbols are given by

$$\mathbf{S} = \begin{bmatrix} 1 & -1 & -1 & 1 & 1 \\ -1 & 1 & 1 & -1 & -1 \end{bmatrix}^T$$

and the $\mathbf{\Pi}_2$ is chosen. Hence, the SM word is given by

$$\mathbf{S}_m = \begin{bmatrix} 1 & -j & -1 & j & j \\ -1 & j & 1 & -j & -j \end{bmatrix}^T$$

The transmitted STSM codeword, using (4), is obtained as follows

$$\mathbf{S}_c = \begin{bmatrix} 1 & 1 & -j & j & -1 & -1 & j & -j & j & -j \\ -1 & 1 & j & j & 1 & -1 & -j & -j & -j & -j \end{bmatrix}^T.$$

IV. STSM FOR JOINT MEDIUM ACCESS AND RATELESS DATA TRANSMISSION

This section describes how STSM can be used in the context of MTC for one-shot, grant-free joint medium access and rateless data transmission, obviating the need for any registration process. The scenario considered here assumes multiple machines that want to communicate with a central access point. However, the approach can be extended to machines that communicate with each other in a non-centralized way (e.g., ad-hoc machine-type networks). It is also assumed that the transmission is ratelessly encoded and, therefore, the only feedback required is the ACK signals, that can also be eliminated if, instead of increasing throughput, we target increasing the probability of correct detection for a given number of transmissions. In the presence of ACK signals, we assume that they are transmitted via a dedicated control channel similar to [38], and that they are perfectly received. Our proposed scheme is not restricted to systems that use rateless, or any other specific family of codes. In practice, there is a plethora of ways to combine and detect the received information [33] (see Fig. 2). Still, rateless coding appears to be one of the most promising ones and, thus, it is employed here [34].

In this direction, and without loss of generality, we have here employed Raptor codes [35] since they are among the most widely used in the literature and among the most practical due to their low complexity, belief-propagation-based decoding. The coded information is modulated, space-time-encoded and transmitted in sets of blocks of a size of L symbols, as shown in Fig. 2. In rateless systems, the transmitted information packets need to be small to avoid transmitting unnecessary bits. Each machine can transmit the blocks either in a continuous manner, or in a random way, since the proposed approach supports both kinds of transmission. Each machine continues transmitting blocks related to the same information sequence, until it receives an ACK from the access point that the corresponding information sequence has been decoded.

To decode the received information the access point needs to know the ID of the machine that transmitted the packet, as well as its relative position in the encoded sequence (see Fig. 2) in order to efficiently combine it. Therefore, together with each packet, some *signature information* (*SI*) needs to be transmitted. In the examined case, this SI consists of two parts. The first set of n_{id} bits provides the ID of the transmitting machine and the second set of n_s bits is used to provide the order of the transmitted packet in the encoding sequence. The n_{id} bits can be either preallocated to machines or they can be randomly selected as in the case of mobile RACH. The way to allocate them and the corresponding consequences are beyond the scope of this work.

Several approaches can be used to transmit all or part of the signature bits. The first approach transmits a preamble (or header) before each data packet, with the preamble bits being encoded with some low rate code. Since advanced channel codes like LDPC codes are not appropriate for such small packet lengths, here traditional convolutional coding is assumed. In addition, for STSM to be applied, it is assumed that the coded packets are also space-time-encoded. A second approach, originates from the approaches currently employed in LTE, where the mobile RACH transmits dedicated preamble sequences that are orthogonal to each other. Similarly to mobile (LTE) RACH, preambles based on Zadoff-Chu (ZC) sequences are considered here. In the mobile RACH of the current LTE system, the eNodeB serves UEs with 64 fixed preambles [36][37]. The corresponding sizes of preambles can support the number of bits to be mapped. Specifically, to transmit q bits, it is required to map them to 2^q sequences of a length of at least 2^q transmission samples. Instead of using ZC sequences, one can use binary sequences with good cross-correlation properties that are based on the Gold Codes (GC), as has been proposed in [38] for transmitting ACK signals. All the aforementioned approaches require transmitting preambles, that as shown in Section V can significantly limit the achievable rate. Instead, the SMBs of STSM can be used to transmit part of or the whole SI, reducing or even eliminating the need for preambles. The trade-offs between these approaches are evaluated in Section V.

For evaluating the gains of STSM over preamble-based approaches, perfect channel estimation and synchronization are herein assumed, for all the evaluated schemes. In practice, however, short pilot sequences will need to be transmitted from each machine to the access point, for synchronization and channel estimation purposes, typically, as part of each transmitted packet. This pilot overhead, which is inherent of any practical coherent system, and it is not an overhead specifically related to STSM, is generally required from all examined schemes, and is typically small compared to the SI preamble overhead (please note that, for AWGN noise, the variance of optimal, unbiased estimators typically decreases by a factor of two any time pilot sequence increases by the same factor) [39]. In addition, the corresponding channel estimation error can be well approximated as additional AWGN noise [39], the variance of which depends only on the estimator and not on the employed method to transmit the SI bits. Therefore, the channel estimation error results in an error-rate degradation common to all the examined schemes. While the design of appropriate pilot sequences and estimation algorithms, and the evaluation of their performance is a very interesting topic, it is still beyond the scope of this work.

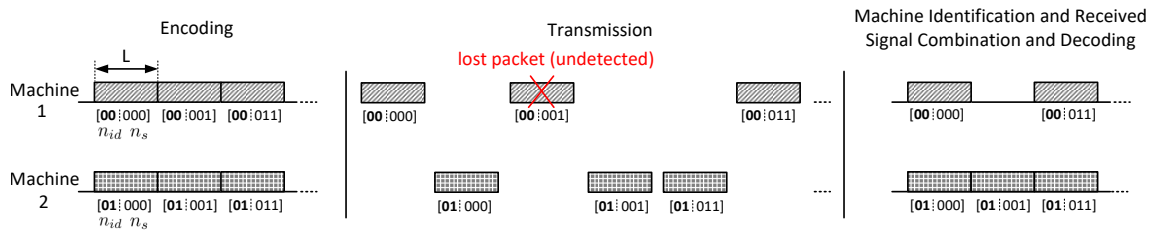


Fig. 2: Transmission of ratelessly coded information packets for multiuser systems in collision-free environment.

V. EVALUATION

Here, the concept of STSM is validated and its performance is evaluated via simulations. In Section V-A, it is shown that STSM enables the transmission of an additional low-rate and highly reliable information stream (i.e., SMB) on top of traditionally modulated and Space-Time encoded information stream (i.e., CMB), resulting in a significant throughput increase (e.g., 20% for transmitting blocks of 100 bits) without practically affecting the average transmitted error-rate. In the same section, it is shown that the reliability of SMBs can be consistently increased by increasing the block size L or reducing the SMBs n_{sm} , validating our SMB-to-SMP mapping. In addition, it is shown that when decoding the CMBs independently of the SMBs (i.e., blind STSM detection) this would entail a performance loss of about 1 dB only.

In Section V-B, the application of STSM in the context of MTCs is discussed. In particular, it is verified that the CMB subchannel can be efficiently exploited in order to transmit signature packets and therefore in order to enable joint medium access and rateless transmissions, while reducing or even eliminating preamble sequences.

In practical MTC schemes actual packet collisions may happen, especially if the transmission takes place in a grant-free manner. To evaluate the appropriateness of STSM in such practical systems, its performance is examined in the extreme case where two users always collide (similarly to a two-user multiple access channel). In particular, we focus our evaluation on the most challenging case where the two users collide in a synchronous manner. Namely, we focus on the case where for traditional, preamble-based approaches, either the preambles or the payload will interfere with each other. Still, for completeness, we also examine the error-rate performance of the signature information of preamble-based schemes, when the corresponding preambles interfere with data (i.e., in the case of asynchronous transmission), which as we show,

is the less challenging case for such approaches. It is significant to notice that when STSM-based methods are applied, it does not matter if the system is synchronous or asynchronous, since all cases, in the absence of preambles, only the payload part will interfere. Section V-C shows that STSM is robust to collisions, and by exploiting the “rateless” aspects of our system, significant throughput gains can be achieved compared to traditional time-division-multiple access (TDMA) systems that avoid collisions, as predicted in the framework of the multiple access channel.

For the conducted simulations, and since we focus on challenging transmission scenarios, BPSK is used for modulating the conventionally transmitted bits. Still, STSM is directly applicable to two-dimensional constellations. In all performance evaluations, the transmitted power is normalized to unity. We assume no channel knowledge at the transmitter but perfect knowledge at the receiver side. The 2×2 channel is modelled as a temporally and spatially uncorrelated frequency-flat Rayleigh channel, and remains constant within a block-size. For rateless systems, Raptor’s inner LT code is generated according to Raptor RFC 5053 standard [41], and rate 0.95 LDPC pre-code with left regular distribution (node degree 3 for all nodes) and right Poisson (check nodes chosen randomly with a uniform distribution) is used. Belief propagation decoding is performed with forty iterations [35].

A. Performance Gains of Uncoded STSM

In Fig. 3 (a), the total uncoded BER performance (for both CMBs and SMBs) of the STSM scheme is depicted for $L = 100$ and various n_{sm} values. It is illustrated that when transmitting $n_{sm} = 4$ additional bits, that corresponds to throughput increase of $G = 4\%$, no performance loss is observed. Also, it is shown that for $n_{sm} = 20$, a throughput gain of up to 20% can be attained without practically affecting the overall BER performance. Fig. 3 (b) shows the uncoded BER performance for each of the multiplexed information subchannels (i.e., CMBs and SMBs) of previous figure. It is shown that SMBs are more reliable than CMBs. For fixed L , the reliability of SMBs increases for lower n_{sm} due to the increase in Euclidean distance between codewords.

Fig. 4 shows the uncoded BER performance and associated throughput gains for the CMBs and SMBs of STSM with $n_{sm} = 12$ and several block lengths L . The error-rate performance of SMBs improves significantly as the block length is increased due to the increase in the Euclidean distance between codewords in the same manner as in previous figure i.e., as n_{sm} is lowered with fixed L . As L is increased, the detection reliability of conventional modulated symbols (i.e., CMBs) enhances and remains practically the same for block lengths higher than 100 bits.

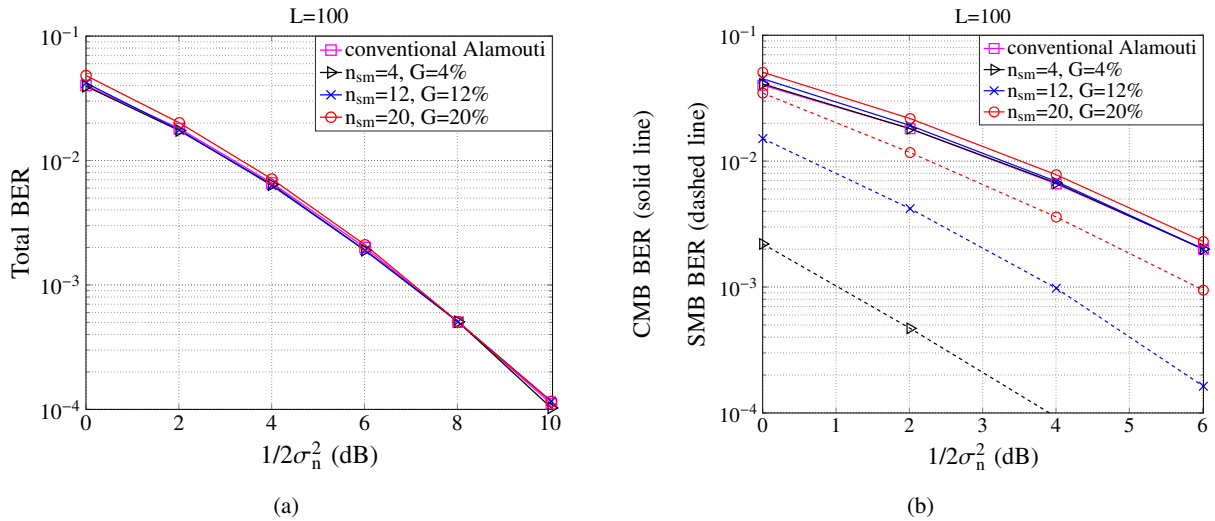


Fig. 3: BER performance of the STSM scheme for several n_{sm} values (resulting in throughput gains G) and $L = 100$ versus traditional Alamouti STBC: (a) Total BER (b) BER of the individual subchannels.

In Fig. 5, the performance of the STSM schemes with a throughput gain of $G = 4\%$ that can be achieved by several L and n_{sm} combinations is depicted. Here, three cases have been considered. Using the analysis in Section III-A, it can be easily verified that the achievable D^2 is the same as long as the n_{sm}/L ratio remains constant. It is verified in Fig. 5 that the performance is only a function of n_{sm}/L . We also showed in Section II-E that the complexity per channel user is independent of the L and n_{sm} values. Therefore, the critical design parameter is the n_{sm}/L ratio and not the exact L and n_{sm} . Furthermore, in Section II-D, it is described how the CMBs can be blindly (and therefore sub-optimally) decoded without knowing L and the employed SMB-to-SMP mapping function at the cost of a BER performance loss. In Fig. 6, this performance loss is evaluated to be around 1 dB for low to high SNR range.

B. Performance of STSM in a Multiuser (Collision-free) Environment

Fig. 7 compares the signature packet error rate (PER) performance of STSM and preamble-based approaches when transmitting $n_{sig} = 4$ and 6 signature bits per packet. In particular, the error-rate performance with STSM and blocks of $L = 200, 600,$ and 1000 against an approach that utilizes ZC preambles of $N_{pr} = 16$ and 64 samples (which is the minimum preamble that supports 4-bit and 6-bit packets) are compared. For the ZC sequences it is assumed that

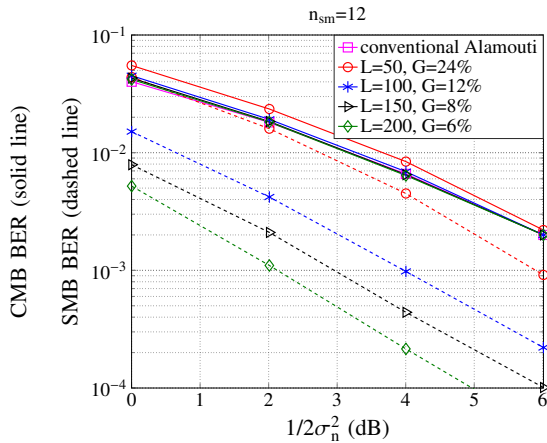


Fig. 4: BER performance of the CMBs and SMBs for several L (resulting in throughput gains G) and $n_{sm} = 12$ versus traditional Alamouti STBC.

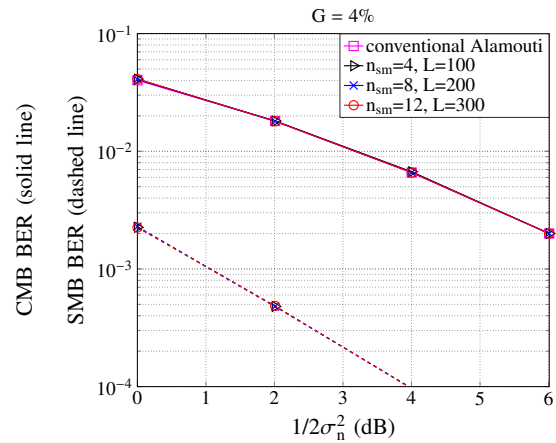


Fig. 5: BER performance of the CMBs and SMBs for constant n_{sm}/L resulting in throughput gain $G = 4\%$ using several n_{sm} and L values versus traditional Alamouti STBC.

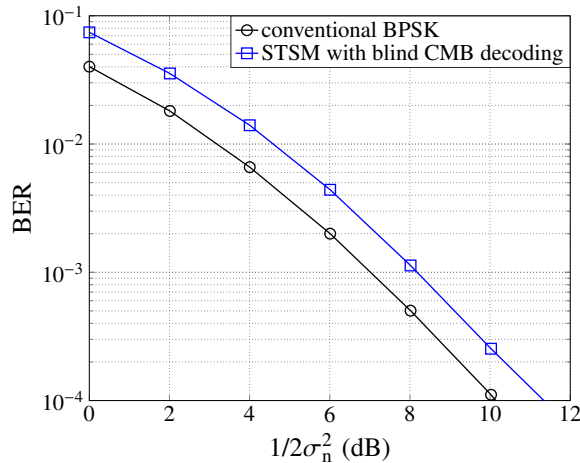


Fig. 6: BER performance of the blindly detected CMBs versus traditional Alamouti STBC.

the sequence is transmitted from one antenna, and at the receiver side coherent detection of the transmitted sequence takes place. The GC correlatable sequences are BPSK modulated and space-time encoded. For the ZC preambles a unity root index has been used and the chosen cyclic shift N_{cs} is set to one [37], to obtain minimum preamble size for supporting 4 and 6-bit packets i.e., $N_{pr}=16$ and 64 samples, respectively. The results for the GC preambles are only generated for $n_{sig} = 6$ (and $N_{pr} = 64$) since for base $2^q - 1$ of preferred pairs used for generation

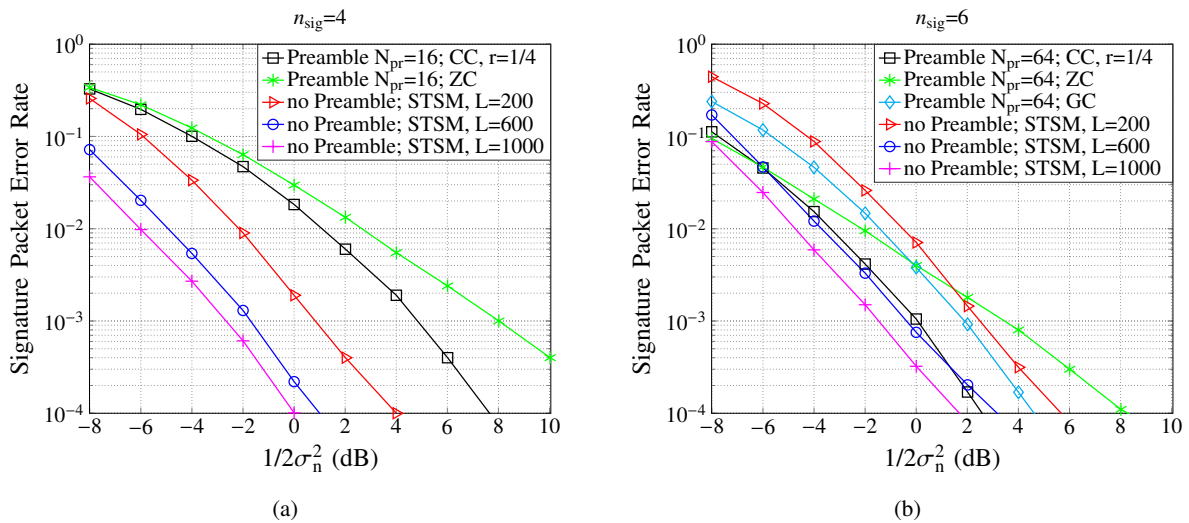


Fig. 7: Comparison of signature information PER for STSM (no preamble) against ZC, GC, and CC preamble-based schemes: (a) preambles with $N_{pr} = 16$ (b) preambles with $N_{pr} = 64$.

of GC, it is required that n is not divisible by 4 [40]. In addition, the signature PER performances of convolutionally coded (CC) preambles of size $N_{pr} = 16$ and 64, are shown when they are also Alamouti space-time encoded.

For the signature packet length of $n_{sm} = 4$, STSM results in superior performance compared to preamble-based schemes for broad range of SNRs. By comparing Figs. 7 (a) and (b), it is shown that while for the higher signature packet length $n_{sig} = 6$, the performance of preamble-based schemes improves, the performances of STSM-based schemes is degraded since the ratio n_{sm}/L becomes smaller (see Sec.III-A). However, even in such a case, as it later shown, STSM can result in throughput gains due to the elimination of the preamble overhead. For $n_{sm} = 6$, the STSM scheme with block of $L = 600$ yields almost the same performance as in CC scheme and better than GC and ZC schemes in most of the SNRs. Also, the STSM scheme with block size $L = 1000$ outperforms all examined preamble-based schemes.

Fig. 8 evaluates the achievable rate of one user in a collision-free environment when using Raptor rateless codes for different methods to encode each packet's SI. All simulations are conducted for message size of 1000 bits. The SI of $n_{sig} = 9$ bits consists of 6 n_{id} bits (to support 64 users as in the mobile RACH), and 3 n_s bits. Eight cases are considered, all targeting the efficient delivery of the $n_{sig} = 9$ signature bits. The case where the signature information is perfectly known, the case where all signature bits are super-modulated and no preamble is used

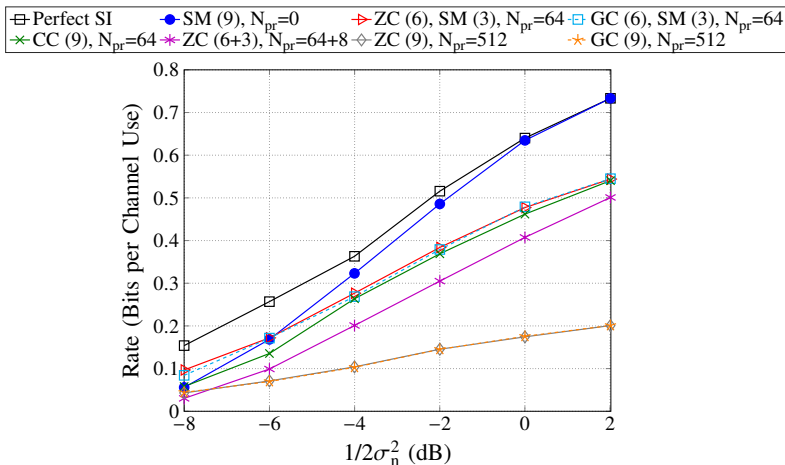


Fig. 8: Achievable rate of the rateless systems with $L = 200$, employing STSM schemes and preamble-based approaches in multiuser collision-free environment.

(denoted by $\text{SM}(9)$, $N_{pr} = 0$), the case where the nine SI bits are transmitted as a preamble of size $N_{pr} = 64$ after being BPSK modulated, convolutionally encoded with a rate $9/64$ and space-time encoded (denoted by $\text{CC}(9)$, $N_{pr} = 64$), the cases where all the nine bits are mapped on ZC or GC preambles of size $N_{pr} = 512$ (denoted by $\text{ZC/GC}(9)$, $N_{pr} = 512$), the case where, in order to reduce the ZC sequence size, we map the 6 n_{id} bits on a ZC of size $N_{pr} = 64$ and the 3 n_s bits on a ZC of size $N_{pr} = 8$, and we transmit them sequentially (denoted by $\text{ZC}(6+3)$ $N_{pr} = 64 + 8$), as well as the cases where a ZC or GC of $N_{pr} = 64$ is used for mapping the n_{id} bits, while the n_s bits are super-modulated (e.g., STSM is used to reduce preamble overhead)¹.

Fig. 8 shows that only STSM-based schemes can approach the “ideal” rateless throughput compared to all other solutions that are solely based on preambles. By super-modulating the signature information, significant throughput gains can be attained from low to high SNRs compared with all other solutions, and the gain reaches more than 35% at high SNRs. This gain is achieved despite the fact that the Signature-Packet Error Rate for STSM is worse than the preamble-based approach for the selected values of n_{sig} and L (see Fig. 7). If larger L (or smaller n_{sig}) values are used, the STSM-based user identification becomes more reliable, and the STSM-based methods outperform the preamble-based methods across the whole SNR regime.

¹In all cases, if the number of packets required to correctly decode the transmitted information exceeds the number of those that can be counted by the available n_s bits, the counting is re-initiated. In addition, if a signature packet sequence is found more than once, the most reliable (in terms of their soft metrics) is used.

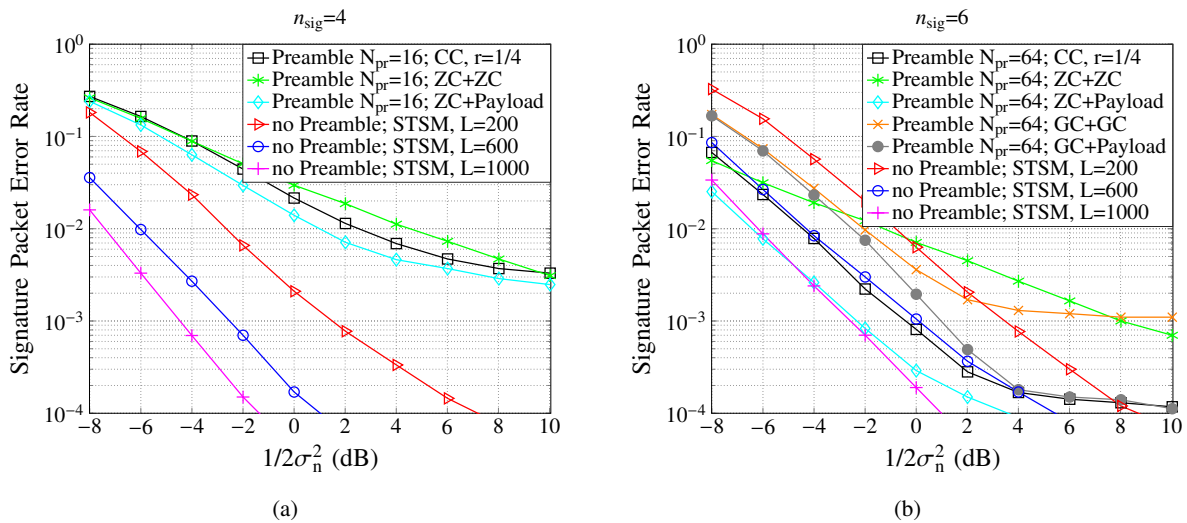


Fig. 9: Comparison of signature information PER for STSM (no preamble) against ZC, GC, and CC preamble-based schemes in multiuser environment with collisions: (a) preambles with $N_{pr} = 16$ (b) preambles with $N_{pr} = 64$.

C. Performance of STSM for Two-Colliding Users

In this section, the performance of STSM under collisions is evaluated. In particular, our target is to evaluate if STSM is capable of exploiting inherent SNR differences between machines/users, to reliably identify and then decode, first the “strong” user in terms of SNR, and then the “weak” one by means of successive-interference cancellation (SIC). Equivalently, we are evaluating if STSM can result in such a reliable SI identification that will enable realizing in practice gains that have been predicted in the theory of the multiple access channel [42]. As already mentioned, for the preamble-based approaches, we focus on the case where the collisions take place in a synchronous manner, but for completeness we also examine the error-rate performance of the signature information of preamble-based schemes, when the corresponding preambles interfere with data (i.e., in the case of asynchronous transmission). As discussed, when STSM-based methods are applied, it does not matter if the system is synchronous or asynchronous, since all cases, in the absence of preambles, only the payload part will interfere.

Fig. 9 shows the signature PER performance for $n_{sig} = 4$ and 6 signature bits per packet, respectively. The performance is shown for the strongest user since when collisions are happening this is the most likely user to be decoded. In addition, when SIC scheme takes place the

strongest user is decoded first. For the schemes that use ZC and GC preambles two scenarios are considered. The first scenario assumes synchronous transmission, where the preambles collide with each other (i.e., ZC+ZC and GC+GC), and the second scenario assumes asynchronous transmission, where the preamble of the strong user collides with the payload of the weak user (i.e., ZC+payload and GC+payload). Fig. 9 shows that the synchronous scenario is more challenging since the existence of multiple correlation peaks makes the user identification more challenging. It also shows that similarly to the collision-free case in Fig. 7, for $n_{sig} = 4$ STSM results in superior performance compared to preamble-based schemes. In addition, similarly to the collision-free case in Fig. 7, for $n_{sig} = 6$, while the performance of preamble-based schemes improves, the performance of STSM-based schemes is degraded due to the smaller n_{sm}/L ratio. Still, it is later shown that STSM can result in throughput gains due to the preamble elimination.

Similarly to Sec. V-B, the attainable sum-rate of two users employing rateless schemes with STSM in multiuser environments under collisions are compared with, synchronous, preamble-based approaches in Fig. 10. As also observed in the collision-free case, despite the fact that STSM-based user identification is not the most reliable for the specific selection of n_{sig} and L parameters, the throughput provided from STSM approaches is the closest to the ideal, due to the preamble elimination. In addition, it is shown that hybrid approaches that use two different methods to transmit SI can be significantly degraded if one of the identification methods is not highly reliable, due to error propagation (e.g., ZC(6+3) or ZC(6), SM(3)).

By exploiting the “rateless” properties of the proposed scheme additional gains can be attained by means of SIC. When the strongest user is successfully decoded, its transmitted signal is reconstructed and removed from the received signal. Then, the detection of the second user is re-attempted. Fig. 11 shows the achievable sum-rate for the two colliding users with and without SIC, and compares the results with a collision-free environments using TDMA [43]. STSM is used to super-modulate the SI bits in all cases. Fig. 11 shows that due to the rateless properties, we can always attempt to decode each user, while treating the other user as noise. Then, gains of up to 26% can be achieved compared to TDMA, where only one user is transmitting at each time instant and further gains of up to 25% can be achieved due to SIC.

VI. CONCLUSIONS

The concept of Space-Time Super-Modulation has been introduced that, for first time, enables joint medium access and rateless transmission for machine-type communications with reduced

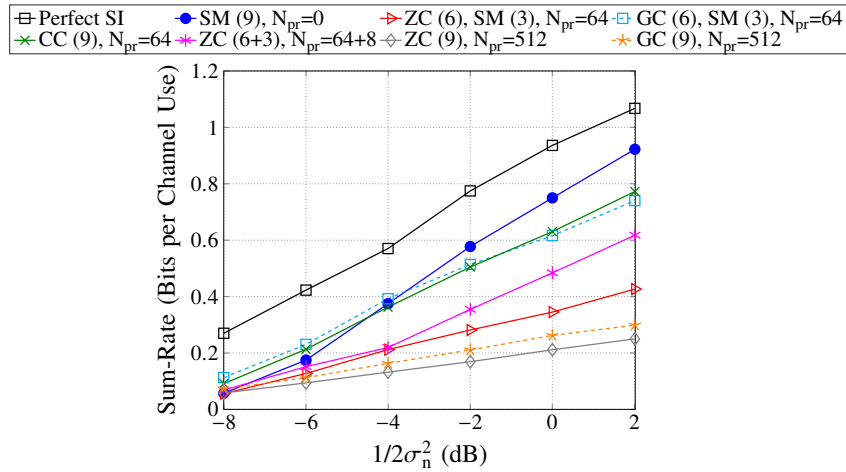


Fig. 10: Achievable sum-rate of the rateless systems with $L = 200$ employing STSM schemes and preamble-based approaches in multiuser environments with two colliding users.

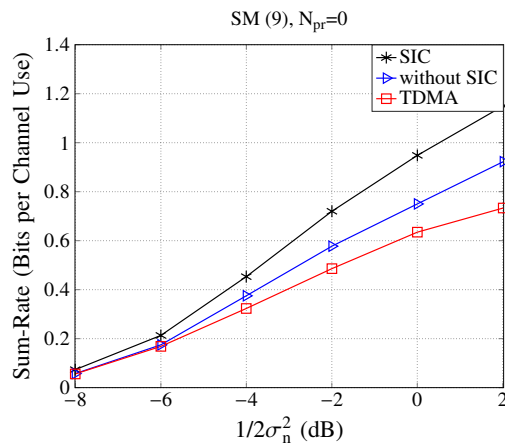


Fig. 11: Achievable sum-rate of the rateless systems with $L = 200$ employing STSM scheme in multiuser environments with two colliding users with and without SIC versus TDMA.

or even no preamble overhead. Due to its rateless properties, such a scheme can exploit collided packets resulting in significant throughput gains compared to systems that try to avoid collisions (e.g., when TDMA is applied), approaching the theoretical gains of multiple access channels.

ACKNOWLEDGMENT

The authors would like to thank the members of the 5G Innovation Centre (5GIC) of the University of Surrey (<http://www.surrey.ac.uk/5GIC>) for their support.

REFERENCES

- [1] A. Ali, et al., "Next generation M2M cellular networks: challenges and practical considerations," *IEEE Commun. Mag.*, vol. 53, no. 9, pp. 18-24, Sep. 2015.
- [2] H. Shariatmadari, et al., "Machine-type communications: current status and future perspectives toward 5G systems," *IEEE Commun. Mag.*, vol. 53, no. 9, pp. 10-17, Sep. 2015.
- [3] M. Hasan, et al., "Random access for machine-to-machine communication in LTE-Advanced networks: issues and approaches," *IEEE Commun. Mag.*, vol. 51, no. 6, pp. 86-93, Jun. 2013.
- [4] A. Rajandekar and B. Sikdar, "A survey of MAC layer issues and protocols for machine-to-machine communications," *IEEE Internet Things J.*, vol. 2, no. 2, pp. 175-186, Apr. 2015.
- [5] F. Ghavimi and H.-H. Chen, "M2M communications in 3GPP LTE/LTE-A networks: architectures, service requirements, challenges, and applications," *IEEE Commun. Surveys Tuts.*, vol. 17, no. 2, pp. 525-549, Second Quarter 2015.
- [6] S.-Y. Lien, et al., "Toward ubiquitous massive accesses in 3GPP machine-to-machine communications," *IEEE Commun. Mag.*, vol. 49, no. 4, pp. 66-74, Apr. 2011.
- [7] A. Laya, et al., "Is the random access channel of LTE and LTE-A suitable for M2M communications? a survey of alternatives," *IEEE Commun. Surveys Tuts.*, vol. 16, no. 1, pp. 4-16, First Quarter 2014.
- [8] G. Wu, et al., "M2M: from mobile to embedded internet," *IEEE Commun. Mag.*, vol. 49, no. 4, pp. 36-43, Apr. 2011.
- [9] S. Andreev, et al., "Understanding the IoT connectivity landscape: a contemporary M2M radio technology roadmap," *IEEE Commun. Mag.*, vol. 53, no. 9, pp. 32-40, Sep. 2015.
- [10] "Requirement analysis and design approaches for 5G air interface," *METIS Project Deliverable D2.1*, Aug. 2013.
- [11] G. Wunder, et al., "Sparse signal processing concepts for efficient 5G system design," *IEEE Access*, vol. 3, pp. 195-208, Mar. 2015.
- [12] "Final 5G NOW transceiver and frame structure concept D.3.3," *5G NOW Technical Report*, May 2015.
- [13] M. Wang, et al., "The evolution of LTE physical layer control channels," *IEEE Commun. Surveys Tuts.*, vol. 18, no. 2, pp. 1336-1354, Second Quarter 2016.
- [14] H. V. Nguyen, et al., "Near-capacity wireless system design principles," *IEEE Commun. Surveys Tuts.*, vol. 17, no. 4, pp. 1806-1833, Fourth Quarter 2015.
- [15] G. Wang, et al., "TiM: fine-grained rate adaptation in WLANs," *IEEE Trans. Mobile Comput.*, vol. 15, no. 3, pp. 748-761, Mar. 2016.
- [16] C.-Y. Li, et al., "An energy efficiency perspective on rate adaptation for 802.11n NIC," *IEEE Trans. Mobile Comput.*, vol. 15, no. 6, pp. 1333-1347, Jun. 2016.
- [17] L. Wang, et al., "Exploring smart pilot for wireless rate adaptation," *IEEE Trans. Wireless Commun.*, vol. 15, no. 7, pp. 4571-4582, Jul. 2016.
- [18] N. Bonello, et al., "Myths and realities of rateless coding," *IEEE Commun. Mag.*, vol. 49, no. 8, pp. 143-151, Aug. 2011.
- [19] N. Bonello, et al., "Low-density parity-check codes and their rateless relatives," *IEEE Commun. Surveys Tuts.*, vol. 13, no. 1, pp. 3-26, First Quarter 2011.
- [20] S. M. Alamouti, "A simple transmit diversity technique for wireless communications," *IEEE J. Sel. Areas Commun.*, vol. 16, no. 8, pp. 1451-1458, Oct. 1998.
- [21] V. Tarokh, H. Jafarkhani, and A. R. Calderbank, "Space-time block coding for wireless communications: performance results," *IEEE J. Sel. Areas Commun.*, vol. 17, no. 3, pp. 451-460, Mar. 1999.
- [22] ———, "Space-time block codes from orthogonal designs," *IEEE Trans. Inf. Theory*, vol. 45, no. 5, pp. 1456-1467, Jul. 1999.

- [23] H. Imai and S. Hirakawa, "A new multilevel coding method using error correcting codes," *IEEE Trans. Inf. Theory*, vol. 23, no. 3, pp. 371-377, May 1977.
- [24] U. Wachsmann, et al., "Multilevel codes: theoretical concepts and practical design rules," *IEEE Trans. Inf. Theory*, vol. 45, no. 5, pp. 1361-1391, Jul. 1999.
- [25] G. Ungerboeck, "Trellis coded modulation with redundant signal sets, part I: introduction," *IEEE Commun. Mag.*, vol. 25, no. 2, pp. 5-11, Feb. 1987.
- [26] ———, "Trellis coded modulation with redundant signal sets, part II: state of the art," *IEEE Commun. Mag.*, vol. 25, no. 2, pp. 12-21, Feb. 1987.
- [27] R. Y. Mesleh, et al. "Spatial modulation," *IEEE Trans. Veh. Technol.*, vol. 57, no. 4, pp. 2228-2241, Jul. 2008.
- [28] B. Wang, et al., "Dynamic compressive sensing-based multi-user detection for uplink grant-free NOMA," *IEEE Commun. Lett.*, vol. 20, no. 11, pp. 2320-2323, Nov. 2016.
- [29] Z. Zhang, et al., "Grant-free rateless multiple access: a novel massive access scheme for Internet of Things," *IEEE Commun. Lett.*, vol. 20, no. 10, pp. 2019-2022, Oct. 2016.
- [30] H. Jafarkhani and N. Seshadri, "Super-orthogonal space-time trellis codes," *IEEE Trans. Inf. Theory*, vol. 49, no. 4, pp. 937-950, Apr. 2003.
- [31] K. Nikitopoulos, F. Mehran, and H. Jafarkhani, "Space-time super-modulation and its application to joint medium access and rateless transmission," in *Proc. IEEE Globecom 2016*.
- [32] V. Tarokh, N. Seshadri, and A. R. Calderbank, "Space-time codes for high data rate wireless communication: performance analysis and code construction," *IEEE Trans. Inf. Theory*, vol. 44, no. 2, pp. 744-765, Mar. 1998.
- [33] H. Chen, et al., "A survey and tutorial on low-complexity turbo coding techniques and a holistic hybrid ARQ design example," *IEEE Commun. Surveys Tuts.*, vol. 15, no. 4, pp. 1546-1566, Fourth Quarter 2013.
- [34] F. Mehran, K. Nikitopoulos, et al., "Rateless wireless systems: gains, approaches, and challenges," in *Proc. 3rd IEEE ChinaSIP*, 2015.
- [35] O. Etesami and A. Shokrollahi, "Raptor codes on binary memoryless symmetric channels," *IEEE Trans. Inf. Theory*, vol. 52, no. 5, pp. 2033-2051, May 2006.
- [36] H. S. Jang, et al., "Message-embedded random access for cellular M2M communications," *IEEE Commun. Lett.*, vol. 20, no. 5, pp. 902-905, May 2016.
- [37] H. S. Jang, et al., "Spatial group based random access for M2M communications," *IEEE Commun. Lett.*, vol. 18, no. 6, pp. 961-964, Jun. 2014.
- [38] E. Magistretti, et al., "802.11ec: collision avoidance without control messages," in *Proc. ACM MobiCom'12*, pp. 65-76, Aug. 2012.
- [39] S. M. Kay, *Fundamentals of Statistical Signal Processing: Estimation Theory*, Prentice Hall Signal Processing Series, 1993.
- [40] D. V. Sarwate, et al., "Crosscorrelation properties of pseudorandom and related sequences," in *Proc. IEEE*, vol. 68, no. 5, pp. 593-619, May. 1980.
- [41] M. Luby, A. Shokrollahi, M. Watson, and T. Stockhammer, "Raptor forward error correction scheme for object delivery," RFC 5053, Oct. 2007.
- [42] D. Tse and P. Viswanath, *Fundamentals of Wireless Communication*, Cambridge University Press, 2005.
- [43] A. F. Molisch, *Wireless Communications, 2nd Edition*, John Wiley & Sons Ltd., Wiley-IEEE Press, 2011.