Faster-than-Nyquist Signaling through Spatio-temporal Symbol-level Precoding for the Multiuser MISO Downlink Channel

Danilo Spano, *Student Member, IEEE*, Maha Alodeh, *Member, IEEE*, Symeon Chatzinotas, *Senior Member, IEEE,* Björn Ottersten, *Fellow Member, IEEE*

Abstract

This paper deals with the problem of the interference between multiple co-channel transmissions in the downlink of a multi-antenna wireless system. In this framework, symbol-level precoding is a promising technique which is able to constructively exploit the multi-user interference and to transform it into useful power at the receiver side. While previous works on symbol-level precoding were focused on exploiting the multi-user interference, in this paper we extend this concept by jointly handling the interference both in the spatial dimension (multi-user interference) and in the temporal dimension (inter-symbol interference). Accordingly, we propose a novel precoding method, referred to as spatio-temporal symbol-level precoding. In this new precoding paradigm, faster-than-Nyquist signaling can be applied over multi-user MISO systems, and the inter-symbol interference can be tackled at the transmitter side, without additional complexity for the user terminals. While applying faster-than-Nyquist signaling, the proposed optimization strategies perform a sum power minimization with Quality-of-Service constraints. Numerical results are presented in a comparative fashion to show the effectiveness of the proposed techniques, which outperform the state of the art symbol-level precoding schemes in terms of symbol error rate, effective rate, and energy efficiency.

Danilo Spano, Maha Alodeh, Symeon Chatzinotas and Björn Ottersten are with Interdisciplinary Centre for Security Reliability and Trust (SnT) at the University of Luxembourg, Luxembourg. E-mails:{danilo.spano@uni.lu, maha.alodeh@uni.lu, symeon.chatzinotas @uni.lu, and bjorn.ottersten@uni.lu}.

This work is supported by H2020 project SANSA (Shared Access Terrestrial-Satellite Backhaul Network enabled by Smart Antennas), FNR projects PROSAT (on-board PROcessing techniques for high throughput SATellites) and SATSENT (SATellite SEnsor NeTworks for spectrum monitoring), FNR-EPSRC project CI-PHY (Exploiting interference for physical layer security in 5G networks), and FNR-AFR project BroadSat. An initial version of the approach herein presented has been published in *ICASSP 2017*.

I. Introduction

Current research in the context of wireless communications is facing the need to break the existent throughput gridlock, in order to fulfill the ever-increasing demand for interactive services and multimedia content delivery. Since the wireless spectrum is a scarce resource, which is becoming more and more congested, a main challenge is to find novel system architectures and advanced signal processing techniques able to stretch the data rate achievable utilizing the available bandwidth. In this direction, one solution relies on the use of multi-antenna transmitters, which allow aggressive reuse of the frequency spectrum by exploiting the additional degrees of freedom given by the spatial dimension. This architecture allows to serve different co-channel users sharing the same time and frequency resources, through a space division multiple access scheme [1]. However, full frequency reuse schemes necessitate advanced signal processing techniques able to handle the multi-user interference (MUI) arising between the simultaneous transmissions towards the different co-channel users. In this framework, precoding schemes are an effective way to manage the MUI, while guaranteeing specific system performance requirements.

Conventional precoding techniques use knowledge of the channel state information (CSI) to mitigate the MUI, therefore they can be referred to as channel-level precoding. In this class of techniques, the generic scheme relies on the design of a precoding weight matrix (or precoder), which depends only on the CSI. As a consequence, the precoder remains constant for a whole block of symbols whose length is related to the coherence time of the channel. In this framework, several strategies have been considered for the precoder design [2]–[6], including power minimization schemes with Quality-of-Service (QoS) costraints, as well as *max-min fair* approaches. The latter ones aim at increasing the fairness of the system, by maximizing the minimum signal-to-interference-plus-noise ratio (SINR) across the users. In the recent years, a new paradigm has been developed in the context of precoding, known as symbol-level precoding (SLP). This term refers to a class of precoding schemes where the transmitted signals are designed based on the knowledge of both the CSI and the data information, constituted by the symbols to be delivered to the users [7]–[16]. Differently from the conventional channel-level schemes, the aim of symbol-level precoding is not to cancel the interference, but rather to control it so to have a constructive interference effect at each user. In [7] the classification of the interference as constructive or destructive was given, and a selective channel inversion scheme was proposed in order to eliminate the destructive interference. A more advanced symbol-level

precoding scheme was proposed in [8], based on the rotation of the destructive interference so as to transform it into useful power. Similarly to the channel-level case, also in this approach different optimization strategies have been considered in the literature. In [9] the sum power minimization and the *max-min fair* problem were solved for PSK modulations. Extensions of such works include optimization strategies for multi-level modulations [10] and more flexible approaches for exploiting the constructive interference [11]. Furthermore, symbol-level precoding has been considered also in relation to physical layer multicasting [11], and taking into account the imperfect knowledge of the CSI [12] and the channel non-linearities [13]–[16]. The concept of constructive interference has been applied also in the context of code-division multiple-access (CDMA) systems [17]. The reader is referred to [18] for a more detailed review of SLP.

In parallel to the full frequency reuse architectures, another strategy for increasing the wireless spectral efficiency which has attracted considerable interest is the so called faster-than-Nyquist (FTN) signaling [19]–[25]. The key idea of FTN signaling is a reduction of the time spacing between two adjacent pulses (the symbol period) below the one satisfying the Nyquist condition. In other words, in FTN signaling the data rate is increased by accelerating the transmitted pulses in the temporal dimension (time packing), thus introducing controlled inter-symbol interference (ISI) which needs to be handled. The FTN concept was firstly introduced in the mid 70s by Mazo in [19], where it was shown that, given a fixed bandwidth, it is possible to accelerate binary sinc pulses up to a factor of 0.802 with respect to the Nyquist[1] limit without damaging the error rate. Although this result was initially received with skepticism and was not developed for many years, the interest in FTN has grown in the last decade. In [20] it was shown that the FTN concept applies also with squared root raised cosine (SRRC) pulses, which allow a higher acceleration thanks to their excess bandwidth. In [21] the achievable rate regions for FTN broadcast were investigated, considering SRRC pulses. Further, FTN has also been extended in the frequency domain, by squeezing the signals together in frequency just as they were accelerated in time [22]. The FTN principle has also been applied jointly in two dimensions, time and frequency, for multi-carrier systems [23], [24], showing improved achievable rate performance. A review of the work on FTN signaling can be found in [25]. The main problem of FTN signaling is the need to cope with the introduced ISI, which in turn results in complex receivers relying on trellis decoders as well as ad hoc equalization schemes, which are often prohibitive in practical

[1]This means considering a symbol period equal to 0.802 times the one allowed by the Nyquist condition.

applications.

In this paper, we propose a novel transmission technique which allows to merge the strategies discussed above, namely the aggressive frequency reuse relying on precoding and the FTN signaling. Considering a generic multi-user MISO system, the main idea is to extend the concept of symbol-level precoding in order to tackle at the transmitter side not only the interference in the spatial dimension (the MUI), but also the interference in the temporal dimension (the ISI). Such an extension allows FTN signaling in a multi-user MISO framework and, at the same time, solves the problem of complex FTN receivers, as the ISI is completely handled at the transmitter. Hereafter, we will refer to this concept as spatio-temporal SLP. It is important to note that this approach allows to exploit in a constructive fashion the interference both in the temporal and in the spatial dimensions, thus gleaning benefits from both the domains. Overall, the main contribution of this paper can be summarized as follows:

- A novel SLP framework, named spatio-temporal SLP, is introduced and formalized, so as to constructively exploit both the MUI and the ISI;
- This new framework is used to apply FTN signaling over a multi-user MISO system, considering SRRC pulses and coping with the ISI at the transmitter side: in particular, we propose a FTN SLP scheme performing sum power minimization under QoS constraints for a generic multi-level modulation, where the data streams are divided in blocks of symbols and the interference (MUI and ISI) is tackled within each block;
- Further, we propose a more advanced sum power minimization scheme, which tackles not only the interference within each block of symbols but also the inter-block ISI arising between adjacent data blocks, borrowing concepts from dirty paper coding [26] and precoding under interference constraints. This scheme results in lower complexity as long frames can be broken down in shorter symbol blocks and processed separately.

It should be mentioned that an initial spatio-temporal FTN SLP scheme has already been presented by the authors in [27], [28]. However, the scheme proposed therein considers sinc pulses rather than SRRC ones, and it models the signals solely in the symbol domain, without accounting for oversampling. On the other hand, in this work the optimization is performed accounting for the oversampled waveforms produced by the pulse shaping operation. An additional novelty with respect to [27], [28] is the proposed optimization scheme for tackling the inter-block interference arising between subsequent data blocks, which results in improved performance and

lower complexity.

The remainder of this paper is organized as follows. In Section II, the system and signals model for spatio-temporal SLP is delineated. In Section III, the problem of FTN SLP with QoS constraints is formalized and solved, processing a single block of data information. In Section IV, the FTN SLP problem is extended taking into account the inter-block interference between adjacent data blocks. In Section V the proposed approaches are validated through simulation results. Finally, in Section VI conclusions are drawn.

Notation: We use upper-case and lower-case bold-faced letters to denote matrices and vectors, respectively. $(\cdot)^T$ denotes the transpose of $(\cdot)$, while and $(\cdot)^*$ and $(\cdot)^\dagger$ denote the conjugate and the conjugate transpose of $(\cdot)$, respectively. $|\cdot|$ and $\angle(\cdot)$ denote the amplitude and the phase of $(\cdot)$, respectively, while $\mathrm{Re}(\cdot)$ and $\mathrm{Im}(\cdot)$ are the real and imaginary parts of $(\cdot)$. j is used to denote the imaginary unit, while $\|\cdot\|$ and $\|\cdot\|_F$ represent the Euclidean norm of a vector and the Frobenius norm of a matrix, respectively. $\otimes$ denotes the Kronecker product, $\circ$ is used to denote the element-wise Hadamard operations, while $\mathbb{E}_i[\cdot]$ represents the statistical expectation with respect to the index i. Finally, $\mathrm{vec}(\cdot)$ denotes the vectorization of a matrix, while $\mathbf{1}_{a\times b}$ and $\boldsymbol{I}_a$ denote the matrix of all ones of size $a \times b$ and the identity matrix of size $a \times a$, respectively.

II. System Model

Let us consider a single-cell multiple-antenna downlink scenario, where a base station delivers K independent data streams to K single-antenna user terminals through N transmit antennas, with $N \geq K$. Each data stream is divided in blocks of S symbols, and the channel is assumed to be quasi-static flat fading. Considering a data block, we can define the data information matrix $\boldsymbol{S} = [\boldsymbol{s}_1^T \dots \boldsymbol{s}_K^T]^T \in \mathbb{C}^{K\times S}$ which aggregates the symbol streams to be conveyed to the different users, taken from a constellation having unit average power. Similarly, we aggregate in the matrix $\boldsymbol{D} = [\boldsymbol{d}_1^T \dots \boldsymbol{d}_N^T]^T \in \mathbb{C}^{N\times S}$ the precoded symbol streams which feed the transmit filters. In fact, each symbol stream has to undergo pulse shaping before the actual transmission. The pulse shaping operation is performed using a unit-power symmetric pulse waveform $\alpha(t)$, having duration $2\eta T$, with T being the symbol period[2]. This implies that $\alpha(t) = 0$ for $|t| > \eta T$. Moving to a discrete time representation, and considering an oversampling factor n_s, the pulse

[2]For infinite pulses, η is defined by the time (in symbol periods) at which the pulse amplitude decades below a sufficiently low level so that it can be considered negligible.

waveform can be represented through its samples spaced by $t_s = \frac{T}{n_s}$, i.e., $\alpha[mt_s]$, with the index m such that $|m| \leq \eta n_s$ (accounting for the pulse duration).

With the introduced formalism, it is possible to write the expression of the discrete samples at the output of the pulse shaping filter of the generic n-th antenna, as follows:

$$x_n[(m-1)t_s] = \sum_{i=1}^{S} d_{ni}\alpha[(m-1)t_s - (i-1)T], \quad m = 1, \ldots, n_s S, \tag{1}$$

where d_{ni} is the i-th element of the symbol vector $\boldsymbol{d}_n$, which in turn is the n-th row of $\boldsymbol{D}$. Such relation can be rewritten in a compact matrix form as:

$$\boldsymbol{x}_n = \boldsymbol{d}_n \boldsymbol{A}_{\text{TX}}, \tag{2}$$

where $\boldsymbol{x}_n \in \mathbb{C}^{1 \times n_s S}$ represents the output data stream (in the oversampled domain) from the n-th antenna and $\boldsymbol{A}_{\text{TX}} \in \mathbb{R}^{S \times n_s S}$ is a block Toeplitz matrix modeling the pulse shaping operation, with its (i, m)-th element being:

$$[\boldsymbol{A}_{\text{TX}}]_{(i,m)} = \alpha[(m-1)t_s - (i-1)T]. \tag{3}$$

Further, we can aggregate the output signals from all the antennas in a matrix $\boldsymbol{X} = [\boldsymbol{x}_1^T \ldots \boldsymbol{x}_N^T]^T \in \mathbb{C}^{N \times n_s S}$, which allows us to model the pulse shaping operation for all the N antennas simply as $\boldsymbol{X} = \boldsymbol{D}\boldsymbol{A}_{\text{TX}}$.

According to the well-known multiuser MISO channel model, the received symbols at the users can be written in matrix form as:

$$\tilde{\boldsymbol{Y}} = \boldsymbol{H}\boldsymbol{X} + \tilde{\boldsymbol{Z}}$$

where the matrix $\tilde{\boldsymbol{Y}} = [\tilde{\boldsymbol{y}}_1^T \ldots \tilde{\boldsymbol{y}}_K^T]^T \in \mathbb{C}^{K \times n_s S}$ represents the received samples at the K users, $\boldsymbol{H} = [\boldsymbol{h}_1^T \ldots \boldsymbol{h}_K^T]^T \in \mathbb{C}^{K \times N}$ is the quasi-static flat fading channel[3] matrix modeling the MUI among the different data streams, and $\tilde{\boldsymbol{Z}} = [\tilde{\boldsymbol{z}}_1^T \ldots \tilde{\boldsymbol{z}}_K^T]^T \in \mathbb{C}^{K \times n_s S}$ models the additive white Gaussian noise (AWGN). In order to obtain the received signals at the users in the symbol domain, the matched filtering and downsampling operation needs to be modeled. We apply this

[3]In principle, it is possible to extend the model accounting also for frequency selective fading channels, by representing the resulting channel induced ISI. Nevertheless, in order not to further complicate the system model, in this contribution we focus on flat fading channels, while the extension to the general case will be considered in the future work.

operation again in a matrix form, using the block Toeplitz matrix $\boldsymbol{A}_{\text{RX}} \in \mathbb{R}^{n_s S \times S}$, which can be defined in the same fashion of (3). Overall, grouping the received symbols at the K users in a matrix $\boldsymbol{Y} = [\boldsymbol{y}_1^T \dots \boldsymbol{y}_K^T]^T \in \mathbb{C}^{K \times S}$, we can write the global communication model as:

$$\boldsymbol{Y} = \tilde{\boldsymbol{Y}} \boldsymbol{A}_{\text{RX}} = \boldsymbol{H} \boldsymbol{X} \boldsymbol{A}_{\text{RX}} + \tilde{\boldsymbol{Z}} \boldsymbol{A}_{\text{RX}} = \boldsymbol{H} \boldsymbol{D} \boldsymbol{A} + \boldsymbol{Z}, \tag{4}$$

where $\boldsymbol{A} = \boldsymbol{A}_{\text{TX}} \boldsymbol{A}_{\text{RX}} \in \mathbb{R}^{S \times S}$ is a matrix representing the convolution of the filters at the transmitter and at the receiver, while $\boldsymbol{Z} = \tilde{\boldsymbol{Z}} \boldsymbol{A}_{\text{RX}} \in \mathbb{C}^{K \times S}$ is the noise in the symbol domain, having power σ_z^2. If we denote by $\beta(t)$ the impulse response of the overall filter composed by the convolution of the pulse shaping and the matched filtering, it can be seen that $\boldsymbol{A}$ is a symmetric Toeplitz matrix whose first row is $\boldsymbol{a} = [\beta[0] \beta[T] \dots \beta[2\eta T] 0 \dots 0]$[4]. It should be stressed that the model in (4) takes into account the interference both in the spatial dimension (the MUI), through the spatial channel matrix $\boldsymbol{H}$, and in the temporal dimension (the ISI), through the temporal channel matrix $\boldsymbol{A}$. The complete system model is represented in the block scheme of Fig. 1, where it is clear how the symbol matrix $\boldsymbol{D}$ is obtained as output of a spatio-temporal precoding module, which takes as input the CSI, i.e. an estimate of $\boldsymbol{H}$, the filters matrices $\boldsymbol{A}_{\text{TX}}$ and $\boldsymbol{A}_{\text{RX}}$ and the data information matrix $\boldsymbol{S}$. Regarding the channel estimation, the receivers estimate their respective channel vectors $\boldsymbol{h}_1, \dots, \boldsymbol{h}_K$ by exploiting a training sequence (pilot symbols) included in the framing structure of the communication system[5], and the resulting CSI is transmitted back to the base station through a feedback channel, in order to be available for the precoding operation. Alternatively, in systems using time division duplexing (TDD) the channel can be directly estimated at the transmitter based on the uplink-downlink reciprocity principle. The reader is referred to [3] for a more detailed overview of the different channel estimation strategies. In the remainder of the paper we assume a perfect knowledge of the CSI, as the robustness to CSI errors falls out of the scope of this contribution. Nonetheless, the effect of an imperfect CSI on the performance of the proposed schemes is assessed in the numerical results section.

The aim of the symbol-level precoding scheme is to optimize the matrix $\boldsymbol{D}$, namely the precoded symbol streams feeding the transmit filters, by constructively exploiting the interference

[4]To ease the notation, we are here assuming that $S > 2\eta$.

[5]In the presented system model, we have not explicitly included pilots because this would further complicate the notation without changing the main conclusions of the work.

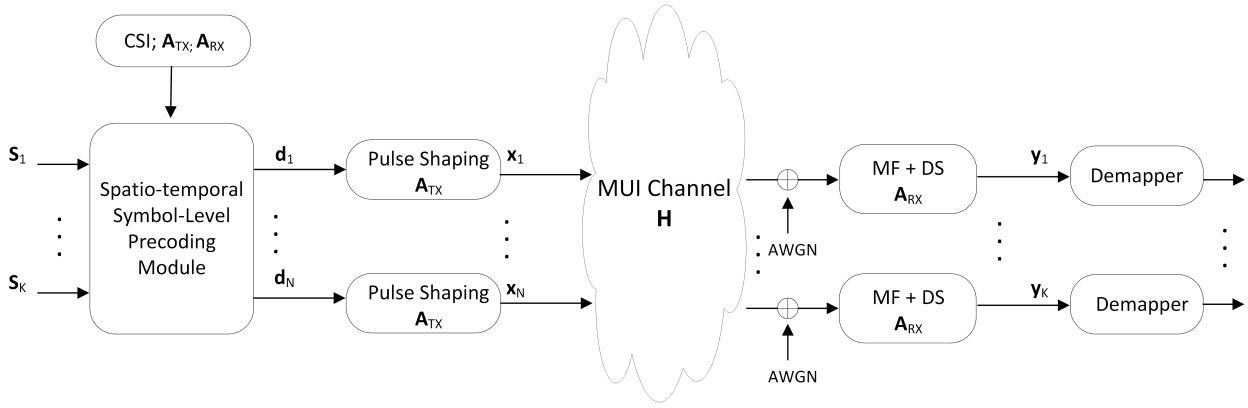

Figure 1: Block scheme of the considered system model relying on spatio-temporal symbol-level precoding.

in the spatial and in the temporal dimension. While this optimization procedure will be explained in detail in the following section, it is now convenient to further manipulate the model of (4) by vectorizing the introduced signal matrices over the temporal dimension (rows first). Accordingly, we model the data information streams through the vector $\boldsymbol{s} = \text{vec}(\boldsymbol{S}^T) = [\boldsymbol{s}_1 \dots \boldsymbol{s}_K]^T \in \mathbb{C}^{KS\times 1}$, the designed symbol streams through $\boldsymbol{d} = \text{vec}(\boldsymbol{D}^T) = [\boldsymbol{d}_1 \dots \boldsymbol{d}_N]^T \in \mathbb{C}^{NS\times 1}$, the transmitted signals through $\boldsymbol{x} = \text{vec}(\boldsymbol{X}^T) = [\boldsymbol{x}_1 \dots \boldsymbol{x}_N]^T \in \mathbb{C}^{Nn_sS\times 1}$, the noise through $\boldsymbol{z} = \text{vec}(\boldsymbol{Z}^T) = [\boldsymbol{z}_1 \dots \boldsymbol{z}_K]^T \in \mathbb{C}^{KS\times 1}$, and the received symbols through $\boldsymbol{y} = \text{vec}(\boldsymbol{Y}^T) = [\boldsymbol{y}_1 \dots \boldsymbol{y}_K]^T \in \mathbb{C}^{KS\times 1}$. It is straightforward to check that the relation between $\boldsymbol{d}$ and $\boldsymbol{x}$ can be written as:

$$\boldsymbol{x} = (\boldsymbol{I}_N \otimes \boldsymbol{A}_{\text{TX}}^T)\boldsymbol{d}. \tag{5}$$

Further, by introducing the matrix $\hat{\boldsymbol{X}} = \boldsymbol{X}\boldsymbol{A}_{\text{RX}} = \boldsymbol{D}\boldsymbol{A} \in \mathbb{C}^{N\times S}$ and its vectorized version $\hat{\boldsymbol{x}} = \text{vec}(\hat{\boldsymbol{X}}^T) \in \mathbb{C}^{NS\times 1}$, and by accounting for (4), it is easy to check that $\hat{\boldsymbol{x}} = (\boldsymbol{I}_N \otimes \boldsymbol{A}^T)\boldsymbol{d}$ and that $\boldsymbol{y} = (\boldsymbol{H} \otimes \boldsymbol{I}_S)\hat{\boldsymbol{x}} + \boldsymbol{z}$. Finally, using the mixed-product property of the Kronecker product we have $(\boldsymbol{H} \otimes \boldsymbol{I}_S)(\boldsymbol{I}_N \otimes \boldsymbol{A}^T) = (\boldsymbol{H} \otimes \boldsymbol{A}^T)$, therefore we can write the global communication model as follows:

$$\boldsymbol{y} = (\boldsymbol{H} \otimes \boldsymbol{A}^T)\boldsymbol{d} + \boldsymbol{z} = \boldsymbol{G}\boldsymbol{d} + \boldsymbol{z}. \tag{6}$$

The matrix $\boldsymbol{G} = \boldsymbol{H} \otimes \boldsymbol{A}^T \in \mathbb{C}^{KS\times NS}$ incorporates both the spatial channel matrix $\boldsymbol{H}$ and the temporal one $\boldsymbol{A}$, thus we will refer to it as spatio-temporal channel matrix. Ultimately, the model of (6) allows us to represent in a very simple way both the MUI and the ISI of the system. In

the next section, the optimization problem modeling the proposed spatio-temporal SLP scheme will be detailed.

A. *Faster-than-Nyquist*

As already mentioned, FTN signaling manages to pack more information in the time domain by reducing the symbol period T below the minimum allowed by the Nyquist criterion, thus introducing controlled ISI. This is graphically shown in Fig. 2 for the case of a SRRC pulse. In the system model definition, we have not made any assumptions on the symbol-rate so far. It can be easily seen that if we do not apply FTN, then the Toeplitz matrix $\boldsymbol{A}$ simply reduces to a scaled identity, i.e., $\boldsymbol{A} = \beta[0]\boldsymbol{I}_S$. In this case there is no ISI and the model in (4) boils down to the classic multiuser MISO case.

Now, let us assume that we apply a signaling acceleration factor $\tau \in [0, 1]$, so that the effective symbol period is $T = \tau T_{\text{ny}}$, with T_{ny} indicating the minimum symbol period allowed by the Nyquist criterion. It can be easily seen that the lower is the acceleration factor τ (i.e., the more we accelerate the transmissions) the larger is the number of non-zero values in the matrix $\boldsymbol{A}$, thus the higher is the ISI level in the system. This can be easily explained by the fact that more pulses are packed in the time domain. Furthermore, as the total duration of the individual pulse remains constant, the value of η increases as τ decreases[6]. The schemes proposed in this paper apply regardless of the chosen pulse. In the numerical results section, SRRC pulses will be considered, as they are the most used in practical applications. It should be noted that no complex equalization or decoding is needed at the receivers, as they are oblivious to the FTN operation. The receivers have to be informed only about the baud-rate of the communication for sampling purposes.

III. Faster-than-Nyquist SLP for Sum Power Minimization with QoS Constraints

In this section a novel SLP scheme accounting for FTN signaling is presented, which exploits in a constructive fashion [9] the interference both in the spatial and in the temporal domain. The novelty with respect to the previous SLP solutions [7]–[15] lies in the ability of the new

[6]In the remainder of the paper we will refer solely to the acceleration factor τ to represent the FTN operation, as η is directly dependent on it for a given pulse.

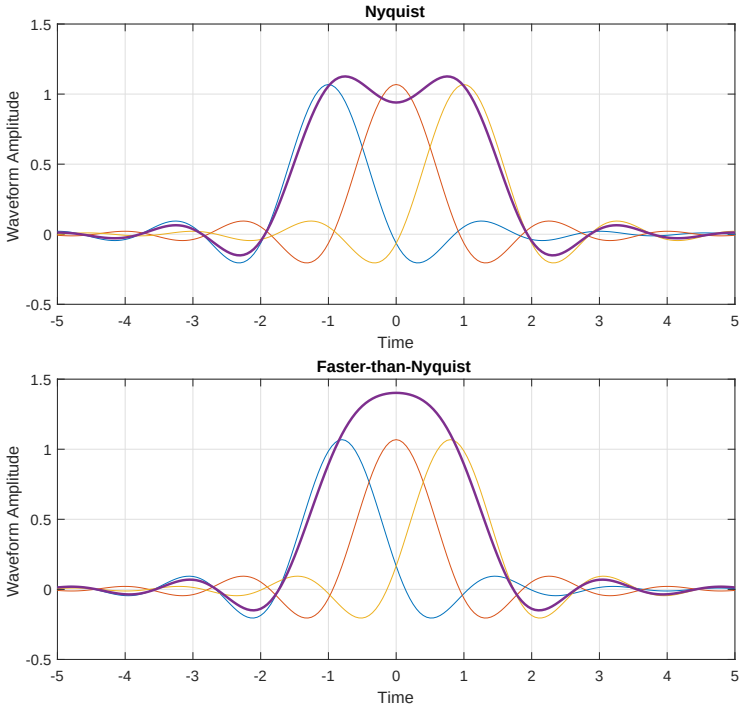

Figure 2: Nyquist vs. Faster-than-Nyquist for a SRRC pulse, in the time domain.

scheme to model and handle the ISI, together with the MUI, and in the consequent capacity of exploting the potential of FTN signaling in the context of multiuser MISO systems. More specifically, in this work we propose a sum power minimization scheme under QoS constraints. As common in the precoding literature, the QoS constraints are expressed in terms of target SINR, which represents a predefined per-user SINR threshold to be guaranteed at the receivers' side. The target SINR, which is an input parameter to the precoding scheme, has to be selected in order to ensure a sufficiently good performance (e.g. in terms of symbol error rate or bit error rate) for the considered application, and strongly depends on the adopted modulation as well as on the use of forward error correction (FEC) schemes. Since we aim at minimizing the total transmit power P_{tot}, it is convenient to explicit its expression before formalizing the proposed scheme, as follows:

$$P_{\text{tot}} = \frac{1}{n_s S} \sum_{m=1}^{nsS} \sum_{n=1}^{N} |[\boldsymbol{X}]_{(n,m)}|^2 = \frac{1}{n_s S} \|\boldsymbol{X}\|_F^2 = \frac{1}{n_s S} \|\boldsymbol{x}\|^2. \tag{7}$$

Accordingly, taking also into account that $\boldsymbol{X} = \boldsymbol{D}\boldsymbol{A}_{\text{TX}}$, a general formulation of the optimization problem for the proposed scheme, using the matrix notation introduced in Section II, is the

following:

$$\begin{aligned} \boldsymbol{D}(\boldsymbol{S},\boldsymbol{H},\boldsymbol{A}_{\text{TX}},\boldsymbol{A}_{\text{RX}},\boldsymbol{\gamma}) = \arg\min_{\boldsymbol{D}} \quad & \|\boldsymbol{D}\boldsymbol{A}_{\text{TX}}\|_F^2 \\ \text{s.t.} \quad & \boldsymbol{H}\boldsymbol{D}\boldsymbol{A} \triangleright \sigma_z \boldsymbol{Q} \circ \boldsymbol{S}, \end{aligned} \tag{8}$$

where the quantities in brackets are given as input to the optimization problem. Among them, the vector $\boldsymbol{\gamma} = [\gamma_1 \dots \gamma_K]^T \in \mathbb{R}^{K\times 1}$ represents the target SINR for all the users and appears in the optimization problem through the matrix $\boldsymbol{Q} = \sqrt{\boldsymbol{\gamma}} \otimes \mathbf{1}_{1\times S}$. The operator $\triangleright$ used in the constraint, which applies element-wise, imposes that the received symbols at each user (represented respectively by the elements of $\boldsymbol{HDA}$) lie in their correct detection regions. Further, the correct detection regions, which clearly depend on the data information matrix $\boldsymbol{S}$, are scaled accounting for the target SINR $\boldsymbol{\gamma}$. The introduced constraint allows the exploitation of the constructive interference effect, in the same fashion as [10]. In order to express it in a more explicit form it is necessary to refer to a specific modulation scheme, as we will do later in this section. The problem (8) can be rewritten resorting to the vectorized formalism and the spatio-temporal channel matrix introduced in (5)-(6), as follows:

$$\begin{aligned} \boldsymbol{d}(\boldsymbol{s},\boldsymbol{H},\boldsymbol{A}_{\text{TX}},\boldsymbol{A}_{\text{RX}},\boldsymbol{\gamma}) = \arg\min_{\boldsymbol{d}} \quad & \|\boldsymbol{x}\|^2 \\ \text{s.t.} \quad & \boldsymbol{G}_k\boldsymbol{d} \triangleright \sigma_z\sqrt{\gamma_k}\boldsymbol{s}_k^T,\ k=1,\dots,K, \end{aligned} \tag{9}$$

where $\boldsymbol{G}_k = [\boldsymbol{g}_k[1]^T \dots \boldsymbol{g}_k[S]^T]^T \in \mathbb{C}^{S\times NS}$ is a submatrix of $\boldsymbol{G}$ denoting the spatio-temporal channel matrix for the k-th user, and $\boldsymbol{s}_k \in \mathbb{C}^{1\times S}$ represents the data information related to the k-th user. The optimization problem can be further manipulated by using (5), and by specifying the constraints for each symbol slot, as follows:

$$\begin{aligned} \boldsymbol{d}(\boldsymbol{s},\boldsymbol{H},\boldsymbol{A}_{\text{TX}},\boldsymbol{A}_{\text{RX}},\boldsymbol{\gamma}) = \arg\min_{\boldsymbol{d}} \quad & \|(\boldsymbol{I}_N \otimes \boldsymbol{A}_{\text{TX}}^T)\boldsymbol{d}\|^2 \\ \text{s.t.} \quad & \boldsymbol{g}_k[i]\boldsymbol{d} \triangleright \sigma_z\sqrt{\gamma_k}\boldsymbol{s}_k[i],\ k=1,\dots,K,\ i=1,\dots,S. \end{aligned} \tag{10}$$

It should be stressed that the quantity $\boldsymbol{g}_k[i]\boldsymbol{d}$ represents the received symbol at the k-th user terminal in the i-th symbol slot, and that the imposed constraint forces it to lie in the correct detection region of the corresponding data information symbol $\boldsymbol{s}_k[i]$. It is also worth highlighting that the SINR corresponding to the k-th user is given by $\mathbb{E}_i\left[\frac{|\boldsymbol{g}_k[i]\boldsymbol{d}|^2}{\sigma_z^2}\right]$, therefore the scaling factor $\sigma_z\sqrt{\gamma_k}$ introduced in the constraint allows to guarantee the target SINR γ_k. It is important now to take a further step in the formalization of the power minimization problem at hand,

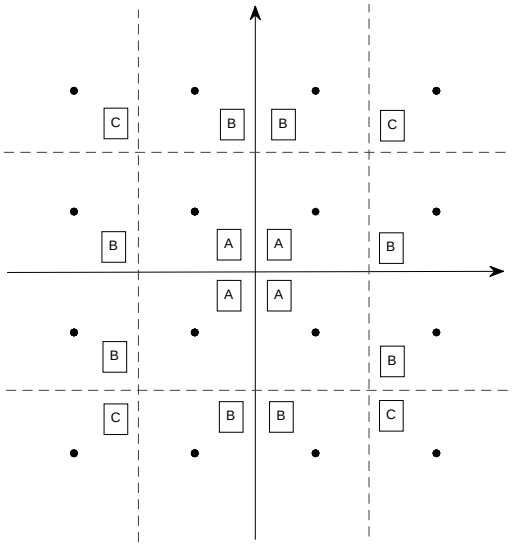

Figure 3: Classification of the constellation points into inner (A), outer (B) and outermost (C), for a 16-QAM modulation scheme.

by expressing the introduced constraints in an explicit form. To this aim, we focus on a QAM modulation scheme for the data information. In this case the optimization problem can be written by decomposing the constraints along the in-phase and quadrature components of the symbols, as follows:

$$\begin{aligned} \boldsymbol{d}(\boldsymbol{s}, \boldsymbol{H}, \boldsymbol{A}_{\text{TX}}, \boldsymbol{A}_{\text{RX}}, \boldsymbol{\gamma}) = \arg\min_{\boldsymbol{d}} \quad & \|(\boldsymbol{I}_N \otimes \boldsymbol{A}_{\text{TX}}^T)\boldsymbol{d}\|^2 \\ \text{s.t.} \quad \mathcal{C}1: \quad & \text{Re}(\boldsymbol{g}_k[i]\boldsymbol{d}) \gtreqless \sigma_z\sqrt{\gamma_k}\,\text{Re}(\boldsymbol{s}_k[i]),\ k=1,\ldots,K,\ i=1,\ldots,S, \\ \mathcal{C}2: \quad & \text{Im}(\boldsymbol{g}_k[i]\boldsymbol{d}) \gtreqless \sigma_z\sqrt{\gamma_k}\,\text{Im}(\boldsymbol{s}_k[i]),\ k=1,\ldots,K,\ i=1,\ldots,S, \end{aligned} \tag{11}$$

where the notation $\gtreqless$ denotes a generalized inequality, which shall be read as $>$,$<$ or $=$ depending on the position of the data $\boldsymbol{s}_k[i]$ within the QAM constellation and, accordingly, on its detection region. A detailed formulation of the constraints $\mathcal{C}1$, $\mathcal{C}2$ is explained hereafter:

- For the inner constellation symbols, which are labeled by A in the 16-QAM example of Fig. 3, the constraints $\mathcal{C}1$, $\mathcal{C}2$ should guarantee that the received signal achieves the exact constellation point. Hence, the constraints are equality constraints, as follows:

$$\begin{aligned} \mathcal{C}1: \quad & \mathrm{Re}(\boldsymbol{g}_k[i]\boldsymbol{d}) = \sigma_z\sqrt{\gamma_k}\,\mathrm{Re}(\boldsymbol{s}_k[i]), \\ \mathcal{C}2: \quad & \mathrm{Im}(\boldsymbol{g}_k[i]\boldsymbol{d}) = \sigma_z\sqrt{\gamma_k}\,\mathrm{Im}(\boldsymbol{s}_k[i]). \end{aligned} \tag{12}$$

- For the outer constellation symbols, which are labeled by B in the 16-QAM example of Fig. 3, the constraints $\mathcal{C}1$, $\mathcal{C}2$ guaranteeing the correct detections and exploiting the constructive interference are:

$$\begin{aligned} \mathcal{C}1: \quad & \mathrm{Re}(\boldsymbol{g}_k[i]\boldsymbol{d}) \geq \sigma_z\sqrt{\gamma_k}\,\mathrm{Re}(\boldsymbol{s}_k[i]), \text{ if } \mathrm{Re}\,\boldsymbol{s}_k[i] > 0, \\ & \mathrm{Re}(\boldsymbol{g}_k[i]\boldsymbol{d}) \leq \sigma_z\sqrt{\gamma_k}\,\mathrm{Re}(\boldsymbol{s}_k[i]), \text{ if } \mathrm{Re}\,\boldsymbol{s}_k[i] < 0, \\ \mathcal{C}2: \quad & \mathrm{Im}(\boldsymbol{g}_k[i]\boldsymbol{d}) = \sigma_z\sqrt{\gamma_k}\,\mathrm{Im}(\boldsymbol{s}_k[i]), \end{aligned} \tag{13}$$

  when $\boldsymbol{s}_k[i]$ lies on the left/right side of the constellation, and:

$$\begin{aligned} \mathcal{C}1: \quad & \mathrm{Re}(\boldsymbol{g}_k[i]\boldsymbol{d}) = \sigma_z\sqrt{\gamma_k}\,\mathrm{Re}(\boldsymbol{s}_k[i]), \\ \mathcal{C}2: \quad & \mathrm{Im}(\boldsymbol{g}_k[i]\boldsymbol{d}) \geq \sigma_z\sqrt{\gamma_k}\,\mathrm{Im}(\boldsymbol{s}_k[i]), \text{ if } \mathrm{Im}\,\boldsymbol{s}_k[i] > 0, \\ & \mathrm{Im}(\boldsymbol{g}_k[i]\boldsymbol{d}) \leq \sigma_z\sqrt{\gamma_k}\,\mathrm{Im}(\boldsymbol{s}_k[i]), \text{ if } \mathrm{Im}\,\boldsymbol{s}_k[i] < 0, \end{aligned} \tag{14}$$

  when $\boldsymbol{s}_k[i]$ lies on the upper/lower side of the constellation.

- Finally, for the outermost symbols lying on the edges of the QAM constellation, which are labeled by C in the 16-QAM example of Fig. 3, the constraints $\mathcal{C}1$, $\mathcal{C}2$ can exploit the constructive interference effect by imposing inequalities along both the in-phase and the quadrature components of the signals, as follows:

$$\begin{aligned} \mathcal{C}1: \quad & \mathrm{Re}(\boldsymbol{g}_k[i]\boldsymbol{d}) \geq \sigma_z\sqrt{\gamma_k}\,\mathrm{Re}(\boldsymbol{s}_k[i]), \text{ if } \mathrm{Re}\,\boldsymbol{s}_k[i] > 0, \\ & \mathrm{Re}(\boldsymbol{g}_k[i]\boldsymbol{d}) \leq \sigma_z\sqrt{\gamma_k}\,\mathrm{Re}(\boldsymbol{s}_k[i]), \text{ if } \mathrm{Re}\,\boldsymbol{s}_k[i] < 0, \\ \mathcal{C}2: \quad & \mathrm{Im}(\boldsymbol{g}_k[i]\boldsymbol{d}) \geq \sigma_z\sqrt{\gamma_k}\,\mathrm{Im}(\boldsymbol{s}_k[i]), \text{ if } \mathrm{Im}\,\boldsymbol{s}_k[i] > 0, \\ & \mathrm{Im}(\boldsymbol{g}_k[i]\boldsymbol{d}) \leq \sigma_z\sqrt{\gamma_k}\,\mathrm{Im}(\boldsymbol{s}_k[i]), \text{ if } \mathrm{Im}\,\boldsymbol{s}_k[i] < 0. \end{aligned} \tag{15}$$

The final optimization problem in (11) presents a convex quadratic objective function and affine constraints, therefore it is convex and can be solved resorting to the standard convex optimization tools [29]. As a final remark, it should be mentioned that, although herein we solely focused on QAM constellations for the sake of brevity, the optimization problem (10) can be straightforwardly expressed for different constellations, by tailoring the constraints to the different detection regions. In particular, if the data information symbols belong to a PSK or to an

APSK constellation, it is convenient to express the constraints in (10) focusing on the amplitude and the phase of the symbols, rather than their in-phase and quadrature components, as shown in [13], [14], [16]. Although in this case the optimization problem appears more constrained with respect to (11), due to the different detection regions, it is still possible to exploit the constructive interference effect of SLP.

IV. Sequential Faster-than-Nyquist SLP: processing subsequent blocks

So far we have considered a single data block of S symbols per stream, and we have devised an approach to constructively handle the interference both in the spatial and in the temporal dimension within the block. However, it is clear that the optimization problem (11) cannot handle any arbitrary block length S, as the dimension of the involved optimization variables, as well as the number of constraints, linearly grow with S[7]. This implies that in a practical system the scheme needs to process subsequently different information blocks of a manageable length S. If the framing structure of the system allows to neglect the mutual interference between adjacent blocks (for instance, because subsequent information blocks are separated by a sufficient number of non-precoded signaling symbols, such as headers and pilots), then the problem formulation in (11) still applies. Nevertheless, in general we also need to account for the ISI arising between subsequent blocks, i.e., the inter-block ISI. In this section we extend the proposed approach so as to cope with the problem of inter-block interference, thus moving closer to the practical application of the discussed FTN SLP scheme.

As a first step towards extending the previous scheme, we need to model the ISI between two adjacent blocks. In particular, if we denote the current block under processing by an index l, we need to model the residual ISI coming from the previous $(l-1)$-th block, as well as the ISI that the current l-th block is causing to the $(l-1)$-th one[8]. This inter-block interference can be taken into account by extending the communication model in (6) as follows:

$$\begin{bmatrix} \boldsymbol{y}_{l-1} \\ \boldsymbol{y}_{l} \end{bmatrix} = \begin{bmatrix} \boldsymbol{G} & \boldsymbol{G}_U \\ \boldsymbol{G}_P & \boldsymbol{G} \end{bmatrix} \begin{bmatrix} \boldsymbol{d}_{l-1} \\ \boldsymbol{d}_{l} \end{bmatrix} + \begin{bmatrix} \boldsymbol{z}_{l-1} \\ \boldsymbol{z}_{l} \end{bmatrix}, \tag{16}$$

[7]A numerical evaluation of the complexity of (11), as a function of S, is given in Section V.

[8]In principle, it could be possible to consider also the $(l-2)$-th block or even previous blocks in the model. However, in most practical scenarios the residual ISI coming from such blocks would be negligible. Thus, also in order to keep the complexity of the scheme to a manageable level, we choose to consider only two adjacent blocks.

where $\boldsymbol{G}_P = \boldsymbol{H} \otimes \boldsymbol{A}_P^T \in \mathbb{C}^{KS \times NS}$ and $\boldsymbol{G}_U = \boldsymbol{H} \otimes \boldsymbol{A}_U^T \in \mathbb{C}^{KS \times NS}$ respectively, and the matrices $\boldsymbol{A}_P \in \mathbb{R}^{S \times S}$ and $\boldsymbol{A}_U \in \mathbb{R}^{S \times S}$ model the ISI coming from the previous block and the ISI caused to the previous block, respectively. It is straightforward to observe that $\boldsymbol{A}_P$ is a Toeplitz matrix whose first column is given by $\boldsymbol{a}_{Pc} = [0 \dots 0 \ \beta[-2\eta T] \ \dots \beta[-T]]^T$ and whose first row is a zero-entries row vector. Similarly, $\boldsymbol{A}_U$ is a Toeplitz matrix whose last column is given by $\boldsymbol{a}_{Uc} = [\beta[T] \dots \beta[2\eta T] \ 0 \dots 0]^T$ and whose last row is a zero-entries row vector.

Now, assuming that blocks are serially processed, the ISI caused by the $(l-1)$-th block to the l-th one will be represented by the vector $\boldsymbol{v} = \boldsymbol{G}_P \boldsymbol{d}_{l-1} \in \mathbb{C}^{KS \times 1}$, which is known and can be used in the optimization scheme designing $\boldsymbol{d}_l$. It should be noted that this concept is similar to the dirty paper coding principle [26], where a known state is taken into account while designing the transmit signal. Analogously to the other introduced vectorized quantities, $\boldsymbol{v}$ can also be decomposed by indexing the components related to the different users, i.e., $\boldsymbol{v} = [\boldsymbol{v}_1 \dots \boldsymbol{v}_K]^T$. Besides accounting for the ISI coming from the previous block, we should also try to minimize the ISI that the l-th block is causing to the previous one, which is represented by the vector $\boldsymbol{G}_U \boldsymbol{d}_l$. This is achieved by setting interference constraints towards the previous block, following a strategy which resembles cognitive precoding schemes [30]–[32]. With these considerations, we can finally formalize the following sequential FTN SLP optimization problem, performing a sum power minimization with QoS constraints:

$$\begin{aligned} \boldsymbol{d}_l(\boldsymbol{s}, \boldsymbol{H}, \boldsymbol{A}_{\text{TX}}, \boldsymbol{A}_{\text{RX}}, \boldsymbol{\gamma}, \boldsymbol{v}, \epsilon) = \arg\min_{\boldsymbol{d}_l} \quad & \|\boldsymbol{x}_l\|^2 \\ \text{s.t.} \quad \mathcal{C}1: \quad & \boldsymbol{G}_k \boldsymbol{d}_l + \boldsymbol{v}_k^T \triangleright \sigma_z \sqrt{\gamma_k} \boldsymbol{s}_k^T, \ k = 1, \dots, K, \\ \mathcal{C}2: \quad & |\boldsymbol{G}_U \boldsymbol{d}_l|^{\circ 2} \leq \epsilon \mathbf{1}_{KS \times 1}, \end{aligned} \tag{17}$$

where the Hadamard notation $\circ$ in $\mathcal{C}2$ is used to indicate that the squaring operation applies element-wise (moreover, the inequality has also to be considered element-wise). The novelty with respect to the problem in (9) is in the fact that the constraints $\mathcal{C}1$ is now also accounting the ISI coming from the previous block, so as to guarantee constructive interference at each user. Further, the constraint $\mathcal{C}2$ imposes a maximum level of ISI that the l-th block is causing to the previous one, through a predefined threshold ϵ. Although the best in terms of ISI reduction would be to impose a zero-forcing condition in $\mathcal{C}2$, i.e., fixing $\epsilon = 0$, a numerical analysis has shown how such choice would make the problem unfeasible for most scenarios due to the lack of degrees of freedom. In the numerical results section different values for ϵ will be considered

and discussed.

The constraint $\mathcal{C}2$ can be decomposed by considering for each user the relative submatrix of the matrix $\boldsymbol{G}_U$, similarly to what was done in (9) for $\boldsymbol{G}$. Accordingly, $\boldsymbol{G}_{Uk} = [\boldsymbol{g}_{Uk}[1]^T \dots \boldsymbol{g}_{Uk}[S]^T]^T \in \mathbb{C}^{S \times NS}$ is a submatrix of $\boldsymbol{G}_U$ related to the k-th user, hence the problem (17) can be rewritten as:

$$\begin{aligned} \boldsymbol{d}_l(\boldsymbol{s}, \boldsymbol{H}, \boldsymbol{A}_{\text{TX}}, \boldsymbol{A}_{\text{RX}}, \boldsymbol{\gamma}, \boldsymbol{v}, \epsilon) = \arg\min_{\boldsymbol{d}_l} \quad & \|\boldsymbol{x}_l\|^2 \\ \text{s.t.} \quad \mathcal{C}1: \quad & \boldsymbol{G}_k \boldsymbol{d}_l + \boldsymbol{v}_k^T \triangleright \sigma_z \sqrt{\gamma_k} \boldsymbol{s}_k^T, \ k = 1, \dots, K, \\ \mathcal{C}2: \quad & |\boldsymbol{G}_{Uk} \boldsymbol{d}_l|^{\circ 2} \leq \epsilon \mathbf{1}_{S \times 1}, \ k = 1, \dots, K, \end{aligned} \tag{18}$$

and, expressing the constraints for each symbol slot, in the same fashion of (10), as:

$$\begin{aligned} \boldsymbol{d}_l(\boldsymbol{s}, \boldsymbol{H}, \boldsymbol{A}_{\text{TX}}, \boldsymbol{A}_{\text{RX}}, \boldsymbol{\gamma}, \boldsymbol{v}, \epsilon) = \arg\min_{\boldsymbol{d}_l} \quad & \|(\boldsymbol{I}_N \otimes \boldsymbol{A}_{\text{TX}}^T) \boldsymbol{d}_l\|^2 \\ \text{s.t.} \quad \mathcal{C}1: \quad & \boldsymbol{g}_k[i] \boldsymbol{d}_l + \boldsymbol{v}_k[i] \triangleright \sigma_z \sqrt{\gamma_k} \boldsymbol{s}_k[i], \ k = 1, \dots, K, \ i = 1, \dots, S, \\ \mathcal{C}2: \quad & |\boldsymbol{g}_{Uk}[i] \boldsymbol{d}_l|^2 \leq \epsilon, \ k = 1, \dots, K, \ i = 1, \dots, S. \end{aligned} \tag{19}$$

Interestingly, it can be noticed how in this sequential scheme the SINR corresponding to the k-th user is given by $\mathbb{E}_i \left[\frac{|\boldsymbol{g}_k[i] \boldsymbol{d}_l + \boldsymbol{v}_k[i]|^2}{\sigma_z^2} \right]$, since the ISI coming from the previous block, modeled by $\boldsymbol{v}_k$, also contributes to the constructive interference effect.

Furthermore, by focusing on a QAM modulation scheme, the problem can be finally written decomposing the constraints $\mathcal{C}1$ along the in-phase and quadrature components of the symbols, as follows:

$$\begin{aligned} \boldsymbol{d}_l(\boldsymbol{s}, \boldsymbol{H}, \boldsymbol{A}_{\text{TX}}, \boldsymbol{A}_{\text{RX}}, \boldsymbol{\gamma}, \boldsymbol{v}, \epsilon) = \arg\min_{\boldsymbol{d}_l} \quad & \|(\boldsymbol{I}_N \otimes \boldsymbol{A}_{\text{TX}}^T) \boldsymbol{d}_l\|^2 \\ \text{s.t.} \quad \mathcal{C}1: \quad & \text{Re}(\boldsymbol{g}_k[i] \boldsymbol{d}_l + \boldsymbol{v}_k[i]) \gtrless \sigma_z \sqrt{\gamma_k} \, \text{Re}(\boldsymbol{s}_k[i]), \ k = 1, \dots, K, \ i = 1, \dots, S, \\ \mathcal{C}2: \quad & \text{Im}(\boldsymbol{g}_k[i] \boldsymbol{d}_l + \boldsymbol{v}_k[i]) \gtrless \sigma_z \sqrt{\gamma_k} \, \text{Im}(\boldsymbol{s}_k[i]), \ k = 1, \dots, K, \ i = 1, \dots, S, \\ \mathcal{C}3: \quad & |\boldsymbol{g}_{Uk}[i] \boldsymbol{d}_l|^2 \leq \epsilon, \ k = 1, \dots, K, \ i = 1, \dots, S. \end{aligned} \tag{20}$$

In the last proposed formulation, the same considerations given in the previous section apply for the constraints $\mathcal{C}1$, $\mathcal{C}2$, while the constraints introduced in $\mathcal{C}3$ are convex quadratic ones. Therefore, the final optimization problem in (20) is convex and can be solved resorting to the standard convex optimization tools [29].

V. NUMERICAL RESULTS

This section presents numerical results to show the effectiveness of the proposed scheme with respect to the classical SLP approach [9], [10], which can handle only the interference in the spatial dimension. As shown in [10], [12], SLP already outperforms the conventional block-level precoding schemes and this is why it was selected as benchmark. The performance of the proposed approaches will be assessed in terms of symbol error rate (SER), total transmit power, effective rate and energy efficiency, considering different acceleration factors τ applied to the FTN system. Further, bit error rate (BER) results will be presented to assess the performance of the techniques when a FEC scheme is used.

All the results presented in the remainder of this section are obtained assuming a 16-QAM modulation scheme for the data information, while the number of antennas N and the number of users K are both fixed to 4 (unless specified otherwise). As to the pulse shaping operation, we have modeled it using a SRRC pulse waveform, which is the most used in practical systems, with a roll-off factor of 0.25. The considered oversampling factor n_s is 20. The target SINR is assumed the same for all the users for the sake of simplicity, and it is fixed to 12 dB for all the results[9], while the noise variance σ_z^2 is assumed unitary. Unless specified otherwise, the results are obtained by averaging over several realizations of the spatial channel matrix which is generated, for the generic user k, as $\boldsymbol{h}_k \sim \mathcal{CN}(0, \sigma_h^2 \boldsymbol{I}_N)$, with $\sigma_h^2 = 1$. Finally, the block length S has been set to 50 symbols. In the following, first we focus on the scheme presented in Section III, which does not consider any inter-block interference. Then, the sequential FTN SLP scheme of Section IV is evaluated.

A. Performance Analysis without Inter-block Interference

We focus herein on a scenario with no ISI between multiple symbol blocks. This is the case when a single data block is handled by the SLP scheme, or when the framing structure ensures that the adjacent blocks do not mutually interfere. Accordingly, we focus on the scheme presented in Section III. The presented results are obtained by averaging over several realizations of the data information matrix $\boldsymbol{S}$.

[9]It can be easily checked numerically how a 12 dB SINR can ensure an error-free BER performance when a FEC scheme is used over a 16-QAM modulation, even though the SER appears considerable. Although for the sake of simplicity most of the results presented in this work are obtained without the use of FEC, the choice of the target SINR has been made in light of this consideration, which should be kept in mind for a proper interpretation of the SER results presented later in this section.

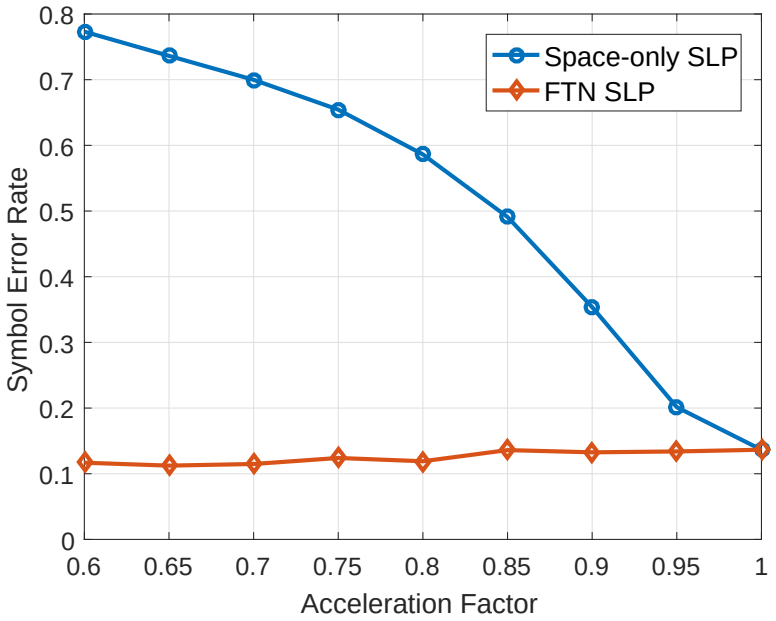

Figure 4: Attained symbol error rate versus acceleration factor.

We start by presenting in Fig. 4 a SER result comparing the proposed FTN SLP approach with the SLP scheme of [10], as a function of the acceleration factor τ. Since the SLP scheme of [10] operates only in the spatial dimension, for each symbol slot, we refer to this technique as space-only SLP. As expected, the SER achieved by the two approaches is the same when no acceleration is applied, since in this case there is no ISI so the schemes are equivalent. When τ is reduced[10], it is apparent how the space-only approach severely suffers the introduced ISI, which is not handled by such scheme, showing a higher and higher SER with decreasing values of τ. On the other hand, the result shows the ability of the proposed technique in managing the ISI for all the considered acceleration factors. Interestingly, the constructive interference effect over the temporal dimension even allows to improve the achieved SER performance[11] when the system is more accelerated.

Another interesting performance metric is the effective sum rate of the system. We define this quantity as:

[10]For the sake of clarity, we stress again that the acceleration factor τ defines the symbol period fraction of the accelerated system compared to the Nyquist system. Thus lower τ means higher acceleration, as discussed in Section II-A.

[11]The attained SER values can be further reduced by increasing the target SINR. However, the application of FEC allows to strongly boost the performance in terms of BER without any SINR increase, as shown later in this section (see also footnote 9).

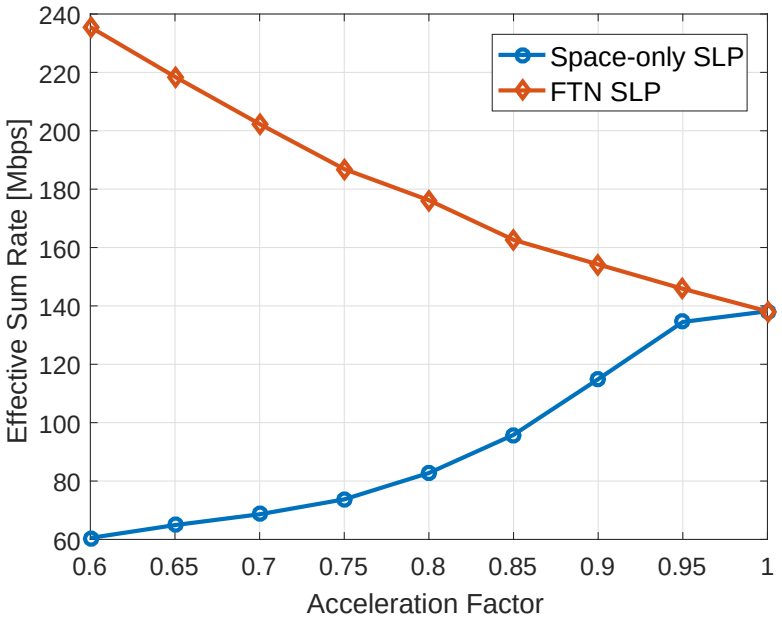

Figure 5: Attained effective sum rate, in Mbps, versus acceleration factor.

$$\bar{R} = \frac{1}{\tau} \sum_{k=1}^{K} R_k (1 - \text{SER}_k), \tag{21}$$

where R_k is the error-free rate for the user k. Such error-free rate (the maximum rate that can be achieved by the used modulation) can be written in turn as $W \log_2(M)$, with W being the user bandwidth and M the modulation order. Considering a user bandwidth of 10 MHz (this value will be also used in the remainder of this section), Fig. 5 compares the effective rate (in Mbps) of the proposed technique with the space-only benchmark, for different acceleration factors. This result shows again the effectiveness of the FTN SLP scheme in exploiting the FTN signaling and handling the ISI.

It is now essential to assess the transmit power required by the proposed scheme, and how it varies with the acceleration factor. The total transmit power, defined in (7), is shown, in dBW, in Fig. 6 as a function of τ. While the transmit power does not depend on τ for the space-only SLP (as this approach does not take into account the acceleration), it is visible how with the FTN SLP it significantly depends on τ. In particular, when τ is too low the required power becomes prohibitive. Nevertheless, if we use a τ not lower than 0.8 (which determines a 25% gain in the rate with respect to the non accelerated case), the power increase is moderate.

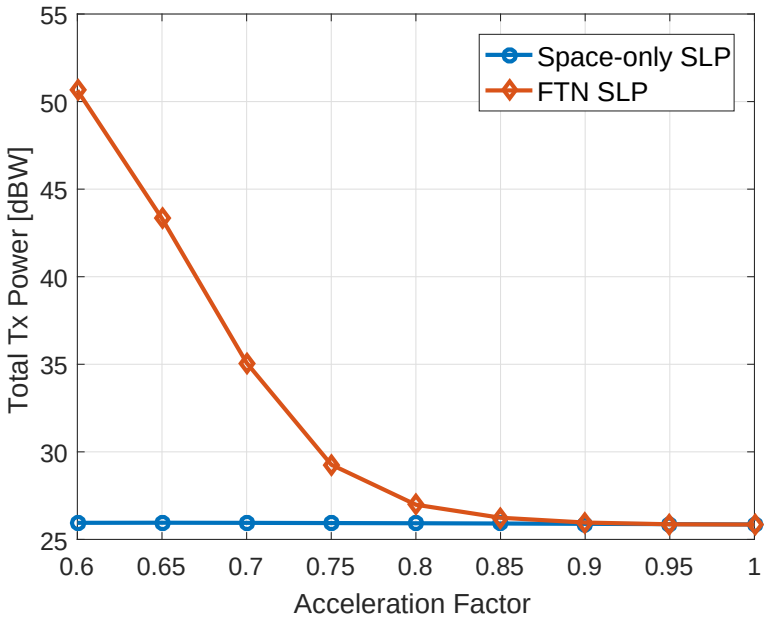

Figure 6: Total transmit power, in dBW, versus acceleration factor.

Finally, a fundamental performance metric which allows to jointly account for the effective rate and the consumed power is the energy efficiency, defined as $\mathcal{E} = \bar{R}/P_{\text{tot}}$, with $\bar{R}$ and P_{tot} defined in (21) and (7) respectively[12]. The related result, in Mbits/J, is shown in Fig. 7. From the result, it turns out that the proposed approach outperforms the space-only scheme in terms of energy efficiency only in a certain range of the acceleration factor, starting from 0.75. Further, it should be highlighted that when $\tau < 0.85$ the energy efficiency of the FTN scheme becomes lower with respect to the non accelerated case, with $\tau = 1$. This means that accelerating the transmission beyond this threshold, although determining a higher rate, is not convenient for the energy efficiency of the system. We can conclude that the optimal τ in this scenario is 0.85, which allows an 18% gain in the rate with respect to the non FTN case.

B. Performance Analysis accounting for the Inter-block Interference

In this section, we consider a scenario where multiple data blocks are sequentially transmitted, without assuming any separation between them. In this case, there is inter-block ISI between

[12]For simplicity, in the computation of the consumed power we are considering only the transmit power, while additional terms due to, for instance, the digital-to-analog conversion or the power amplifiers losses, are neglected. Nevertheless, the main conclusions of this work still apply if a more complicated power model is used.

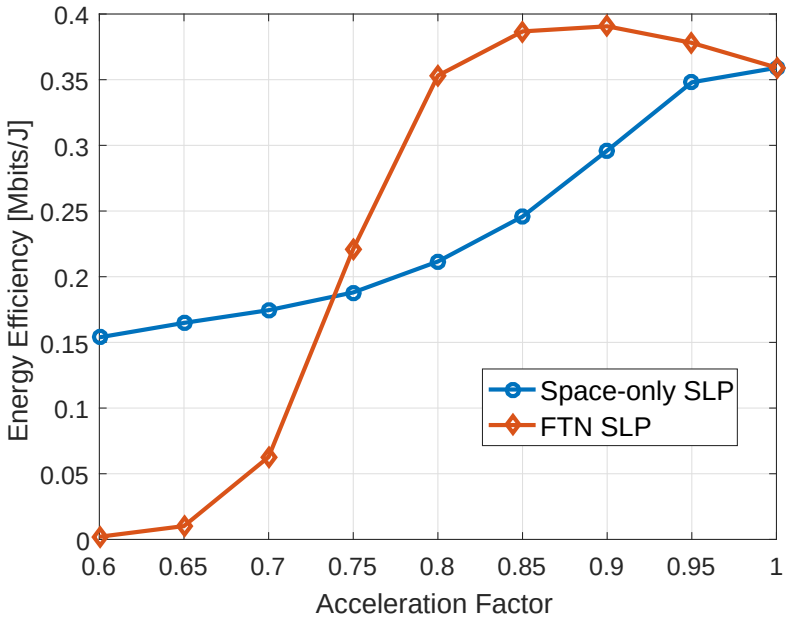

Figure 7: Attained energy efficiency, in Mbits/J, versus acceleration factor.

adjacent blocks, and we can assess the performance of the scheme proposed in Section IV. The presented results are obtained by simulating $L = 10$ sequential blocks. As anticipated, the choice of the parameter ϵ in (20) affects the feasibility of the optimization problem. Moreover, since the problem (20) is more constrained than the one in (11), the permissible range for the acceleration factor τ turns out to be tighter. In fact, the performed numerical analysis has shown that there is a lower bound for τ, which depends on ϵ, under which the problem becomes unfeasible. Hereafter we consider acceleration factors τ in the range $[0.8, 1]$, and we assess the performance of the scheme in (20) for different values of ϵ. In particular, the evaluation is performed for $\epsilon = 3\sigma_z^2$, $\epsilon = 6\sigma_z^2$, and $\epsilon = +\infty$, with the last case corresponding to solving the problem in (20) without the constraint $\mathcal{C}3$.

In Figs. 8-11 we present the obtained numerical results, based on the previously introduced performance metrics. The sequential FTN SLP approach is compared with the non sequential one of Section III, which has to tolerate the inter-block interference arising in the system. Moreover, the space-only SLP is also used as a benchmark. In Figs. 8-9 the obtained SER and the effective sum rate are shown, respectively, as a function of the acceleration factor τ. It is notable how, when the acceleration becomes significant, the non sequential FTN SLP scheme suffers the ISI, showing worse SER and rate performance. On the other hand, the introduced sequential

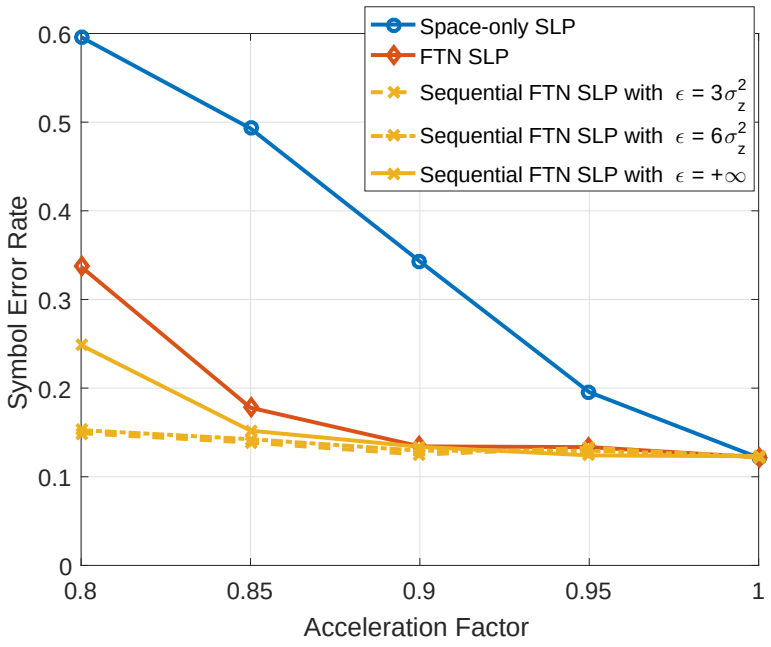

Figure 8: Attained symbol error rate versus acceleration factor, in a multi-block scenario.

FTN SLP scheme is able to handle the inter-block interference, and it does not show any SER degradation when τ is reduced with $\epsilon = 3\sigma_z^2$ and $\epsilon = 6\sigma_z^2$, without a significant difference among these two cases. When the constraint $\mathcal{C}3$ in (20) is relaxed, i.e. for $\epsilon = +\infty$, it can be seen how even the sequential scheme shows a SER degradation for low values of τ, because of the residual inter-block ISI. Nevertheless, it is apparent how the sequential scheme outperforms the benchmarks in terms of SER and effective rate even when ϵ is set to $+\infty$.

As previously discussed, it is fundamental to evaluate the performance also in terms of transmit power and, ultimately, in terms of energy efficiency. The related results are shown in Figs. 10-11. Concerning the total transmit power, it appears from Fig. 10 how the power requirement of the sequential scheme is higher than the non sequential one for $\epsilon = 3\sigma_z^2$ and $\epsilon = 6\sigma_z^2$, especially for low values of τ within the considered range. On the other hand, it emerges how when $\epsilon = +\infty$ the transmit power is comparable to the non sequential scheme, since in this case the problem in (20) has more degrees of freedom. Taking this into account, it is interesting to notice from Fig. 11 how the best approach in terms of energy efficiency is the sequential FTN scheme with $\epsilon = +\infty$, which outperforms the non sequential scheme for all the considered acceleration factors. Conversely, the constrained cases with $\epsilon = 3\sigma_z^2$ and $\epsilon = 6\sigma_z^2$ show a worse energy efficiency due to their higher transmit power. For this reason, in the remainder of this

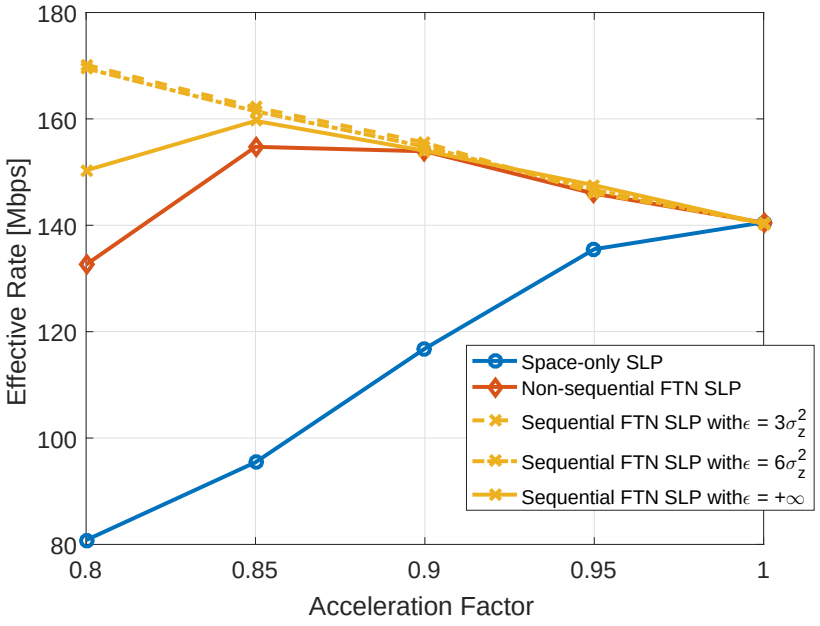

Figure 9: Attained effective sum rate, in Mbps, versus acceleration factor, in a multi-block scenario.

section we will mainly focus on the case with $\epsilon = +\infty$, which is assumed when not specified otherwise. Concerning the acceleration factor, it can be observed how for $\tau \geq 0.85$ the energy efficiency attained by the proposed sequential FTN scheme is not lower than the non accelerated case ($\tau = 1$). Accordingly, also in this multi-block scenario the optimal τ is 0.85, which allows an 18% gain in the rate with respect to the non FTN case.

An additional result is displayed in Fig. 12, which shows how the achieved performance changes when more antennas are utilized at the transmitter. More specifically, the energy efficiency and the effective rate are represented as a function of the number of antennas N, when $\tau = 0.85$. Remarkably, while the effective rate does not show substantial changes, the attained energy efficiency considerably grows with N due to the reduced transmit power required to achieve the target SINR. This effect is related to the improved constructive interference effect taking place when more antennas are used, and has been discussed also in [10].

C. Effect of imperfect CSI and BER analysis with FEC

As previously mentioned, we have assumed so far a perfect knowledge of the CSI for the presented techniques, as a robust version of them to CSI errors falls out of the scope of this contribution. Nevertheless, it is worth analyzing how the different schemes are sensitive

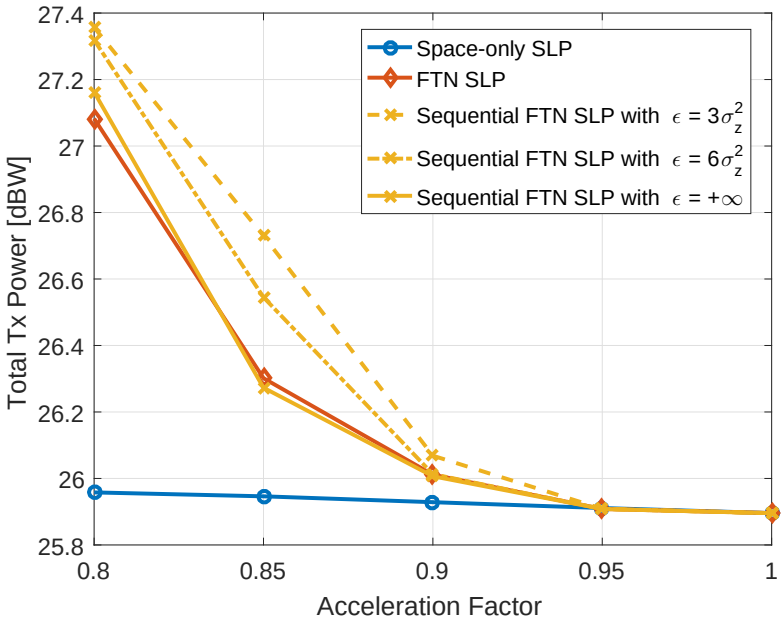

Figure 10: Total transmit power, in dBW, versus acceleration factor, in a multi-block scenario.

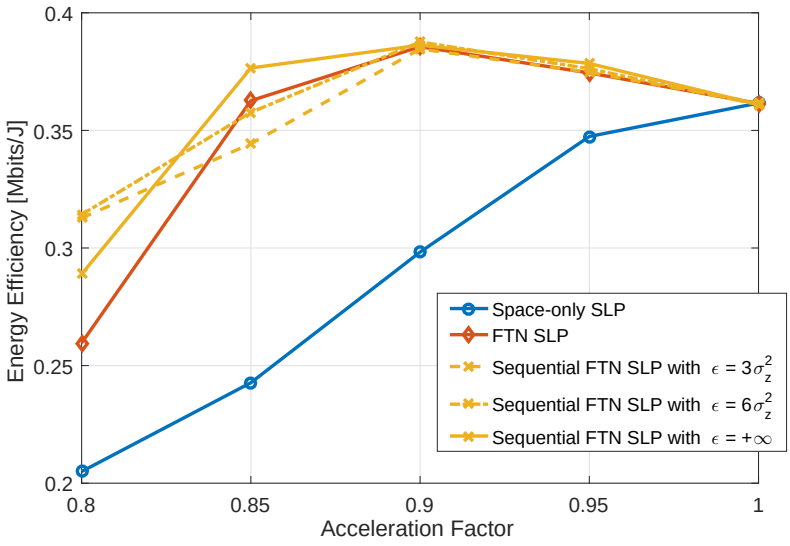

Figure 11: Attained energy efficiency, in Mbits/J, versus acceleration factor, in a multi-block scenario.

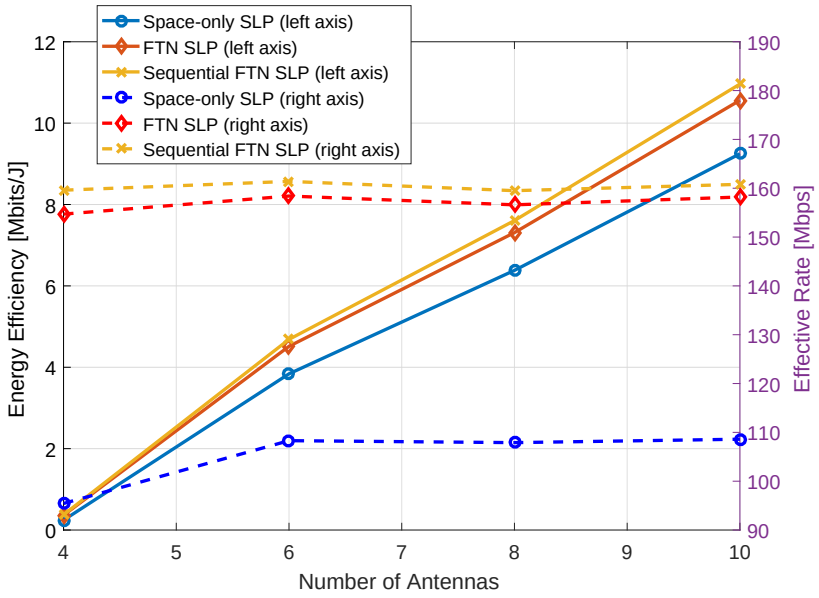

Figure 12: Attained energy efficiency and effective rate versus number of antennas, for $\tau = 0.85$.

to channel uncertainties, which are usually present in practical applications. Consequently we can model the CSI estimate for the k-th user as $\hat{\boldsymbol{h}}_k = \boldsymbol{h}_k + \boldsymbol{e}_k$ (see also [3]), with $\boldsymbol{e}_k$ being an error vector modeled as $\boldsymbol{e}_k \sim \mathcal{CN}(0, \sigma_e^2 \boldsymbol{I}_N)$. In Fig. 13 we present the obtained SER for the introduced schemes for different values of τ and σ_e^2, focusing on the multi-block scenario (moreover, for simplicity from now on we consider a fixed channel matrix $\boldsymbol{H}$, as in (22)). First of all, it can be noticed how the SER degradation induced by the CSI error is limited for $\sigma_e^2 = 10^{-4}$ and $\sigma_e^2 = 10^{-3}$, while when $\sigma_e^2 = 10^{-2}$ the performance is extremely degraded for all the schemes. Notably, the non sequential FTN SLP scheme and the sequential one show a comparable sensitivity to CSI errors, and their SER degradation does not substantially vary with τ. As to the space-only scheme, it shows the same sensitivity of the proposed ones only for high acceleration factors ($\tau > 0.9$), while when τ is decreased the SER degradation is reduced (however, it should be noted that in this case the SER is already very high with perfect CSI, thus the CSI errors are less determinant).

Finally, we present in Fig. 14 a result in terms of BER obtained using a FEC scheme with the proposed techniques. In particular, a low-density parity check (LDPC) code has been used, with a code rate $5/6$. Both the coded BER and the uncoded BER are shown for the different approaches as a function of τ, and the case with $\epsilon = 6\sigma_z^2$ is also considered (besides $\epsilon = +\infty$) for the sequential FTN SLP scheme. Remarkably, it can be seen how the application of LDPC

$$\boldsymbol{H} = \begin{bmatrix} -0.6753 + 0.7875j & -0.9406 + 0.3229j & -0.1814 + 0.2493j & 0.1856 - 0.7855j \\ -0.0214 + 0.3383j & -1.1166 + 1.0522j & 0.0435 - 1.1963j & -0.8914 - 0.2106j \\ 0.1734 - 1.0120j & -1.1399 + 1.5326j & -0.5112 + 1.3084j & -0.8879 + 0.3026j \\ 0.0145 - 0.0956j & -0.5300 + 0.0520j & -0.5749 - 0.0975j & -0.2097 + 0.3532j \end{bmatrix}, \tag{22}$$

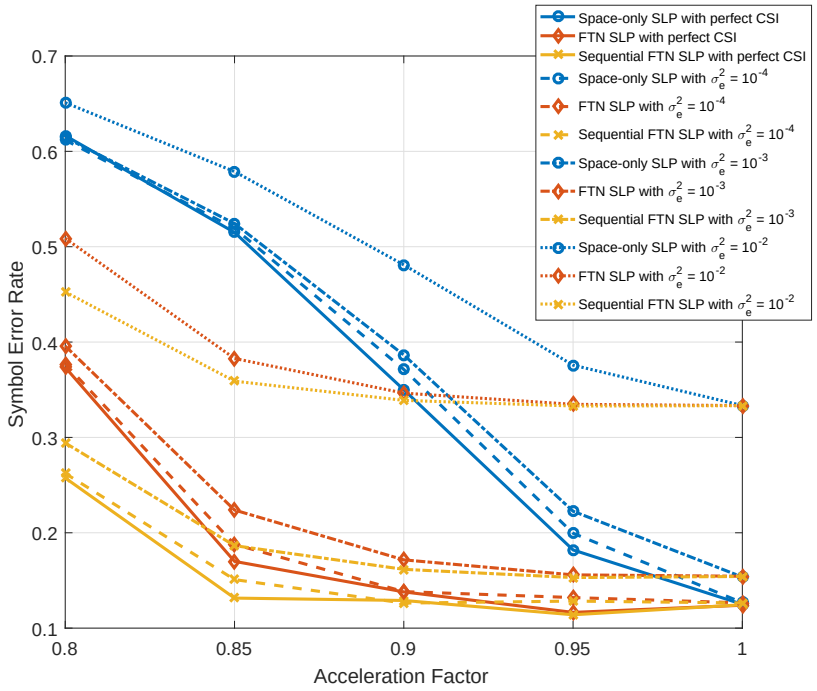

Figure 13: Attained symbol error rate versus acceleration factor, considering CSI errors.

allows to strongly improve the BER performance achieved by the proposed schemes, with an error-free communication for $\tau \geq 0.85$ for the sequential SLP with $\epsilon = 6\sigma_z^2$. As expected, it can be also noticed how the gains given by FEC disappear when the uncoded BER becomes too high due to ISI (in particular, this happens for $\tau \leq 0.9$ for the space-only SLP, and for $\tau = 0.8$ for the non sequential FTN SLP and for the relaxed sequential one).

D. Numerical Evaluation of the Complexity

We conclude this section by providing a complexity evaluation of the proposed approaches as a function of the number of symbols to be processed (for each user stream). Since the proposed optimization problems are tackled resorting to numerical convex optimization tools

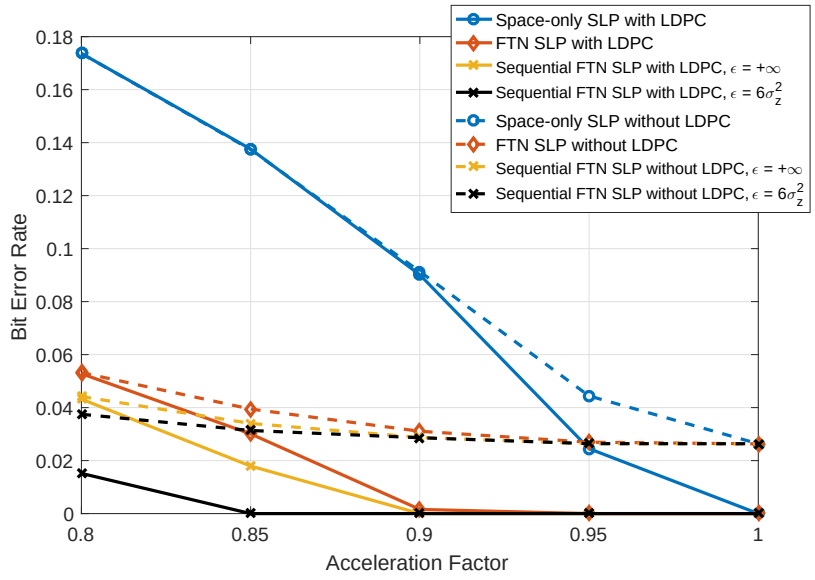

Figure 14: Attained bit error rate versus acceleration factor, considering a LDPC scheme.

[29], analytical expressions for the complexity are hard to derive. Thus herein the complexity is numerically evaluated in terms of average running time of the algorithms over the same machine. Before presenting such evaluation, we should stress that in general complexity is one of the key challenges of SLP with respect to conventional channel-level precoding schemes: in fact, while in the latter ones the precoder optimization is performed once per channel coherence time, SLP (e.g. [10]) presents a much higher switching rate, equal to the baud rate. Nonetheless, the spatio-temporal precoding introduced herein has an advantage in this regard, since the optimization procedure applies once per block and not once per symbol slot. This can be observed in Fig. 15, which presents how the average running time varies with the number of symbols to be processed per stream, for $\tau = 0.85$. We consider the spatio-temporal SLP, which performs the optimization symbol by symbol, the non sequential FTN SLP, which processes the whole symbol stream all at once, and the sequential FTN SLP scheme, which processes the data divided in blocks of $S = 50$ symbols. As expected, it can be observed how both the space-only approach and the sequential FTN one present a linear dependence between the running time and the number of symbols (blocks), but the latter one shows a remarkably reduced running time due to the block processing. On the other hand, it can be also noticed how, in the case of the non sequential FTN SLP, the complexity grows in a superlinear way with the number of processed symbols, and this justifies the need of resorting to a serial processing of multiple sequential blocks. Overall, the

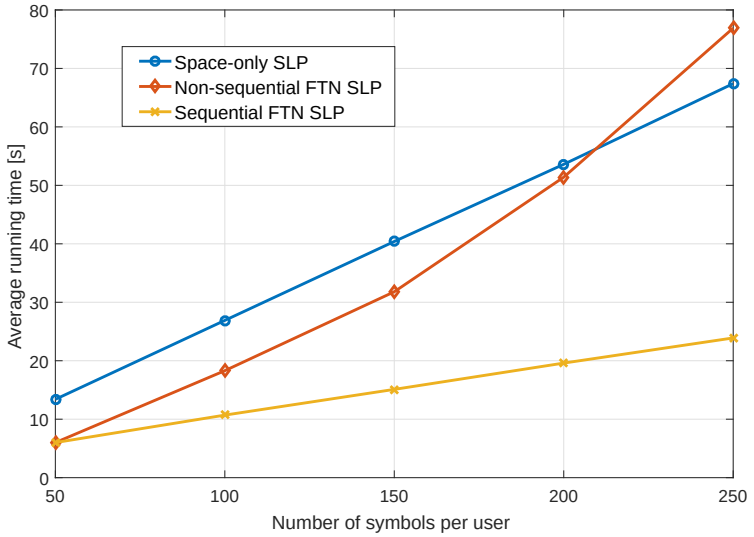

Figure 15: Average running time of the algorithms, in seconds, versus stream length.

sequential FTN SLP scheme appears to be the most suitable in terms of complexity. In general, complexity still remains an open challenge of the SLP framework, especially when the system dimensions (e.g. number of users/antennas) is high. An efficient implementation of SLP schemes aimed at reducing their inherent complexity, and eventually at allowing a real-time processing, is part of the ongoing and future work [33].

VI. Conclusions

In this work, a novel symbol-level precoding strategy has been proposed, which handles, at the transmitter side, not only the interference in the spatial dimension (the MUI), but also the interference in the temporal dimension (the ISI). This new precoding method, named spatio-temporal symbol-level precoding, is used to apply faster-than-Nyquist signaling over multiuser MISO systems. The introduced strategy splits the data streams in blocks of symbols, and processes the blocks so as to exploit in a constructive fashion the interference both in the temporal dimension and in the spatial one, thus gleaning benefits from both the domains. Firstly, we have proposed and formalized a sum power minimization scheme, with QoS constraints, which tackles the interference within each data block. Afterwards, an extension of the optimization scheme has been presented, which tackles a scenario with multiple mutually interfering data blocks, and is able to manage the inter-block ISI arising between adjacent blocks. The performance

of the proposed schemes has been assessed through numerical simulations over a multiuser MISO system using SRRC pulse shaping, in terms of achieved SER, effective sum rate, energy efficiency, and total transmit power. The results have shown how the proposed spatio-temporal precoding outperforms the existent symbol-level precoding solutions in terms of effective rate and energy efficiency, for acceleration factors in the order of 0.8-0.9. Further, the gain in the system rate (thus, in energy efficiency) of the FTN schemes has been discussed with respect to the classical Nyquist transmission. In the future work, a possible extension of the proposed schemes is foreseen, where the spatio-temporal SLP framework is used to tackle the ISI generated by frequency selective channels, or by non-linear components.

REFERENCES

[1] R. Roy and B. Ottersten, "Spatial division multiple access wireless communication systems," May 1996, US Patent 5,515,378. [Online]. Available: https://www.google.com/patents/US5515378

[2] Y.-F. Liu, Y.-H. Dai, and Z.-Q. Luo, "Coordinated beamforming for MISO interference channel: Complexity analysis and efficient algorithms," *IEEE Transactions on Signal Processing*, vol. 59, no. 3, pp. 1142–1157, 2011.

[3] M. Bengtsson and B. Ottersten, "Optimal and suboptimal transmit beamforming," in *Handbook of Antennas in Wireless Communications*. CRC Press, 2001.

[4] M. Schubert and H. Boche, "Solution of the multiuser downlink beamforming problem with individual SINR constraints," *IEEE Transactions on Vehicular Technology*, vol. 53, no. 1, pp. 18–28, Jan 2004.

[5] W. Yu and T. Lan, "Transmitter optimization for the multi-antenna downlink with per-antenna power constraints," *IEEE Transactions on Signal Processing*, vol. 55, no. 6, pp. 2646–2660, June 2007.

[6] G. Dartmann, X. Gong, W. Afzal, and G. Ascheid, "On the duality of the max min beamforming problem with per-antenna and per-antenna-array power constraints," *IEEE Transactions on Vehicular Technology*, vol. 62, no. 2, pp. 606–619, Feb 2013.

[7] C. Masouros and E. Alsusa, "Dynamic linear precoding for the exploitation of known interference in MIMO broadcast systems," *IEEE Transactions on Wireless Communications*, vol. 8, no. 3, pp. 1396–1404, March 2009.

[8] C. Masouros, "Correlation rotation linear precoding for MIMO broadcast communications," *IEEE Transactions on Signal Processing*, vol. 59, no. 1, pp. 252–262, Jan. 2011.

[9] M. Alodeh, S. Chatzinotas, and B. Ottersten, "Constructive multiuser interference in symbol level precoding for the MISO downlink channel," *IEEE Transactions on Signal Processing*, vol. 63, no. 9, pp. 2239–2252, May 2015.

[10] M. Alodeh, S. Chatzinotas, and B. Ottersten, "Symbol-level multiuser MISO precoding for multi-level adaptive modulation," *IEEE Transactions on Wireless Communications*, vol. 16, no. 8, pp. 5511–5524, Aug 2017.

[11] M. Alodeh, S. Chatzinotas, and B. Ottersten, "Energy-efficient symbol-level precoding in multiuser MISO based on relaxed detection region," *IEEE Transactions on Wireless Communications*, vol. 15, no. 5, pp. 3755–3767, May 2016.

[12] C. Masouros and G. Zheng, "Exploiting known interference as green signal power for downlink beamforming optimization," *IEEE Transactions on Signal Processing*, vol. 63, no. 14, pp. 3628–3640, July 2015.

[13] D. Spano, M. Alodeh, S. Chatzinotas, and B. Ottersten, "Per-antenna power minimization in symbol-level precoding," in *2016 IEEE Global Communications Conference (GLOBECOM)*, Dec 2016, pp. 1–6.

[14] D. Spano, S. Chatzinotas, J. Krause, and B. Ottersten, "Symbol-level precoding with per-antenna power constraints for the multi-beam satellite downlink," in *2016 8th Advanced Satellite Multimedia Systems Conference and the 14th Signal Processing for Space Communications Workshop (ASMS/SPSC)*, Sept 2016, pp. 1–8.

[15] D. Spano, M. Alodeh, S. Chatzinotas, J. Krause, and B. Ottersten, "Spatial PAPR reduction in symbol-level precoding for the multi-beam satellite downlink," in *2017 IEEE 18th International Workshop on Signal Processing Advances in Wireless Communications (SPAWC)*, July 2017. [Online]. Available: http://hdl.handle.net/10993/31224

[16] D. Spano, M. Alodeh, S. Chatzinotas, and B. Ottersten, "Symbol-level precoding for the nonlinear multiuser miso downlink channel," *IEEE Transactions on Signal Processing*, vol. 66, no. 5, pp. 1331–1345, March 2018.

[17] C. Masouros and E. Alsusa, "Soft linear precoding for the downlink of ds/cdma communication systems," *IEEE Transactions on Vehicular Technology*, vol. 59, no. 1, pp. 203–215, Jan 2010.

[18] M. Alodeh, D. Spano, A. Kalantari, C. Tsinos, D. Christopoulos, S. Chatzinotas, and B. Ottersten, "Symbol-level and multicast precoding for multiuser multiantenna downlink: A state-of-the-art, classification and challenges," *IEEE Communications Surveys & Tutorials*, 2017, submitted. [Online]. Available: https://arxiv.org/pdf/1703.03617.pdf

[19] J. E. Mazo, "Faster-than-Nyquist signaling," *The Bell System Technical Journal*, vol. 54, no. 8, pp. 1451–1462, Oct 1975.

[20] A. D. Liveris and C. N. Georghiades, "Exploiting faster-than-Nyquist signaling," *IEEE Transactions on Communications*, vol. 51, no. 9, pp. 1502–1511, Sept 2003.

[21] Y. J. D. Kim, J. Bajcsy, and D. Vargas, "Faster-than-Nyquist broadcasting in gaussian channels: Achievable rate regions and coding," *IEEE Transactions on Communications*, vol. 64, no. 3, pp. 1016–1030, March 2016.

[22] F. Rusek and J. B. Anderson, "The two dimensional Mazo limit," in *Proceedings. International Symposium on Information Theory, 2005. ISIT 2005.*, Sept 2005, pp. 970–974.

[23] A. Barbieri, D. Fertonani, and G. Colavolpe, "Time-frequency packing for linear modulations: spectral efficiency and practical detection schemes," *IEEE Transactions on Communications*, vol. 57, no. 10, pp. 2951–2959, October 2009.

[24] A. Modenini, F. Rusek, and G. Colavolpe, "Faster-than-Nyquist signaling for next generation communication architectures," in *2014 22nd European Signal Processing Conference (EUSIPCO)*, Sept 2014, pp. 1856–1860.

[25] J. B. Anderson, F. Rusek, and V. Öwall, "Faster-than-Nyquist signaling," *Proceedings of the IEEE*, vol. 101, no. 8, pp. 1817–1830, Aug 2013.

[26] M. Costa, "Writing on dirty paper," *IEEE Trans. Inf. Theory*, vol. 29, no. 3, pp. 439–441, 1983.

[27] M. Alodeh, D. Spano, S. Chatzinotas, and B. Ottersten, "Faster-than-Nyquist spatiotemporal symbol-level precoding in the downlink of multiuser MISO channels," in *2017 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, March 2017.

[28] M. Alodeh, D. Spano, and S. Chatzinotas, "Spatio-temporal precoding for faster-than-Nyquist signal transmissions," March 2017, Patent LU100110.

[29] S. Boyd and L. Vandenberghe, *Convex optimization*. Cambridge Univ. Press, 2004.

[30] S. Haykin, "Cognitive radio: brain-empowered wireless communications," *IEEE Journal on Selected Areas in Communications*, vol. 23, no. 2, pp. 201–220, Feb 2005.

[31] A. Goldsmith, S. A. Jafar, I. Maric, and S. Srinivasa, "Breaking spectrum gridlock with cognitive radios: An information theoretic perspective," *Proceedings of the IEEE*, vol. 97, no. 5, pp. 894–914, May 2009.

[32] M. Alodeh, D. Spano, S. Chatzinotas, and B. Ottersten, "Peak power minimization in symbol-level precoding for cognitive miso downlink channels," in *2016 IEEE International Conference on Digital Signal Processing (DSP)*, Oct 2016, pp. 240–244.

[33] J. Krivochiza, A. Kalantari, S. Chatzinotas, and B. Ottersten, "Computationally efficient symbol-level precoding communications demonstrator," in *Symposium on Information Theory and Signal Processing in the Benelux*, May 2017.