Robust 3D MIMO-OFDM Channel Estimation with Hybrid Analog-Digital Architecture

Giuseppe Destino, Jani Saloranta, Markku Juntti
Centre for Wireless Communications
University of Oulu,
Oulu, Finland
destino@ee.oulu.fi, jsaloran@ee.oulu.fi and juntti@ee.oulu.fi

Shirish Nagaraj
Next Generation and Standards, Commun. Devices Group,
Intel,
Santa Clara, CA., USA,
shirish.nagaraj@intel.com

***Abstract*—We consider the problem of 3D multiple-input-multiple-output (MIMO) orthogonal frequency-division multiplexing (OFDM) channel estimation, a key for the future development of 3D beamforming techniques. Our main contribution is a novel algorithm, namely the *adaptive*-LASSO, that can jointly exploit the sparsity structure of the MIMO-OFDM channel in the spatial and delay domains. The algorithm is designed to handle large antenna arrays by means of a hybrid analog-digital architecture. In this regard, we describe an effective beam-switching strategy to sample the channel using a few analog beamformers. We investigate the impact of the signal bandwidth, antenna structures, line-of-sight (LOS) and non-line-of-sight (NLOS) conditions via ray-tracing based simulations. Also, we show that the A-LASSO can provide significant improvements with respect to the legacy methods, $e.g.$ least-square technique.**

I. Introduction

The potentials of three dimensional (3D) multiple-input-multiple-output (MIMO) technology to provide high user capacity, better spectral efficiency, less intercell interference and improved coverage are well known [1]. A key is the usage of large-scale antenna arrays for 3D beamforming, $i.e.$, adapt the radiation beam pattern in both elevation and azimuth to provide more degrees of freedom in supporting users. However, as the number of antenna elements grows, the cost of a full digital approach can be prohibitive. For this reason, an hybrid analog-digital architecture is often considered [2].

In addition to radio frequency (RF) implementation challenges, 3D MIMO quests for new channel estimation and beamforming algorithms [3]. In this paper, the focus is on the development of a robust 3D MIMO orthogonal frequency-division multiplexing (OFDM) channel estimation method constrained to the hybrid transceiver architecture shown in Figure 1 [2]. More specifically, we propose the *adaptive*-least absolute shrinkage and selection operator (A-LASSO) algorithm and a beam switching technique to achieve fast sensing with few analog beamformers. Key of the algorithm is the utilization of an ℓ_1 based optimization to exploit the sparsity property of the channel [4]–[6], enhanced with the utilization of large-scale antennas and high carrier frequencies.

The rest of the paper is organized as follows. In Section II, we present the system model and the main assumptions. In Section III, we describe the key contributions. In Sections IV results based on ray tracing simulations are provided, and, finally, in Section V concluding remarks are given.

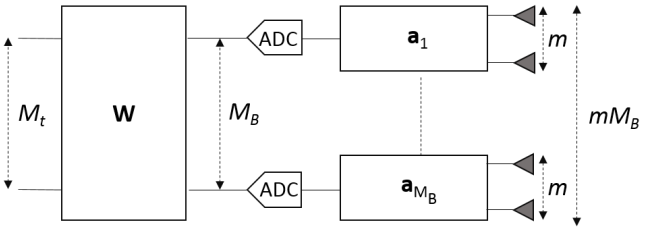


Figure 1. Analog-digital architecture of the receiver

II. System Model

We consider a MIMO-OFDM system in which B is the maximum signal bandwidth, N_{FFT} is the total number of subcarriers, M and one is the number of receiving and transmitting antennas, respectively. At the receiver, we assume the hybrid analog-digital architecture depicted in Figure 1 [2], where M_B is the number of independent digital paths and each digital path consists of an analog-to-digital converter (ADC), amplifiers, an analog beamformer and m antenna elements. Analog beamformers are implemented with phase-shifters and a phase-shifter can control only one antenna element.

Mathematically, the aforementioned receiver architecture can be modeled as follows. Let $\mathbf{a}_i \in \mathbb{C}^m$ be the analog beamformer used in the ith digital path and let $a_{ji} \in \mathbb{C}$, with $|a_{ji}| = 1$, be the jth element of $\mathbf{a}_i$. We denote by $\mathbf{A} \in \mathbb{C}^{M \times M_B}$ the "total" analog beamformer matrix, that is block-diagonal with the ith block given by $\mathbf{a}_i$. The digital beamformer matrix is given by $\mathbf{W} \in \mathbb{C}^{M_B \times M_t}$, where M_t is the number of combining digital beamformers.

The MIMO-OFDM channel is modeled as a frequency-selective wideband clustered channel, in which N_c is the number of clusters and N_q is the number of propagation paths per cluster [4]. For the nth subcarrier, the channel is given by

$$\mathbf{h}_n = \sqrt{\frac{M}{N_c N_q N_{\text{FFT}}}} \sum_{c=1}^{N_c} \sum_{q=1}^{N_q} \alpha_{q,c} \mathbf{a}_R(\phi_{q,c}, \theta_{q,c}) e^{-j2\pi n \frac{\tau_{q,c} B}{N_{\text{FFT}}}}, \tag{1}$$

where the index $(\cdot)_{q,c}$ refers to p-th path of the c-th cluster, $\phi_{q,c}, \theta_{q,c}$ are the azimuth and elevation of the angle of arrival, $\tau_{q,c}$ is path-delay, $\alpha_{p,c} \sim \mathcal{CN}(0, \sigma^2_{p,c})$ is a complex Gaussian

978-0-9928-6265-7/16/$31.00 ©2016 IEEE

random variable modeling the complex gain of the path, $\sigma^2_{p,c}$ is the energy path-gain and $\mathbf{a}_R(x,y)$ is given by

$$\mathbf{a}_R(x,y) = e^{-j\frac{2\pi}{\lambda}\mathbf{k}^{\mathrm{T}}(x,y)\mathbf{P}}, \tag{2}$$

where $(\cdot)^{\mathrm{T}}$ indicates transpose, λ is the carrier wavelength, $\mathbf{P} \in \mathbb{R}^{3\times M}$ is the matrix containing the (3D) location of antenna elements[1] and $\mathbf{k}(x,y) \triangleq [\cos(x)\cos(y), \sin(x)\cos(y), \sin(y)]^{\mathrm{T}}$.

Assume that N subcarriers are used for a MIMO-OFDM transmissions, then the received signal can be obtained as

$$\mathbf{Y} = \sqrt{\rho}\mathbf{W}^{\mathrm{H}}\mathbf{A}^{\mathrm{H}}(\mathbf{H}\mathbf{X} + \mathbf{N}), \tag{3}$$

where ρ is the average received power, $\mathbf{H} \triangleq [\mathbf{h}_1, \cdots, \mathbf{h}_N]$, $\mathbf{X} \in \mathbb{C}^{N\times N}$ is a diagonal matrix in which the diagonal element $X_{ii} \in \mathbb{C}$ is the transmitted symbol over the ith subcarrier and $\mathbf{N} \in \mathbb{C}^{M_t\times N}$ is the noise matrix with $n_{ij} \sim \mathcal{CN}(0,1)$ as a complex-Gaussian random variable.

III. Estimation of the MIMO-OFDM Channel with A-LASSO

We focus on the estimation problem of the MIMO-OFDM channel given that $\mathbf{X}$ is known at the receiver and $\mathbf{H}$ has a sparse representation [4]–[6], *i.e*, the angle of arrivals and path-delays are few, distinct and within finite intervals. Let $\mathbf{\Psi} \in \mathbb{C}^{N_{\mathrm{FFT}}M\times L}$ be a dictionary matrix, that is a matrix used to represent the channel $\mathbf{h} \triangleq \mathrm{vec}(\mathbf{H})$ as a linear combination of L spatial-Fourier-frequencies [7], *i.e*., $\mathbf{h} = \mathbf{\Psi}\mathbf{z}$ where $\mathbf{z} \in \mathbb{C}^L$ is referred to as the representation of $\mathbf{h}$ in $\mathbf{\Psi}$.

Relying on the sparsity structure of the channel, we formulate a robust channel estimator using an ℓ_1-model, *e.g*., the least absolute shrinkage and selection operator (LASSO),

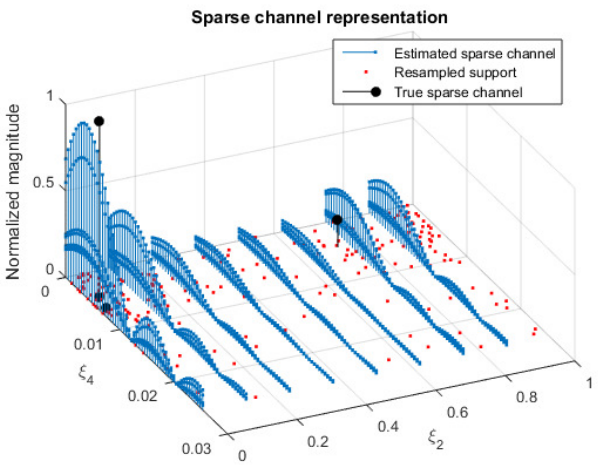
$$\hat{\mathbf{z}} = \min_{\mathbf{z}\in\mathbb{C}^L} \nu\|\mathbf{z}\|_1 + \tfrac{1}{2}\|\mathbf{y} - (\mathbf{X}\mathbf{S}_f^{\mathrm{T}} \otimes \mathbf{W}^{\mathrm{H}}\mathbf{A}^{\mathrm{H}})\mathbf{\Psi}\mathbf{z}\|_2^2, \tag{4}$$

where $\|\cdot\|_q$ is the q-norm, $\otimes$ denotes the Kronecker product, $\mathbf{y} \in \mathbb{C}^{M_t N}$ is the vector form of $\mathbf{Y}$, ν is a parameter that controls the sparsity ($\approx 1/\sqrt{N_{\mathrm{FFT}}}$, due to the fact that the sparsity regularizer should be proportional to the noise spectral density) and $\mathbf{S}_f \in \mathbb{R}^{N_{\mathrm{FFT}}\times N}$ is the "frequency selection" matrix[2] used to select the pilot frequency from the set of N_{FFT} subcarriers.

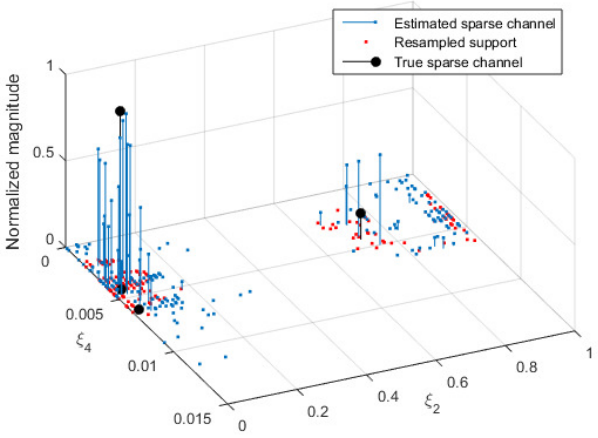
Given a dictionary $\mathbf{\Psi}$, equation (4) can be efficiently solved with numerous algorithms available in the literature, *e.g*. the alternating direction method of multipliers (ADMM) [8]. However, in order to enhance sparsity and minimize the number of observations (the number of elements in $\mathbf{y}$), it is well-known that both and $\mathbf{\Psi}$ and $\mathbf{\Phi} = \mathbf{X}\mathbf{S}_f^{\mathrm{T}} \otimes \mathbf{W}^{\mathrm{H}}\mathbf{A}^{\mathrm{H}}$ need to be optimized. For these reasons, new research is focusing on the development of joint optimization methods of the sparse variable $\mathbf{z}$ and the dictionary-sampling matrix $\mathbf{\Phi}\mathbf{\Psi}$ [9], [10]. We propose a novel algorithm, namely the A-LASSO technique that, in contrast to the formulation in [11], is specialized to the case of MIMO-OFDM channel estimation.

[1]Coordinates are relative respect to the location of one element.

[2]The n-th column of $\mathbf{S}_f$ has only one non-zero element, *i.e*., 1, at the row-index corresponding to the n-th subcarrier index.



(a) A-LASSO step-1



(b) A-LASSO step-3

Figure 2. Progress of the dictionary optimization. Black and blue indicators refer to the true (channel rays) and estimated sparse channel representations. Red dots indicates the new support (atoms) obtained from the re-sampling.

A. A-LASSO Algorithm

The A-LASSO solves the optimization problem

$$\begin{aligned}(\hat{\mathbf{z}}, \hat{\mathbf{\Xi}}) = \min_{\substack{\mathbf{z}\in\mathbb{C}^L \\ \mathbf{\Xi}\in\mathbb{R}^{U\times L}}} & \lambda\|\mathbf{z}\|_1 + \tfrac{1}{2}\|\mathbf{y} - (\mathbf{X}\mathbf{S}_f^{\mathrm{T}} \otimes \mathbf{W}^{\mathrm{H}}\mathbf{A}^{\mathrm{H}})\mathbf{\Psi}\mathbf{z}\|_2^2, \\ \text{s.t.} \quad & \mathbf{\Psi} = \mathcal{D}(\mathbf{\Xi}),\ \underline{\xi_i} \le \xi_{ij} < \bar{\xi}_i,\ \forall ij,\end{aligned} \tag{5}$$

where $\mathcal{D}(\mathbf{\Xi})$ is the "dictionary function" defined as

$$\mathcal{D}(\mathbf{\Xi}) \triangleq \left[\bigotimes_{i=1}^{U}\mathbf{v}(\xi_{i1}, K_i), \cdots, \bigotimes_{i=1}^{U}\mathbf{v}(\xi_{iL}, K_i)\right], \tag{6}$$

with $\bigotimes$ indicating the "total" Kronecker product of U vectors, ξ_{ij} is the ij-th element of dictionary variable $\mathbf{\Xi}$, $\underline{\xi_i}$ and $\bar{\xi}_i$ are the minimum and maximum value of $\xi_{ij}\forall j$, and $\mathbf{v}(x,y)$ is

$$\mathbf{v}(x,K) = \left[1, \cdots, e^{-j2\pi x(K-1)}\right]^{\mathrm{T}}, \tag{7}$$

with $x \in [0,1)$ and $K \in \mathbb{N}$.

More specifically, the dictionary represent a multidimensional Fourier transform in which the columns are not necessarily orthogonal. In fact, each column of the dictionary, hereafter referred to as *atom*, can be uniquely associated to a multi-dimensional Fourier frequency of a channel path characterized by a specific delay, azimuth and elevation angle-of-arrival (AoA) and angle-of-departure (AoD). The number of atoms is given by L, which can be arbitrarily selected to provide a sufficient sampling of the domain of $\boldsymbol{\Xi}$. The parameters U and K are specific to the choice of the antenna structure (both at the transmitter and receiver) and the total number of subcarriers. For instance, in [11], where a MIMO transmission with uniform linear array (ULA) antennas both at the transmitter and receiver was studied, U was equal to 2. In this work, we set $U = 3$: two dimensions to describe the spatial-frequency and one for the Fourier frequency. Finally, if we consider a uniform rectangular array (URA) receiving antenna with $M_y \times M_z$ elements deployed in the yz-plane, then spatial and Fourier frequencies can be defined with vectors of $K_1{=}M_y$, $K_2{=}M_z$ and $K_3{=}N_{\text{FFT}}$ elements, respectively.

The A-LASSO algorithm consists of alternating the minimization of (5) for a fixed $\hat{\boldsymbol{\Psi}}$ with a dictionary optimization until the objective function remains almost invariant (typically three iterations). More specifically, the dictionary is updated based on the randomization method described in [11]. The key is the heuristic that the higher the value of $|z_i|$, the higher is the "importance" of the i-th atom in the dictionary. In fact, the update method is to re-sample the atoms proportionally to their weight and randomize their corresponding parameters $\boldsymbol{\xi}$. A careful implementation of the latter step is needed to avoid atom repetition as well as very close atoms. Our approach is to randomly select the new atoms over a fine grid such that the objective function of (5) decreases.

Figure 2 shows an example of the aforementioned optimization process. The result is obtained with SNR$=$ 10dB, $N_{\text{FFT}} = 2048$, $B = 200$MHz, $M_B = 8$ and $M = 64$. The antenna structure is URA with $M_y = 8$ and $M_z = 8$ and, each digital path controls 8 elements in the z-direction. The black indicators refer to the true channel, whereas the blue ones are obtained from the A-LASSO algorithm. The red dots indicate the new atoms[3] computed with the resampling. Figure 2(a) and 2(b) refer to the first and the third iteration of the A-LASSO optimization. It can be noticed how the support of the channel estimate concentrates around the black indicators.

B. Hybrid Beamforming

Let us focus on the design of $\boldsymbol{\Phi}$ and in particular on the hybrid beamforming $\bar{\boldsymbol{\Phi}} \triangleq \mathbf{W}^{\text{H}}\mathbf{A}^{\text{H}}$. As suggested in [12], the optimal projection matrix $\boldsymbol{\Phi}$ is the matrix that yields

$$\boldsymbol{\Psi}^{\text{H}}\boldsymbol{\Phi}^{\text{H}}\boldsymbol{\Phi}\boldsymbol{\Psi} \approx \mathbf{I}_L, \tag{8}$$

which in our case implies

$$\boldsymbol{\Psi}^{\text{H}}\left(\mathbf{X}\mathbf{S}^{\text{T}}\otimes\bar{\boldsymbol{\Phi}}\right)^{\text{H}}\left(\mathbf{X}\mathbf{S}^{\text{T}}\otimes\bar{\boldsymbol{\Phi}}\right)\boldsymbol{\Psi} = \boldsymbol{\Psi}^{\text{H}}\left(\mathbf{I}\otimes\bar{\boldsymbol{\Phi}}^{\text{H}}\bar{\boldsymbol{\Phi}}\right)\boldsymbol{\Psi} \approx \mathbf{I}_L, \tag{9}$$

[3]Notice that the axis corresponding to $\xi_{3:}$ is not plotted for the sake of clarity. However, they can be recognized with the fact that multiple bars are overlapped are the same location.

in which we have made the assumptions that $|X_{ii}| = 1$, $\forall i$.

This is not a trivial problem as, in the A-LASSO, $\boldsymbol{\Psi}$ is also a variable. However, if the paths are sufficiently isolated and atoms are not highly correlated, it is sensible to assume $\boldsymbol{\Psi}^{\text{H}}\boldsymbol{\Psi} \approx \mathbf{I}$. Hence, (9) can be reduced to

$$\bar{\boldsymbol{\Phi}}^{\text{H}}\bar{\boldsymbol{\Phi}} = \mathbf{A}\mathbf{W}\mathbf{W}^{\text{H}}\mathbf{A}^{\text{H}} \approx \mathbf{I}_M. \tag{10}$$

Next, we write the explicit expression of $\bar{\boldsymbol{\Phi}}^{\text{H}}\bar{\boldsymbol{\Phi}}$ using the block-diagonal structure of $\mathbf{A}$. We obtain

$$\bar{\boldsymbol{\Phi}}^{\text{H}}\bar{\boldsymbol{\Phi}} = \begin{bmatrix} \sum_{i=1}^{M_t} |w_{1i}|^2 \mathbf{a}_1\mathbf{a}_1^{\text{H}} & \cdot & \sum_{i=1}^{M_t} |w_{1i}w_{M_Bi}^*| \mathbf{a}_1\mathbf{a}_{M_B}^{\text{H}} \\ \vdots & \vdots & \vdots \\ \sum_{i=1}^{M_t} |w_{M_Bi}w_{1i}^*| \mathbf{a}_{M_B}\mathbf{a}_1^{\text{H}} & \cdot & \sum_{i=1}^{M_t} |w_{M_Bi}|^2 \mathbf{a}_{M_B}\mathbf{a}_{M_B}^{\text{H}} \end{bmatrix}, \tag{11}$$

where w_{ij} is the ij-th element of $\mathbf{W}$ and $(\cdot)^*$ indicates the complex-conjugate.

We notice that if $\mathbf{W}$ is a diagonal matrix, then $\bar{\boldsymbol{\Phi}}^{\text{H}}\bar{\boldsymbol{\Phi}}$ is block-diagonal. Also, we can recognize that the elements of the i-th diagonal block have a constant magnitude as the analog beamformers are implemented with phase-shifters only.

In the light of the above, it appears very challenging to find a $\bar{\boldsymbol{\Phi}}$ such that $\bar{\boldsymbol{\Phi}}^{\text{H}}\bar{\boldsymbol{\Phi}}$ is almost an identity matrix. In order to circumvent this problem, we propose a modification of the structure of $\mathbf{A}$. More specifically, we consider

$$\bar{\mathbf{A}} = \begin{bmatrix} \mathbf{A}_0 & \mathbf{A}_2 & \cdots & \mathbf{A}_{M_B-1} \end{bmatrix}, \tag{12}$$

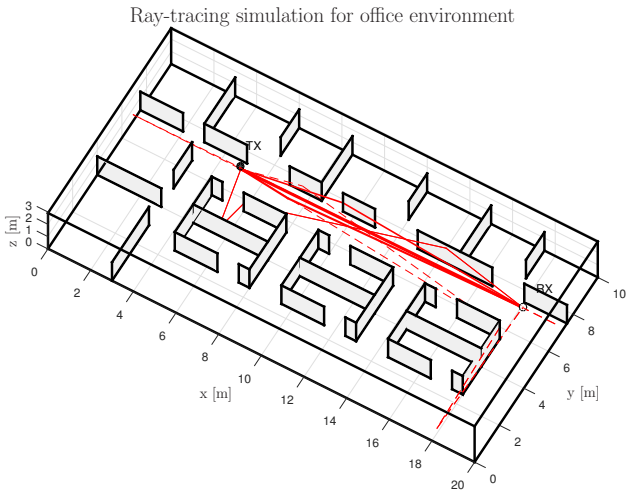
where $\mathbf{A}_k \in \mathbb{C}^{M\times M_B}$ as a block-diagonal matrix with the j-th diagonal block given by $\mathbf{a}_j^k$, where the k indicates a cycle-shift of the index j by k positions, *e.g.* $\mathbf{a}_4^2 = \mathbf{a}_2$.

This structure corresponds to a measurement system with beam switching, *i.e.*, the analog beamformers are changed during the measurement phase. In doing so, the i-th diagonal block of $\bar{\boldsymbol{\Phi}}^{\text{H}}\bar{\boldsymbol{\Phi}}$ is given by $\sum_{j=1}^{M_B} |w_{ii}|^2 \mathbf{a}_j\mathbf{a}_j^{\text{H}}$, in which the kq-th off-diagonal term is no longer constant but a linear combination of $a_{ik}a_{iq}^*$, $\forall i$. Thus, if M_B is sufficiently large and $|\mathbf{a}_i^{\text{H}}\mathbf{a}_j|$ is small, $\forall ij$ with $i \neq j$, then $\bar{\boldsymbol{\Phi}}^{\text{H}}\bar{\boldsymbol{\Phi}} \approx \mathbf{I}_M$.

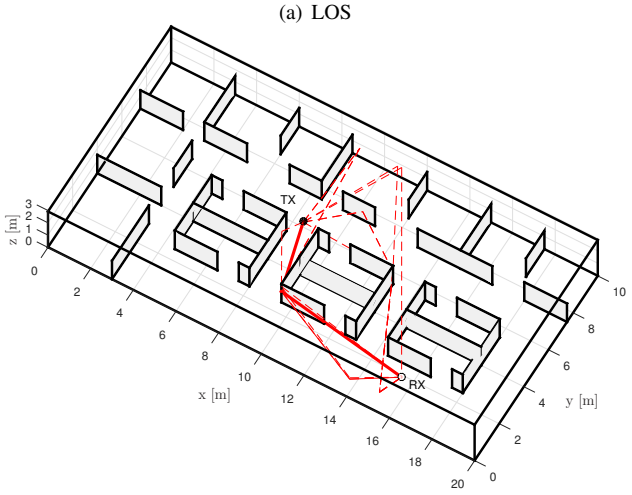
IV. Simulation Results

In this section, we evaluate the performance of the proposed algorithm via ray-tracing based simulations. We consider an office environment and study both LOS (Figure 3(a)) and NLOS (Figure 3(b)) links. The pilot signal is transmitted at the carrier frequency 28GHz, with a bandwidth $B_p \leq B = 200$ MHz. The OFDM symbol is designed with 2048 subcarriers, *i.e.*, $N_{\text{FFT}} = 2048$, of which only N are used for pilot. The antenna at the receiver is modeled with a URA in the yz-plane, with M_z rows of 8 elements deployed in the y-direction. Antenna elements are equispaced by $\lambda/2$ both in the vertical (z-axis) and the horizontal (y-axis) directions. The number of digital paths is $M_B = 8$, and each one controls a column (sub-array in the z-direction) of the URA. Thus, $m = M_z$.

The receiver antenna covers a 3D sector defined with (azimuth) $\phi \in [-\pi/4, \pi/4]$ and (elevation) $\theta \in [-\pi/4, \pi/4]$.



(a) LOS



(b) NLOS

Figure 3. Simulation scenario in office environment

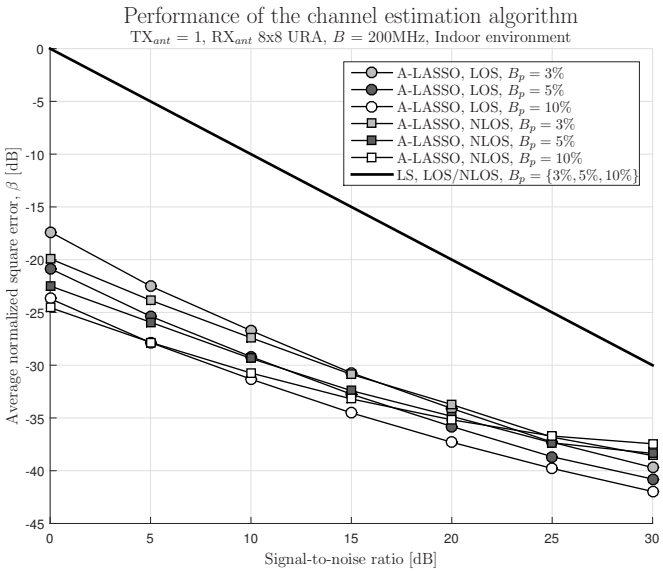
At the transmitter, instead, we assume the capability of transmitting a signal only in half-sphere[4]. With respect to the scenario depicted in Figure 3, the receiver's sector is always facing to the yz-plane with $x = 0$, whereas the transmitter if facing to the yz-plane with $x = 20$.

Due to this sectorization, the effective number of rays is smaller than that one generated by the ray-tracing simulator. For instance, in Figures 3(a) and 3(b), the dashed lines indicate the removed rays. Finally, notice that in NLOS conditions the number of paths is two (effect of diffraction and blockage), whereas in LOS is four (effect of the multipath propagation).

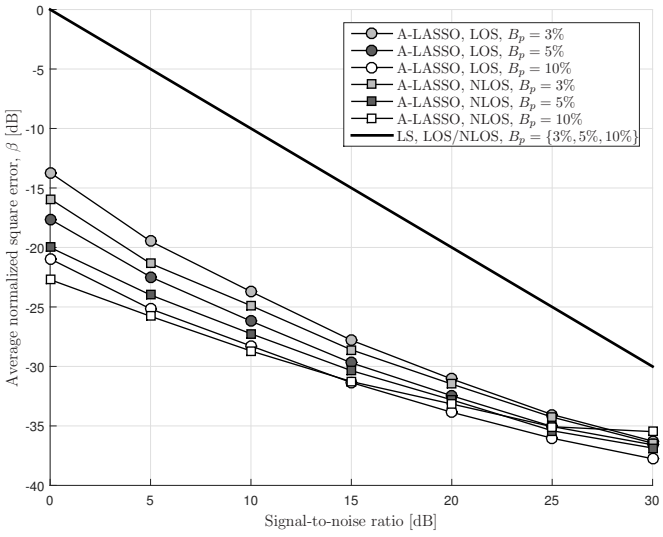
The objectives of our study are: i) evaluate the performance of the proposed A-LASSO method as a function of the Signal-to-Noise Ratio (SNR) using different bandwidth B_p and URA structures; ii) investigate the impact of beam switching by considering a full and half rotation cycle; iii) compare the performance between LOS and NLOS conditions and, iv) compare the A-LASSO technique with a Least Squares (LS) estimator[5] (legacy method).

[4]This level of directionality can be obtained with a back-reflector

[5]$\hat{\mathbf{H}}_{\text{LS}} = (\mathbf{W}^{\text{H}}\mathbf{A}^{\text{H}})^{\dagger}[\mathbf{y}_1 X_{11}^{-1}, \cdots, \mathbf{y}_N X_{NN}^{-1}]$, where †is the pseudoinverse.



(a) Full rotation cycle



(b) Half rotation cycle

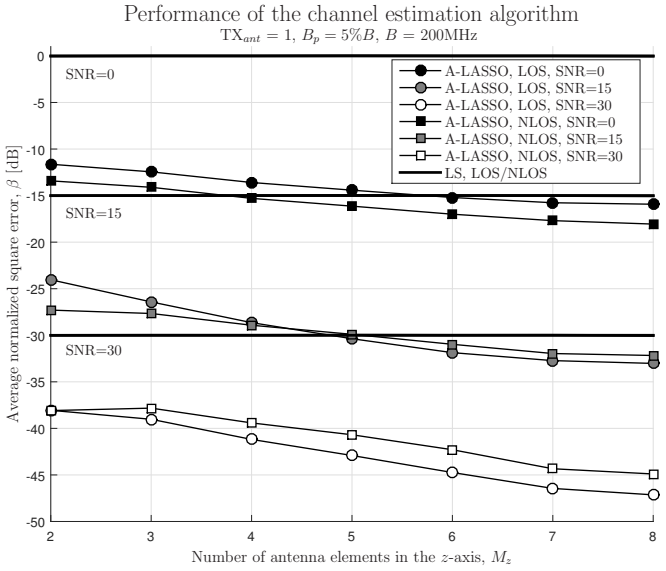
Figure 4. Performance as a function of the SNR and for different channel bandwidth

The performance metric is the average *normalized* squared error (in dB) given by

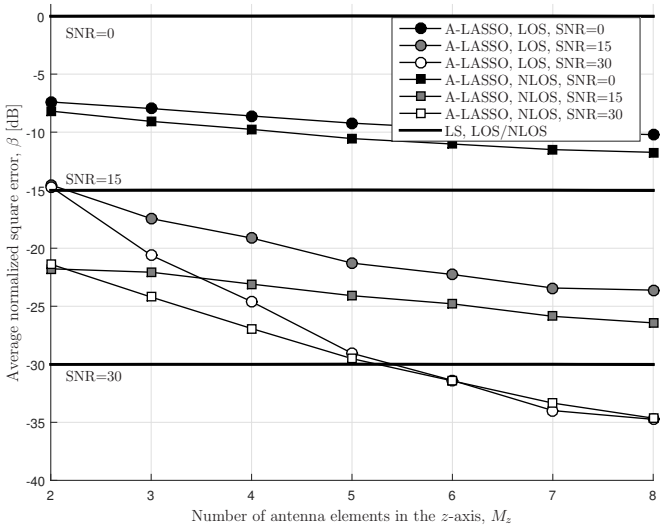
$$\beta = 10 \log \left(\frac{1}{T} \sum_{t=1}^{T} \left\{ \frac{\|\text{vec}(\mathbf{H}^{(t)}) - \hat{\mathbf{h}}^{(t)}\|_2^2}{\|\text{vec}(\mathbf{H}^{(t)})\|_2^2} \right\} \right), \qquad (13)$$

where T is the number of simulations, $\hat{\mathbf{h}} \triangleq \hat{\mathbf{\Psi}}\hat{\mathbf{z}}$ and $\hat{\mathbf{\Psi}} = \mathcal{D}(\hat{\mathbf{\Xi}})$.

Figure 4 shows the estimation error as function of the SNR for both LOS and NLOS scenarios as well as for different values of B_p (expressed as a percentage of B) and with two measurement settings, full cycle (Figure 4(a)) and half cycle (Figure 4(b)). The results are obtained with a receiving antenna URA with $M_z = 8$, thus $M = 64$. Generally, it is shown that a lower estimation error can be achieved with the A-LASSO algorithm. What is more, the gap between the A-LASSO and LS is more prominent in the low SNR régime, since the A-LASSO can remove noise components by optimizing the dictionary.



(a) Full rotation cycle



(b) Half rotation cycle

Figure 5. Performance as a function of the antenna structure.

By comparing the results between LOS and NLOS channel conditions, we notice that the estimation error is similar, although slightly worse in LOS case. This is a consequence of the simulation setting, in which the LOS link has a higher number of paths.

Finally, by comparing Figures 4(a) and 4(b), the error is larger with the half rotation cycle. This is a typical effect in compressive sensing and it indicates that the number of beam switching is not enough to yield $\bar{\mathbf{\Phi}}^{\mathrm{H}}\bar{\mathbf{\Phi}} \approx \mathbf{I}$.

Figure 5 shows the performance of the A-LASSO as a function of the antenna structure. We vary M_z from two to eight, thus gradually allowing higher resolution for the elevation. Likewise the previous set of results, the estimation error is computed with different SNR as well as for LOS and NLOS channel conditions and with a complete and a reduced measurement set. The pilot bandwidth $B_p = 5\%B$. Unlike the LS, the estimation error achieved with the A-LASSO algorithm improves with the increase of M_z. In fact, the A-LASSO algorithm can exploit a higher elevation resolution to separate paths as well as to remove noise. The results in Figure 5(b) show, however, a critical value of M_z, before which the A-LASSO leads to a larger estimation error compared to the LS. This is related to the fact that good sparsifying dictionary can not be found, and subsequently the sparse channel representation is not sufficiently sparse.

V. Conclusions

We considered the problem of channel estimation for a 3D MIMO-OFDM system with hybrid analog-digital transceiver architecture. We leveraged channel sparsity, especially enhanced with 3D MIMO as well as with the usage of high carrier frequencies, into the formulation of a ℓ_1-based estimation algorithm, the A-LASSO. Also, we proposed an efficient beam switching strategy to circumvent the limitations caused by the limited number of digital paths. It was found that the performance of the A-LASSO depends on the number of rays forming the channel. With limited numbers of beam sweeps, the estimation error can increase when the antenna is not sufficiently large. Future work will focus on the study of multiple transmit antennas and precoding design.

VI. Acknowledgment

The research was funded in part by Tekes, Nokia, Esju and CoreHW. We thank to Prof. Aarno Pärssinen for useful discussions and guidance.

References

[1] S. M. Razavizadeh, M. Ahn, and I. Lee, "Three-dimensional beamforming: A new enabling technology for 5g wireless networks," *IEEE Signal Process. Mag.*, vol. 31, no. 6, pp. 94–101, Nov 2014.

[2] S. Han *et al.*, "Large-scale antenna systems with hybrid analog and digital beamforming for millimeter wave 5g," *IEEE Commun. Mag.*, vol. 53, no. 1, pp. 186–194, January 2015.

[3] N. Seifi *et al.*, "Coordinated 3d beamforming for interference management in cellular networks," *IEEE Trans. Wireless Commun.*, vol. 13, no. 10, pp. 5396–5410, Oct 2014.

[4] O. El Ayach *et al.*, "Spatially sparse precoding in millimeter wave mimo systems," *IEEE Trans. Wireless Commun.*, vol. 13, no. 3, pp. 1499–1513, March 2014.

[5] P. Schniter and A. Sayeed, "Channel estimation and precoder design for millimeter-wave communications: The sparse way," in *Proc. Annual Asilomar Conf. Signals, Syst., Comp.*, Nov 2014, pp. 273–277.

[6] H. Deng and A. Sayeed, "Mm-wave mimo channel modeling and user localization using sparse beamspace signatures," in *Proc. IEEE Works. on Sign. Proc. Adv. in Wirel. Comms.* IEEE, 2014, pp. 130–134.

[7] W. Bajwa *et al.*, "Compressed Channel Sensing: A New Approach to Estimating Sparse Multipath Channels," *IEEE Proceedings*, vol. 98, no. 6, pp. 1058–1076, June 2010.

[8] S. Boyd *et al.*, "Distributed Optimization and Statistical Learning via the Alternating Direction Method of Multipliers," *Found. Trends Mach. Learn.*, vol. 3, no. 1, pp. 1–122, Jan. 2011.

[9] A. Soni and J. Haupt, "Learning sparse representations for adaptive compressive sensing," in *Proc. IEEE Int. Conf. Acoust., Speech, Signal Processing*, March 2012, pp. 2097–2100.

[10] R. Rubinstein and M. Elad, "Dictionary learning for analysis-synthesis thresholding," *IEEE Trans. Signal Process.*, vol. 62, no. 22, pp. 5962–5972, Nov 2014.

[11] G. Destino, M. Juntti, and S. Nagaraj, "Leveraging sparsity into massive mimo channel estimation with the adaptive-lasso," in *IEEE GlobalSIP*, 2015.

[12] J. M. Duarte-Carvajalino and G. Sapiro, "Learning to sense sparse signals: simultaneous sensing matrix and sparsifying dictionary optimization," *IEEE Trans. Image Process.*, vol. 18, no. 7, pp. 1395–1408, July 2009.