# Spatio-Temporal Delta-Sigma Modulation for Massive MIMO with Low Resolution DACs

Nicolas Schlegel*†, Chadi Jabbour† and Philippe Ciblat†
* Nokia Bell Labs, Massy, France
† LTCI, Telecom Paris, Institut Polytechnique de Paris

***Abstract*—Going toward extreme multiple input multiple output (MIMO) antenna systems, each individual radio frequency (RF) chain will require cost and energy efficient implementations. Among the power hungry components is the digital-to-analog converter (DAC). While lowering the resolution can reduce power consumption, the introduced quantization error will lead to degraded performance. In this paper, we propose to use spatio-temporal delta sigma modulation (DSM) upstream of the low-resolution converter in a downlink massive MIMO-orthogonal frequency division multiplexing (OFDM) scenario. Its role is to shape quantization noise toward unused frequency resources and directions of space. At the heart lies an optimization based design of the modulator's noise transfer function (NTF) using available channel state information (CSI). A constraint to shape quantization noise away from adjacent frequency bands is studied as well. Numerical results show that the recently thriving spatial DSM is outperformed by its spatio-temporal counterpart. Additional filtering is however still required to reach spectral masks set by current communication standards.**

***Index Terms*—Delta-Sigma modulation, spatio-temporal, MIMO, OFDM, quantization.**

## I. Introduction

Massive multiple input multiple output (MIMO) has become a pillar technology in contemporary communication systems. To provide the further needs for increased capacity, larger antenna arrays are envisioned. Scalability of this approach will require a careful redesign of the transceiver radio frequency (RF) chain to keep power consumption, cost and implementation constraints low. In partial load scenarios, the digital-to-analog converters (DACs) become significant power consumers in the transmit chain [1], rising exponentially with resolution [2]. This motivated the study of low-resolution DACs in massive MIMO systems, with various approaches to deal with the quantization error emerging from the literature. Prohibitive complexity and incompatibility with spectral masks imposed by communication standards still hinder practical implementations.

As massive MIMO systems equipped with low-resolution DACs and conventional linear precoders suffer from important performance degradation [3], a common approach to handle the quantization distortion is through the design of quantized precoders [4]–[7]. In this paradigm, a precoder generates a signal directly on the set of quantization levels. Consequently, no additional quantization error is incurred from passing through the low-resolution DAC. For this task, the best performing precoding methods revolve around finding fast sub-optimal, hyper-parameter dependent solutions to integer optimization problems [5]–[7]. Additionally, the better performing symbol dependent precoders require recomputation for every transmitted orthogonal frequency division multiplexing (OFDM) symbol, leading to high computational complexity. Also often overlooked is the impact of low-resolution DACs on out-of-band (OOB) emissions [8]. Multiple designs dismiss this issue and are therefore incompatible with standard required constraints in this area [9, section 6.6].

Another approach to the problem relies on pre-processing the signal before quantization. In delta sigma modulation (DSM), a higher resolution signal is encoded into a lower resolution counterpart while shaping the quantization noise introduced in the process toward unused resources. When applied to the temporal domain, some oversampling is required to separate the noise into frequency bands unoccupied by the signal. With the recent growth of the antenna arrays, spatial noise shaping has regained attention. The idea is to steer quantization noise into directions that are unoccupied by served users by designing the noise transfer function (NTF) in an appropriate manner. In [10], the NTF design relied on angles of departure (AoDs) or angle sectors occupied by users. This approach was then extended to take full channel state information (CSI) into account [11]. More specifically, flat fading multi-path channels were considered, with the designed NTF leveraging the path loss for each direction. Both methods are symbol agnostic and only require recomputation at every channel coherence period. In [12], a quantized precoding algorithm and a first-order spatial DSM were combined.

Spatio-temporal DSM has received some past attention [13]–[15], showing that quantization noise can indeed be shaped in frequency and space, with a lesser focus on the NTF design itself.

Prior work has only looked at flat fading channels and therefore did not need to consider OFDM modulations. Our contribution in this paper fills this gap and considers spatio-temporal NTF design using full CSI of frequency selective multi-path channels in a downlink massive MIMO-OFDM scenario. Commonly used stability constraints used in the optimization based design of the modulator are compared against each other. In an effort to reach standard compliance, the problem is extended to minimize OOB emissions caused by the coarse quantization.

This paper is structured as follows. Section II introduces the considered system model. Section III exhibits the NTF design

This work has been submitted to the IEEE for possible publication. Copyright may be transferred without notice, after which this version may no longer be accessible.

problem and its solution. Section IV provides numerical results. Section V concludes the paper.

### A. Notations

Scalar values, column vectors and matrices are represented by normal lowercase, bold lowercase and bold uppercase respectively. The transpose and hermitian operations are denoted by $(\cdot)^{\mathsf{T}}$ and $(\cdot)^{\mathsf{H}}$. The imaginary unit is denoted as $j$. Discrete variables appearing in both time and frequency have their latter form denoted with $\tilde{(\cdot)}$. The index $n$ is preferred for the time-domain while the index $k$ is rather used for the frequency domain. The $L_2$ norm associated with a positive definite matrix $\mathbf{A}$ and a vector $\mathbf{x} = [x_1, \cdots, x_{K-1}]^{\mathsf{T}}$ is defined as $\|\mathbf{x}\|_{2,\mathbf{A}} = \sqrt{\mathbf{x}^{\mathsf{H}}\mathbf{A}\mathbf{x}}$. If $\mathbf{A} = \mathbf{Id}$ is the identity matrix, we denote $\|\mathbf{x}\|_{2,\mathbf{Id}}$ simply by $\|\mathbf{x}\|_2$. The $L_{1,\mathsf{IQ}}$ norm is defined as $\|\mathbf{x}\|_{1,\mathsf{IQ}} = \sum_{k=0}^{K-1} |\mathcal{R}(x_k)| + |\mathcal{I}(x_k)|$ and the $L_{\infty,\mathsf{IQ}}$ norm as $\|\mathbf{x}\|_{\infty,\mathsf{IQ}} = \max_{k\in\{0,\cdots,K-1\}} \max\left(|\mathcal{R}(x_k)|, |\mathcal{I}(x_k)|\right)$.

## II. System model

A massive MIMO-OFDM model is considered in the downlink scenario, as illustrated in Fig. 1, where there are one base station composed by $B$ antennas and $U$ users each equipped with one antenna. We assume $U < B$. The output of the base station at time $n$ is a vector $\mathbf{x}[n]$ of length $B$. The receive signal at time $n$ for user $u$ is a scalar denoted by $y_u[n]$. When referring to the $N$-length OFDM frame for a given antenna or user index, the signals are the vectors $\mathbf{x}_b = [x_b[0], \cdots, x_b[N-1]]^{\mathsf{T}}$ and $\mathbf{y}_u$.

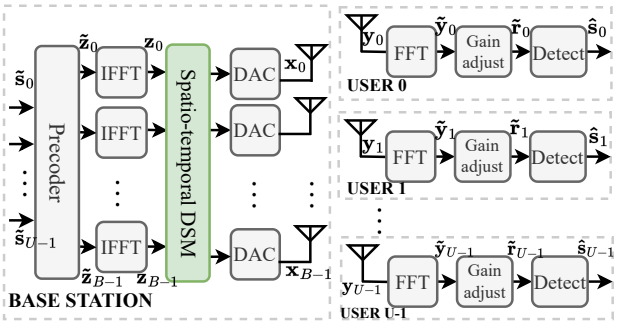


Fig. 1. Block diagram of the system model.

The signal $\mathbf{x}[n]$ is transmitted over a multipath frequency selective channel $\mathbf{H}[t] \in \mathbb{C}^{U\times B}, t \in [\![0, T-1]\!]$, with $T$ the number of taps. A plane wave model, such as in [8], is used. With the antenna elements disposed in a uniform linear array (ULA), the entries of $\mathbf{H}[t]$ are defined as

$$h_{u,b}[t] = \psi_u \gamma_{u,t} e^{-j2\pi\alpha(b-1)\sin(\phi_{u,t})}. \tag{1}$$

$\psi_u^2 = (100/\delta_u)^2$ denotes the large scale fading, with $\delta_u$ being the distance in meters from the base station to user $u$. The inter-antenna spacing in multiples of the carrier wavelengths is $\alpha$. The power delay profile is given by $\gamma_{u,t}$ where $\gamma_{u,0}$ stands for the line of sight (LoS) component while $\gamma_{u,t}$ for $t \neq 0$ stands for the non line of sight (NLoS) components. $\phi_{u,t}$ corresponds to the angle of departure for user $u$ at the time $t$.

Then the received signal $\mathbf{y}[n] = [y_0[n], \cdots, y_{U-1}[n]]^{\mathsf{T}} \in \mathbb{C}^U, n \in [\![0, N-1]\!]$ is expressed as

$$\mathbf{y}[n] = \sum_{t=0}^{T-1} \mathbf{H}[t]\mathbf{x}[n-t] + \boldsymbol{\eta}[n], \tag{2}$$

where $\boldsymbol{\eta}[n]$ is a white circularly-symmetric Gaussian noise. Under the assumption of sufficiently long cyclic prefix added on the frame $[\mathbf{x}[0], \cdots, \mathbf{x}[N-1]]$ of length $N$, taking (2) to the frequency domain gives

$$\tilde{\mathbf{y}}[k] = \tilde{\mathbf{H}}[k]\tilde{\mathbf{x}}[k] + \tilde{\boldsymbol{\eta}}[k], \tag{3}$$

where $\tilde{\mathbf{x}}[k] = [\tilde{x}_0[k], \cdots, \tilde{x}_{B-1}[k]]^{\mathsf{T}}$, with $\tilde{x}_b[k]$ the output at frequency $k$ of the Fast Fourier Transform (FFT) of $\mathbf{x}_b$ for antenna $b$. A similar definition holds for $\tilde{\mathbf{y}}[k]$.

### A. System model without hardware impairments

When the system is equipped with high resolution DACs, the quantization error introduced by the conversion can be neglected. The following development describes the system in the absence of any hardware impairment.

The symbols $\tilde{\mathbf{s}}[k] \in \mathbb{C}^U$, with $k \in [\![0, N-1]\!]$, are transmitted. As OFDM will be used, the index $k$ can be interpreted as a frequency bin. Not all frequency bins carry data. A partition of the set is defined : the bins in $\mathcal{D}$ carry useful symbols while the bins in $\mathcal{G}$ (guard bands) and $\mathcal{A}$ (adjacent bands) are empty, with $\mathcal{A}$ subject to spectral emission restrictions. Finally $\mathcal{D} \cup \mathcal{G} \cup \mathcal{A} = [\![0, N-1]\!]$.

The symbols pass through the linear precoder $\mathbf{P} \in \mathbb{C}^{NB\times NU}$. Working frequency bin per frequency bin, the precoder has a block diagonal structure where the $k$-th block in the diagonal is denoted by $\mathbf{P}[k] \in \mathbb{C}^{B\times U}, k \in [\![0, N-1]\!]$ and its application to the symbols results in

$$\tilde{\mathbf{z}}[k] = \mathbf{P}[k]\tilde{\mathbf{s}}[k]. \tag{4}$$

As the focus is not on the precoder, a zero forcing (ZF) precoder is chosen for simplicity, although the later development can as well be applied with other precoding strategies. The ZF precoder is defined by

$$\mathbf{P}_{\text{ZF}}[k] = \frac{1}{\beta}\tilde{\mathbf{H}}[k]^{\mathsf{H}}(\tilde{\mathbf{H}}[k]\tilde{\mathbf{H}}[k]^{\mathsf{H}})^{-1}, \tag{5}$$

with $\beta > 0$ a scaling factor to ensure that the resulting transmit signal meets a given power constraint.

In order to implement OFDM, an Inverse Fast Fourier Transform (IFFT) per antenna $b$ is applied on $\tilde{\mathbf{z}}_b$. Then we obtain the sample $z_b[n]$ at time $n$ for antenna $b$. Stacking the samples with respect to the antenna indices gives

$$\mathbf{z}[n] = [z_1[n], \cdots, z_B[N-1]]^{\mathsf{T}}, \quad n \in [\![0, N-1]\!]. \tag{6}$$

In the absence of quantization, we have $\mathbf{x}[n] = \mathbf{z}[n]$. Finally, the precoder scaling factor $\beta$ is chosen such that the transmit signal verifies $\frac{1}{|D|}\sum_{n=0}^{N-1} \mathbb{E}\left[\|\mathbf{x}[n]\|_2^2\right] \leq E_t$.

On the receive side, for each user, the cyclic prefix is removed and the FFT operation is applied. Then the corresponding samples $\tilde{\mathbf{y}}[k]$, given in (3), are just scaled back, leading to

$$\tilde{\mathbf{r}}[k] = \beta \tilde{\mathbf{y}}[k]. \tag{7}$$

Given that a ZF precoder was considered, combining equations (3), (4) and (5) into (7) leads to

$$\tilde{\mathbf{r}}[k] = \tilde{\mathbf{s}}[k] + \beta \tilde{\boldsymbol{\eta}}[k], \quad k \in [\![0, N-1]\!], \tag{8}$$

on which a threshold detector may be applied to retrieve the symbols.

### B. System model with low-resolution DACs

We now consider the case with low-resolution DACs. Consequently, $\mathbf{x}[n] \neq \mathbf{z}[n]$ due to the coarse quantization operation. In order to reduce the impact of quantization error on the resources of interest, a spatio-temporal DSM encodes the signal onto the quantization levels prior to the conversion. The goal of this paper is to propose a design of this DSM such that the quantization noise is :

- not cast into angles targeting users. This property may be satisfied thanks to the "beamforming" ability of the $B$ transmit antennas.
- not cast into adjacent frequency bands regardless of the angles. This property is imposed to remain compliant with current communication standards imposing spectral masks.

The quantization noise shaping is dictated by the NTF. To summarize these ideas, Fig. 2 schematically represent the ideal NTF where the red area is forbidden while the green one is authorized for the quantization noise location.

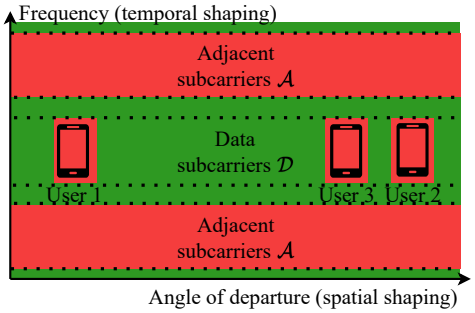


Fig. 2. Illustration of a target NTF for the spatio-temporal DSM. The areas in green can be used for casting quantization noise, while the others should be avoided

DSM can be represented thanks to the so-called error-feedback architecture [16, section 7.2.2] as illustrated in Fig. 3, where $Q(\cdot)$ is the uniform scalar quantization operation applied separately to the real and imaginary parts, and $g$ is the feedback filter. As the NTF is driven by $g$, the goal of this paper is actually to design $g$.

The filter $g$ is assumed to be 2-D finite impulse response (FIR) with support $[\![0, T_{\text{ant}}-1]\!] \times [\![0, T_{\text{time}}-1]\!]$, with $T_{\text{ant}} \leq B$

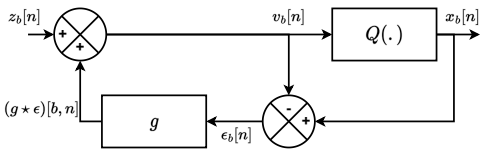


Fig. 3. Corresponding feedback architecture for the STP-aided quantizer.

and $T_{\text{time}} \leq N$. Moreover, the operations mentioned in Fig. 3 can be written as follows

$$\begin{aligned} x_b[n] &= Q(v_b[n]) = v_b[n] + \epsilon_b[n] \\ v_b[n] &= z_b[n] + (g \star \epsilon)_b[n] \end{aligned} \tag{9}$$

where $\star$ is the 2-D convolution defined by

$$(g \star \epsilon)_b[n] = \sum_{\ell=0}^{T_{\text{ant}}-1} \sum_{m=0}^{T_{\text{time}}-1} g_\ell[m] \epsilon_{b-\ell}[n-m]. \tag{10}$$

Defining $f = \delta + g$, with $\delta$ the 2-D Dirac pulse, the input-output relationship of DSM takes the expression

$$x_b[n] = z_b[n] + (f \star \epsilon)_b[n]. \tag{11}$$

In the rest of the paper, we make the common assumption that $\epsilon_b[n]$ is white and independent of the input signal $z_b[n]$, see e.g. [10], [16].

Applying a 2-D Fourier transform (FT) on equation (11) results in

$$X(\psi, \nu) = Z(\psi, \nu) + F(\psi, \nu) E(\psi, \nu), \tag{12}$$

where

$$F(\psi, \nu) = \sum_{b=0}^{T_{\text{ant}}-1} \sum_{n=0}^{T_{\text{time}}-1} f_b[n] e^{-2i\pi\psi b} e^{-2i\pi\nu n}, \tag{13}$$

and the other 2-D FT are defined similarly. The NTF to be optimized can now be identified to $F(\psi, \nu)$.

As DSM incorporates a feedback loop, it is prone to instability, which deteriorates the noise shaping capabilities. The next section focuses design constraints to ensure stability of the modulator.

### C. Stability of DSM

By construction, the the output of DSM is bounded because constrained by the quantizer's full scale. By stability we refer to the boundedness of signals internal to the DSM. Hereafter, we define two conditions to enforce this property.

Condition 1, also used in [10], [11], forces the quantization error signal $\epsilon_b[n]$ to be bounded. With $L$ be the number of quantization levels, $\Delta$ be the quantization step and $\mathbf{x}$ the sequence in which we stack all the elements $x_b[n]$ for any $b, n \in \mathbb{N}$, it is enounced in the 2-D case as :

*Condition 1 (No overload):* We assume that the initial quantizer input $v_b[0]$ satisfies $2\|v_0[0]\|_{\infty,\text{IQ}} \leq L\Delta$ for any $b \in [\![0, \cdots, B-1]\!]$. If $\Delta\|\mathbf{g}\|_{1,\text{IQ}} + 2\|\mathbf{z}\|_{\infty,\text{IQ}} \leq L\Delta$, then the quantization error $\epsilon_b[n]$ is guaranteed to incur no overload, i.e. $2\|\boldsymbol{\epsilon}\|_{\infty,\text{IQ}} \leq \Delta$.

The second condition is a bound on the maximal value of the NTF [17]. This condition is not a sufficient nor a necessary condition for stability, but is widely used in practice, such as in the popular DSM toolbox [18].

*Condition 2 (Lee):* Let $\gamma$ be a design parameter. We assume that

$$\max_{(\psi,\nu)\in[0,1]^2} \quad |F(\psi,\nu)| \leq \gamma.$$

The choice of $\gamma$ can be determined through extensive simulations.

## III. Problem formulation and solving

This section establishes the design of the NTF through the filter $f$. The approach taken formulates an optimization problem ensuring the resulting DSM satisfies the desired noise shaping behavior as well as some stability conditions.

In the noiseless case, the received signal at user $u$ can be expressed as

$$y_u[n] = \sum_{t=0}^{T-1} \sum_{b=0}^{B-1} h_{u,b}[t] z_b[n-t] + d_u[n], \tag{14}$$

where $d_u[n]$ is a distortion component introduced by the DSM, given by

$$d_u[n] = \sum_{t=0}^{T-1} \sum_{b=0}^{B-1} h_{u,b}[t] (f \star \epsilon)_b[n-t]. \tag{15}$$

On the receiver side, only the data carriers $\mathcal{D}$ require low distortion. We will only work with the 1-D FFT over the time index which leads to

$$\tilde{d}_u[k] = \sum_{n=0}^{N-1} d_u[n] e^{-2j\pi \frac{k}{N} n} \tag{16}$$

for any $u \in [\![0, U-1]\!]$ and $k \in [\![0, N-1]\!]$.

Under the assumption that $\epsilon_b[n]$ is white noise independent of the modulator input $\mathbf{z}$ and of variance $\sigma^2$, after some straightforward calculations omitted for brevity, the energy of the distortion in the frequency domain is given by

$$\mathbb{E}[|\tilde{d}_u[k]|^2] = N\sigma^2 \|\bar{\mathbf{H}}_u[k] \tilde{\mathbf{f}}[k]\|_2^2, \tag{17}$$

with $\tilde{\mathbf{f}}[k] = [\tilde{f}_0[k], \ldots, \tilde{f}_{T_{\text{ant}}-1}[k]]^\mathsf{T}$ where , $\tilde{f}_\ell[k]$ is the 1-D FFT over the time index of $f$, and $\bar{\mathbf{H}}_u[k] \in \mathbb{C}^{B \times T_{\text{ant}}}$ a Hankel matrix expressed as

$$\bar{\mathbf{H}}_u[k] = \begin{bmatrix} \tilde{h}_{u,0}[k] & \tilde{h}_{u,1}[k] & \ldots & \tilde{h}_{u,T_{\text{ant}}-1}[k] \\ \tilde{h}_{u,1}[k] & \tilde{h}_{u,2}[k] & \ldots & \tilde{h}_{u,T_{\text{ant}}}[k] \\ \vdots & \iddots & & \vdots \\ \tilde{h}_{u,B-T_{\text{ant}}}[k] & & \ldots & \tilde{h}_{u,B-1}[k] \\ \vdots & \iddots & & 0 \\ \vdots & & \iddots & \vdots \\ \tilde{h}_{u,B-1}[k] & 0 & \cdots & 0 \end{bmatrix}. \tag{18}$$

### A. Reducing in-band distortion

To communicate reliably, the filter $f$ minimizes the in-band distortion that is largest over the users. This leads to the following optimization problem called $\mathcal{P}1$.

*Problem 1:*

$$\min_{f,\mu_1} \quad \mu_1 \tag{19a}$$

$$\text{s.t.} \sqrt{\sum_{k\in\mathcal{D}} \|\bar{\mathbf{H}}_u[k]\tilde{\mathbf{f}}[k]\|_2^2} \leq \mu_1 \quad \forall u \in [\![0, U-1]\!] \tag{19b}$$

$$f_0[0] = 1, \tag{19c}$$

$$C(f) \leq \gamma. \tag{19d}$$

Equation (19c) forces at least one delay in the feedback loop, ensuring its computability. Equation (19d) is related to one stability constraint (condition 1 or 2 depending on the chosen function $C(\cdot)$).

When Condition 1 is selected as a stability condition, we have

$$C(f) = \|\mathbf{f} - \boldsymbol{\delta}\|_{1,\mathsf{IQ}} \tag{20}$$

and $\gamma = L + 1 - 2\|\mathbf{z}\|_{\infty,\mathsf{IQ}}$ where $\boldsymbol{\delta}$ is the vector deduced from the 2-D Dirac pulse $\delta$ offering the same dimension as $\mathbf{f}$.

When Condition 2 is selected as a stability condition, the semi-infinite constraint needs to be discretized. With uniform sampling of the space, the collection of constraints is given by

$$C_{k_1,k_2}(f) = |F(k_1/N_{\text{ant}}, k_2/N_{\text{time}})| \tag{21}$$

for $(k_1, k_2) \in [\![0, \cdots, N_{\text{ant}}-1]\!] \times [\![0, \cdots, N_{\text{time}}-1]\!]$ where $N_{\text{ant}}$ and $N_{\text{time}}$ are the number of samples in each dimension of the grid.

As both stability constraints are linear, and as (19b) corresponds to a second-order cone constraint, problem $\mathcal{P}1$ is convex and more precisely a second-order cone program which can be solved using standard optimization toolboxes.

### B. Reducing in-band distortion and out-of-band emissions

Besides ensuring reliable communication in-band, it is necessary to mitigate OOB emissions. Focusing on restricting emissions on the adjacent band $\mathcal{A}$, the energy of the NTF on that band is expressed as

$$\begin{aligned} &\int_{\mathcal{A}} \left| \sum_{n=0}^{T_{\text{time}}-1} f_b[n] e^{-2i\pi n\nu} \right|^2 d\nu \\ &= \|\mathbf{f}_b\|_{\mathbf{K}}^2, \end{aligned} \tag{22}$$

where $\mathbf{f}_b = [f_b[0], \ldots, f_b[T_{\text{time}}-1]]^\mathsf{T}$ and the positive definite matrix $\mathbf{K}$ is given by $[\mathbf{K}]_{n_1,n_2} = \int_{\mathcal{A}} e^{-2i\pi(n_2-n_1)\nu} d\nu$. This leads to the following optimization problem called $\mathcal{P}2$.

*Problem 2:*

$$\min_{f,\mu_1,\mu_2} \quad w\mu_1 + (1-w)\mu_2 \tag{23a}$$

$$\text{s.t. constraints } (19b) - (19c) - (19d) \text{ hold}, \tag{23b}$$

$$\|\mathbf{f}_b\|_{\mathbf{K}} \leq \mu_2 \quad \forall b \in [\![0, T_{\text{ant}}-1]\!]. \tag{23c}$$

where the objective is a sum weighted by $w$ to balance between distortion going into user directions within the band or in all directions on adjacent bands. As (23c) corresponds to second-order cone constraint, $\mathcal{P}2$ is also a convex second-order cone program.

## IV. Numerical results

The simulation setup is for a massive MIMO system with $B = 128$ antennas, $U = 4$ users and $N = 4096$ subcarriers per OFDM system. To simulate a 20 MHz scenario with 30 kHz subcarrier spacing, 612 carriers are reserved for data and are surrounded on each side by 27 guard carriers. The channel model uses $T = 3$ paths, corresponding to a delay spread of 66 µs and the Rician factor is 4. The constellation is 64-QAM. Simulations are achieved over 100 realizations of the channel. Other parameters are specified in each simulation case. When using Lee's stability condition, simulations have shown that $N_{\text{ant}} = 5T_{\text{ant}}$ and $N_{\text{time}} = 5T_{\text{time}}$ ensure sufficient sampling of the criterion. The signal to noise ratio (SNR) is defined as $E_t/N_0$ with $N_0$ the noise level of the additive noise. Error vector magnitude (EVM) calculations are performed without the additive noise. Six scenarios will be considered :

- "1 bit quantizer" : 1 bit quantizer without any noise shaping.
- "1 bit spatial" : 1 bit DSM optimized with $\mathcal{P}1$, but the error feedback filter has $(T_{\text{ant}}, T_{\text{time}}) = (10, 1)$. No frequency-domain noise shaping is possible.
- "1 bit spatio-temporal $\mathcal{P}1$" : 1 bit DSM optimized with $\mathcal{P}1$ at $(T_{\text{ant}}, T_{\text{time}}) = (10, 10)$. Here OOB emissions are disregarded. Only in-band distortion is relevant.
- "1 bit spatio-temporal $\mathcal{P}2$" : 1 bit DSM optimized with $\mathcal{P}2$ at $(T_{\text{ant}}, T_{\text{time}}) = (15, 15)$. OOB emissions are taken into consideration.
- "2 bit spatio-temporal $\mathcal{P}1$" : Same as "1 bit spatio-temporal $\mathcal{P}1$", except with 2 bits.
- "No quantizer" : Ideal transmitter without quantization.

### A. Maximum stable input range

The impact of the stability condition is analyzed using scenario "Spatio-temporal $\mathcal{P}1$". As the stability of the modulator is highly dependant on the input signal's backoff to the DSM's internal quantizer full scale, we proceed to search for the optimal input range. The DSM input $z_b[n]$ corresponds to the output of an IFFT due to OFDM. Consequently, it can be assumed to be Gaussian distributed. Defining the clipping probability $p_{\text{clip}}$ as the probability that $|z_b[n]|$ is larger than the full scale of the quantizer $(L-1)\Delta$, it relates to it by $(L-1)\Delta = 2\sqrt{2E_t}\text{erfc}^{-1}(p_{\text{clip}})$ where erfc is the complementary error function.

In Fig. 4, we plot the bit error rate (BER) versus $p_{\text{clip}}$ at medium SNR (3 dB) for both stability conditions coming from (20) and (21). With low clipping probability, BER is high because the signal power is small compared to the introduced quantization noise power. At large $p_{\text{clip}}$, the modulator faces instability. Overall, condition 2 with $\gamma = 1.7$ performs best and is retained for the remaining simulations. Similar conclusions were reached at low and high SNR.

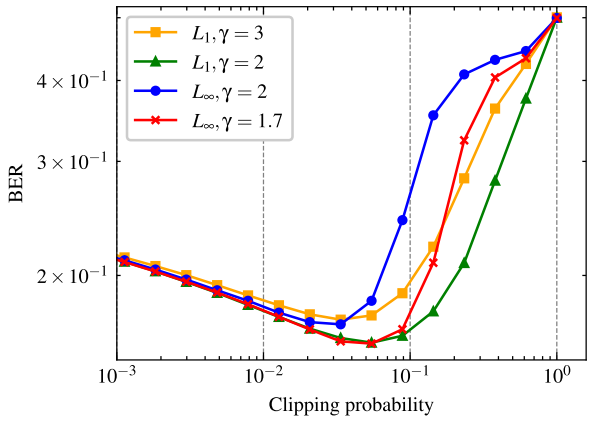


Fig. 4. EVM versus $p_{\text{clip}}$ for different stability conditions and $\gamma$.

### B. Discussion on the design of problem $\mathcal{P}2$

We analyze the impact of the weight $w$ in the objective function under the scenario "Spatio-temporal $\mathcal{P}2$". Fig. 5 illustrates the adjacent channel leakage ratio (ACLR) and EVM tradeoff existing between different choices of $w$. To determine

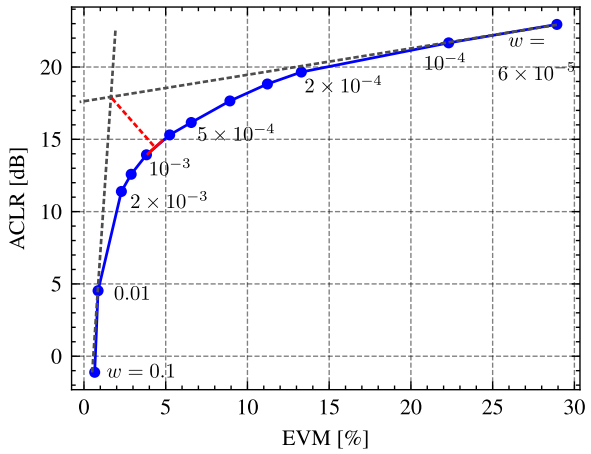


Fig. 5. ACLR versus EVM for multiple values of $w$.

the chosen $w$, the distance to the intersection between the asymptotes in EVM $\to 0\%$ and EVM $\to \infty\%$ is taken. The resulting point, $w = 10^{-3}$, offers an ACLR of 13.9 dB and an EVM of 3.8 %. The target of an ACLR of 45 dB as advocated by 3GPP [9] is not achieved. Choosing smaller values of $w$ in hopes of improving ACLR causes significant EVM degradation. To overcome this issue, temporal oversampling and higher resolution DACs can be an option.

### C. Performance analysis

In Fig. 6, the NTFs for three scenarios (spatial, $\mathcal{P}1$, $\mathcal{P}2$) are plotted versus the angles and the frequency bins. The adjacent bands and users are also overlaid on top. Without

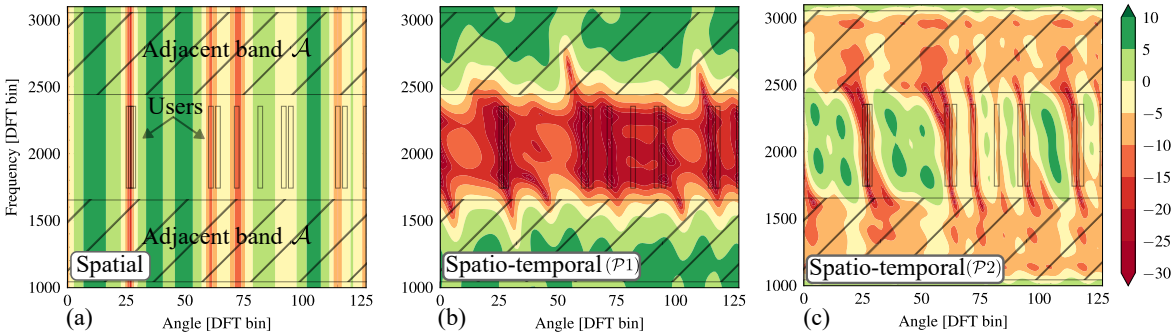


Fig. 6. NTF level versus angles and frequency bins for scenarios (a) spatial, (b) $\mathcal{P}1$, and (c) $\mathcal{P}2$.

any constraint on OOB emissions, $\mathcal{P}1$ enables deeper notches on the user locations, and thus better in-band performance than $\mathcal{P}2$. The spatial-only NTF has no capability of providing different noise shaping over the frequency axis.

In Fig. 7, the BER curve is plotted for the six scenarios. The spatio-temporal $\mathcal{P}1$ scenario outperforms its spatial counterpart. When taking into account OOB emissions, as done with $\mathcal{P}2$, the in-band noise shaping is less sharp and a slight BER degradation is observed. Finally, raising the resolution to $2\,\text{bits}$ enables the proposed method to come within $2\,\text{dB}$ of the perfect case.

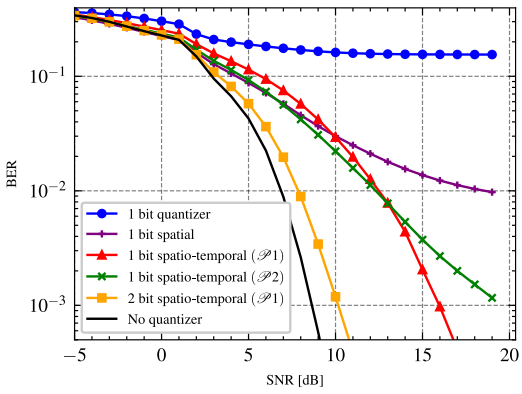


Fig. 7. BER performance between the various scenarios.

## V. Conclusion

In this work, a spatio-temporal DSM was optimized in the context of a massive MIMO-OFDM system. A flexible method balancing between in-band and OOB quantization noise mitigation was proposed. With large antenna arrays, simulation results at low resolution show promising BER performance. However, when it comes to reducing OOB emissions, further investment is required to reach typical targets set by communication standards. Future works may study the impact of time-domain oversampling or higher resolution modulators in hopes of improving ACLR performance.

## References


[1] H. Halbauer, A. Weber, D. Wiegner, and T. Wild, "Energy efficient massive mimo array configurations," in *IEEE Globecom Workshops*, 2018, pp. 1–6.

[2] O. Morales Chacón, J. J. Wikner, C. Svensson, L. Siek, and A. Alvandpour, "Analysis of energy consumption bounds in cmos current-steering digital-to-analog converters," *Analog Integr. Circuits Signal Process.*, vol. 111, no. 3, pp. 339–351, Jun 2022.

[3] S. Jacobsson, G. Durisi, M. Coldrey, and C. Studer, "Linear precoding with low-resolution dacs for massive mu-mimo-ofdm downlink," *IEEE Trans. Commun.*, vol. 18, no. 3, pp. 1595–1609, 2019.

[4] S. Jacobsson, G. Durisi, M. Coldrey, T. Goldstein, and C. Studer, "Quantized Precoding for Massive MU-MIMO," *IEEE Trans. Commun.*, vol. 65, no. 11, pp. 4670–4684, 2017.

[5] A. Mezghani and R. W. Heath, "Massive mimo precoding and spectral shaping with low resolution phase-only dacs and active constellation extension," *IEEE Trans. Wireless Commun.*, vol. 21, no. 7, pp. 5265–5278, 2022.

[6] S. Jacobsson, O. Castañeda, C. Jeon, G. Durisi, and C. Studer, "Nonlinear precoding for phase-quantized constant-envelope massive mu-mimo-ofdm," in *Proc. IEEE Int. Conf. Telecomun.*, 2018, pp. 367–372.

[7] Y. Karabacakoğlu, A. B. Üçüncü, and G. M. Güvensen, "An iterative distortion-aware precoding for quantized upsampled wideband massive mimo," in *Proc. IEEE Int. Commun. Conf.*, 2023, pp. 6498–6503.

[8] S. Jacobsson, G. Durisi, M. Coldrey, and C. Studer, "On out-of-band emissions of quantized precoding in massive mu-mimo-ofdm," in *Proc. IEEE Asilomar Conf. Signals, Syst., Comput.*, 2017, pp. 21–26.

[9] 3GPP, "3rd Generation Partnership Project; 5G; NR; Base Station (BS) radio transmission and reception," 3GPP, Tech. Rep. TS 38.104 version 15.14.0 Release 15, 2021.

[10] W.-Y. Keung and W.-K. Ma, "Spatial sigma-delta modulation for coarsely quantized massive mimo downlink: Flexible designs by convex optimization," *IEEE Open J. Signal Process.*, vol. 5, pp. 520–538, 2024.

[11] ——, "Spatial sigma-delta modulation for few-bit mimo precoding: Quantization error suppression according to channel state information," in *Proc. European Signal Process. Conf.*, 2024, pp. 2072–2076.

[12] M. Shao, W.-K. Ma, Q. Li, and A. L. Swindlehurst, "One-bit sigma-delta mimo precoding," *IEEE J. Sel. Areas Commun.*, vol. 13, no. 5, pp. 1046–1061, 2019.

[13] D. Scholnik and J. Coleman, "Joint spatial and temporal delta-sigma modulation for wideband antenna arrays and video halftoning," vol. 5, 2001, pp. 2941–2944 vol.5.

[14] ——, "Space-time vector delta-sigma modulation," in *Proc. Int. Symp. Circuits Syst. (ISCAS)*, vol. 3, 2002, pp. III–III.

[15] D. Scholnik, J. Coleman, D. Bowling, and M. Neel, "Spatio-temporal delta-sigma modulation for shared wideband transmit arrays," in *Proc. IEEE Radar Conf.*, 2004, pp. 85–90.

[16] R. Schreier and G. C. Temes, *The First Order Delta Sigma Modulator*, 2005, pp. 21–62.

[17] L. W. L., "A novel higher order interpolative modulator topology for high resolution oversampling a/d converters," *Master's Thesis, Massachusetts Institute of technology*, 1987.

[18] R. Schreier, "Delta sigma toolbox," https://www.mathworks.com/matlabcentral/fileexchange/19-delta-sigma-toolbox, 2025, MATLAB Central File Exchange. Retrieved March 31, 2025.