Estimation of Very Large MIMO Channels Using Compressed Sensing

Daniel C. Araújo⋆, André L. F. de Almeida⋆, João C. M. Mota⋆, Dennis Hui†
⋆GTEL-Wireless Telecommunication Research Group
Department of Teleinformatics Engineering, Federal University of Ceará, Fortaleza, Brazil
E-mails: {araujo,andre,mota}@gtel.ufc.br
†Ericsson Research, San José, California, USA, Email: dennis.hui@ericsson.com

Abstract— **In this paper, we propose an efficient pilot-assisted technique for the estimation of very-large MIMO (multiple-input multiple-output) channels exploiting the inherent sparsity of the channel. We first obtain an appropriate sparse decomposition model from a virtual channel representation of the very-large MIMO channel. Based on this model, we capitalize on a fundamental result of the compressed sensing (CS) to show that the channel matrix can be accurately estimated from very short training sequences compared to the number of used transmit antennas. We compare the normalized mean square error (NMSE) obtained using the proposed CS-based channel estimator, the least-square (LS) estimator and the Cramer-Rao lower bound (CRLB). The simulation results show that the proposed estimator obtains good performance, being 5 dB from the CRLB.**

Keywords— **Very-large MIMO channels, compressed sensing, sparsity, channel estimation, matching pursuit.**

I. Introduction

Very-large multiple-input multiple-output (VL-MIMO) systems have recently attracted significant research interests due to the potential to achieve high data rates and its robustness against interference, fading and antenna unit failures [1], [2]. Furthermore, according to [1], [2], in cellular networks the use of very-large arrays is an alternative to cell-size shrinking, which is the traditional way of increasing the network capacity . It is also expected that VL-MIMO systems potentially reduce uplink and downlink transmit powers as discussed in [3], [4].

The promised benefits of VL-MIMO systems are strongly dependent on the quality of the channel state information (CSI) available at the receiver and/or transmitter. Conventional channel estimation approaches rely on training sequences [1], [5], [6] and most of them use the least square (LS) approach to estimate the channel. However, this technique requires that the length of the training sequences be at least equal to the number of transmit antennas. It may be too restrictive in VL-MIMO systems where the number of transmit antennas is an order of magnitude higher than in traditional MIMO systems. Thus, it can lead to excessive use of critical communication resources such as energy and spectrum.

One way to overcome this problem in VL-MIMO systems is exploiting the sparsity of the channel to obtain CSI at transmitter and/or receiver using less communication resource. The compressed (or compressive) sensing (CS) theory is very attractive in problems where the signal of interest has a sparse representation [7], [8]. This subject has been under intensive study and several works have exploited this concept in different areas: sampling theory [9], ultrawideband channel estimation [10], [11] and radar [12], just to mention a few. The use of compressed sensing techniques holds for signals which are sparse in the standard coordinate basis or sparse with respect to some other orthonormal basis. In our context, the VL-MIMO channel matrix itself is not sparse. However, considering that the VL-MIMO channel is modelled by the virtual channel model [13], the channel matrix can be mapped into a sparse matrix using transmit/receive Fourier bases. The same occurs to consider the Weichselberger's channel model [14] which is a stochastic MIMO channel model that combines the advantages of the Kronecker model [15] and the virtual channel model [13].

In literature, some works have proposed channel estimators based on CS theory [10], [16]. In [17], the authors present a channel estimator based on CS theory considering the Dantzig Selector algorithm. However, it is not found any simulation comparison with another technique, even with the CRLB one. In this work, we propose a pilot-assisted technique based on CS theory to estimate the VL-MIMO CSI. Different from the estimator in [17], our approach uses the well-known greedy algorithm called orthogonal matching pursuit (OMP). The VL-MIMO channel is represented in a compressed space which is a sparse decomposition model based on a virtual channel representation of the very large channel. Through normalized mean square error (NMSE) curves, we show that sparse signal reconstruction methods, such as those based on matching pursuits, guarantee the recovery of the spatial structure of the VL-MIMO channel with high probability [7], [8]. Moreover, by exploiting sparsity we show that the channel matrix can be accurately estimated from very short training sequences compared to the number of used transmit antennas. As a contribution to the literature, we compare the (CS)-based estimator with the LS-based estimator and Cramer-Rao lower bound.

II. Compressive Sensing Prerequisites

First, let us introduce the concept of a sparse vector [7], [8].

This work is supported by Ericsson Research and Development Centre, Ericsson Telecommunications S.A., Brazil. The authors are partially supported by CAPES and CNPq.

***Definition** 1:* Let $\mathbf{h} \in \mathbf{R}^{N\text{x}1}$, if there exist a $K < N$ nonzero entries, then it is a K-sparse vector.

Define $\mathbf{x} \in \mathbf{R}^{N\text{x}1}$ as the signal of interest. Consider an orthonormal basis $\mathbf{\Psi} \in \mathbf{R}^{N\text{x}N}$ such that $\theta = \mathbf{\Psi}^T\mathbf{x}$ is a K-sparse vector. The matrix $\mathbf{\Psi}$ compresses $\mathbf{x}$ and is referred to as the "codebook". However, in general, we do not acquire $\mathbf{x}$ directly but rather acquire $M < N$ linear measurements $\mathbf{y} = \mathbf{\Phi x}$ using an $M \times N$ measurement matrix $\mathbf{\Phi}$. This kind of acquisition effectively compresses the signal. The acquired signal can then be written as $\mathbf{y} = \mathbf{\Phi x} = \mathbf{\Phi \Psi}\theta$, where the vector $\mathbf{y} \in \mathbf{R}^{M\text{x}1}$ contains the compressed measurements. The sparsity of vector θ enables that $\mathbf{x}$ can be accurately recovered from the following equation [7], [8], [18]

$$\hat{\theta} = \text{argmin}||\theta||_1 \text{ s.t. } ||\mathbf{y} - \mathbf{\Upsilon}\theta||_2 < \epsilon. \tag{1}$$

where, ϵ is upper-bounded by the l_2-norm of a noisy vector. The solution of Eq.(1) relies on the fact that the θ has to be sufficiently sparse. Furthermore, the matrix $\mathbf{\Upsilon}$ has to obey the restricted isometry property (RIP) to guarantee the stability of the problem [19].

***Definition** 2:* Consider each integer K=1, 2, $\ldots$, and l_0 norm of a vector θ being $||\theta||_0 = K$, for all $\theta \in \mathbf{R}^{N\text{x}1}$. The K-restricted isometric constant δ_K of the matrix Υ

$$(1 - \delta_K)||\theta||_2^2 \leq ||\Upsilon\theta||_2^2 \leq (1 + \delta_K)||\theta||_2^2.$$

Defining the constant δ_K, it is possible to guarantee the recovery of the information contained in θ [16].

Theorem 1: If the matrix Υ has RIP constant $\delta_{2K} < \sqrt{2}-1$, the solution $\hat{\theta}$ to Eq. (1) [16], [19] obeys

$$||\theta - \hat{\theta}||_2^2 \leq C_0 K^{-1/2}||\theta - \theta_{\mathbf{K}}||_1 + C_1\epsilon.$$

where, the constants C_0 and C_1 rely only on δ_{2k}. θ_K represents the approximation of vector θ retaining only the K most significant values.

Theorem 1 guarantees that sparse signals can be recovered even from noisy measurements. Unfortunately, the RIP constant calculation for a given matrix generally implies in a high computational complexity. Some classes of matrices have been studied in the literature. A case of interest is a matrix in $\mathbf{R}^{M\text{x}N}$ with independent Gaussian entries, having zero mean and variance $1/M$. From Theorem 1 the condition $K \leq \mathcal{O}(M/log(N/M))$ guarantees the correct estimation of the vector θ with high probability [7].

III. Channel Model

Consider a MIMO system with N transmit antennas and M receive antennas, where both N and M are very large, e.g. in the order of hundreds. Assume that the receiver does not have any knowledge about the communication channel. However, for a reliable communication, the receiver needs to estimate the channel using, e.g. a known training sequence sent by the transmitter. The received signal is denoted by

$$\mathbf{Y} = \mathbf{HS} + \mathbf{Z}. \tag{2}$$

The matrix $\mathbf{S} \in \mathbf{C}^{N\text{x}\tau}$ contains the training sequences sent by each one of the N transmit antennas, where τ denotes the training time, $\mathbf{H}$ is a $M \times N$ channel matrix and $\mathbf{Z}$ is the additive white Gaussian noise matrix whose entries have variance $N_0/2$. The matrix $\mathbf{Y} \in \mathbf{C}^{M\text{x}\tau}$ and its entries, $[\mathbf{Y}]_{m,\tau}$, stands for the discrete-time baseband received signal at m-th antenna and τ-th training symbol period.

As we have mentioned previously, it is a common practice in the literature to use i.i.d entries for modelling the MIMO channel. However, this assumption leads us to overestimate the spatial degrees of freedom of the MIMO channel. Thus, for a precise characterization of the VL-MIMO channel we abandon the i.i.d. assumption and adopt a spatially structured channel model and evaluate the impact of increasing the number of antennas on a structured model.

In [14], a stochastic MIMO channel model, inspired by Kronecker model and virtual channel model, was proposed. Combining the advantages of both models, the so-called Weichselberger's model shows enhanced capabilities to model the spatial multipath structure of the MIMO channel. However, the price for this better matching with the real channel is the knowledge of the transmit and receive one-side correlation matrices. Let us first introduce the eigenvalue decomposition of the one-side correlation matrices

$$\begin{aligned}\mathbf{R}_r &= \mathbf{E}\{\mathbf{HH}^H\} = \mathbf{U}_r\Lambda_r\mathbf{U}_r^H, \\ \mathbf{R}_t &= \mathbf{E}\{\mathbf{H}^H\mathbf{H}\}^T = \mathbf{U}_t\Lambda_t\mathbf{U}_t^H,\end{aligned}$$

where, the matrices $\mathbf{U}_r$ (resp. $\mathbf{U}_t$) and $\mathbf{\Lambda}_r$ (resp. $\mathbf{\Lambda}_t$) are the receive (resp. transmit) eigenvector and eingenvalue matrices.

In [14], Weichselberger proposed the following channel decomposition model

$$\mathbf{H} = \mathbf{U}_r\left(\mathbf{\Omega} \odot \mathbf{H}_w\right)\mathbf{U}_t^H \tag{3}$$

where $\mathbf{\Omega}$ is a positive and real-valued matrix, being equal to the square-root of the power coupling between transmit and receive eigenmodes, $\mathbf{H}_w$ has i.i.d entries, and $\odot$ denotes the Hadamard product. A full-rank matrix $\mathbf{\Omega}$ means a scattering-rich environment with maximum diversity. In this situation, model (5) is equivalent to the Kronecker model. In [14], typical examples for the structure of $\mathbf{\Omega}$ are given for different propagation environments.

In work [13], [14], it is shown that whether the number of antennas elements goes to infinity, the discrete Fourier transform (DFT) matrix serves as asymptotically optimal eigenvectors matrix in Eq. (4) for the channel matrix [13], [14]. This can be confirmed empirically in the Fig. 1, where it depicts the normalized l_2-norm of the error

$$e_{norm} = \frac{||\mathbf{W} - \mathbf{U}_r||_2^2}{MN} \tag{4}$$

where $\mathbf{W}$ is the DFT matrix and $\mathbf{U}_r$ is the measured eigenvector matrix of a one-side correlation matrix $\mathbf{R}_x = \mathbf{E}\{\mathbf{HH}^H\}$.

Figure 1 has been obtained from hundreds of eigenvector decompositions of the i.i.d channels and taken a mean matrix $\mathbf{U}_r$. If we consider a specific channel, this curve will be different, in other words this depends on the environment. Thus, how much antennas are necessary for the MIMO channel to be considered "very large"?. It is a difficult question, since each scenario has a different spatial structure.

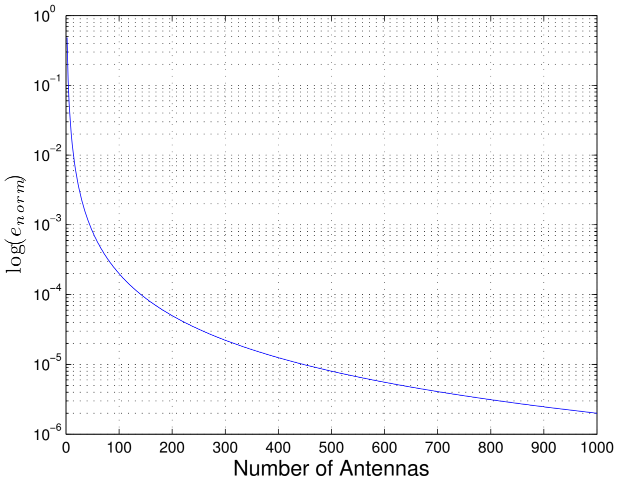

Fig. 1. Convergence of the error between the eigenbasis of the one-sided correlation matrix and the DFT matrix of the same dimensions as the number of antennas grows large.

Based on Fig. 1, we propose to model for VL-MIMO as follows

$$\begin{aligned}\mathbf{H} &= [\mathbf{W}_r + \mathbf{\Delta}_r]\,(\mathbf{\Omega} \odot \mathbf{H}_w)\,[\mathbf{W}_t + \mathbf{\Delta}_t]^H \\ &= \tilde{\mathbf{W}}_r\,(\mathbf{\Omega} \odot \mathbf{H}_w)\,\tilde{\mathbf{W}}_t^H, \end{aligned} \tag{5}$$

where $\tilde{\mathbf{W}}_t = \mathbf{W}_t + \mathbf{\Delta}_t$ and $\tilde{\mathbf{W}}_r = \mathbf{W}_r + \mathbf{\Delta}_r$. $\mathbf{W}_t$ and $\mathbf{W}_r$ are DFT matrix $N \times N$ and $M \times M$ respectively. The matrices $\mathbf{\Delta}_t$ and $\mathbf{\Delta}_r$ are the error model due to the approximation of eigenvectors of the one side matrices correlation $\mathbf{R}_r$ and $\mathbf{R}_t$ in DFT matrices. Their entries are zero mean white Gaussian random variables with $\mathbf{E}\{vec(\mathbf{\Delta}_t)vec(\mathbf{\Delta}_t)^H\} = \sigma_t^2\mathbf{I}_{M^2}$, $\mathbf{E}\{vec(\mathbf{\Delta}_r)vec(\mathbf{\Delta}_r)^H\} = \sigma_r^2\mathbf{I}_{N^2}$ and $\mathbf{E}\{vec(\mathbf{\Delta}_t)vec(\mathbf{\Delta}_r)^H\} = \mathbf{0}_{MN}$. The values of σ_r^2 and σ_t^2 depend on the number of antenna and the characteristics of the scenario like position of the scattering. If the number of antennas increases $\tilde{\mathbf{W}}_t \approx \mathbf{W}_t$ and $\tilde{\mathbf{W}}_r \approx \mathbf{W}_r$, thus

$$\mathbf{H} \approx \mathbf{W}_r\,(\mathbf{\Omega} \odot \mathbf{H}_w)\,\mathbf{W}_t^H. \tag{6}$$

Eq. (6) represents a virtual channel model for the VL-MIMO channel. Therefore,the asymptotic growth of the number of transmit and receive antennas, the Weichselberger's model converges to the virtual channel model. Such a convergence can be faster or not depending on the spatial structure of the environment.

IV. Proposed channel estimation technique

The channel estimation via LS impose that $\tau \geq N$ to obtain the VL-MIMO channel. However, this restriction is not practical, whereas a desirable situation is that $\tau \leq N$. This condition lead us a systems with less equation than variable resulting in a system of equations with infinitely many solutions. Using compressed sensing it is possible to estimate the VL-MIMO channel even under the condition $\tau \leq N$. First it is necessary to find an appropriate codebook (also known as "dictionary") $\mathbf{\Psi}$ to obtain a sparse representation of $\mathbf{H}$.

Consider Eq. (2) and use the $vec(.)$ operator to obtain

$$\begin{aligned} vec(\mathbf{Y}) &= vec(\mathbf{HS}) + vec(\mathbf{Z}) \\ &= (\mathbf{S} \otimes \mathbf{I}_M)\,vec(\mathbf{H}) + vec(\mathbf{Z}), \end{aligned} \tag{7}$$

where we have used the property $vec(\mathbf{ABC}) = (\mathbf{C}^T \otimes \mathbf{A})vec(\mathbf{B})$ of the Kronecker product. Substituting Eq. (5) in Eq. (7) and using again this property, yields

$$\begin{aligned} vec(\mathbf{Y}) &= (\mathbf{S}^T \otimes \mathbf{I}_M)vec(\mathbf{U}_r\,(\mathbf{\Omega}_w \odot \mathbf{H}_w)\,\mathbf{U}_t^H) + vec(\mathbf{Z}) \\ &= (\mathbf{S}^T \otimes \mathbf{I}_M)\left(\tilde{\mathbf{W}}_t \otimes \tilde{\mathbf{W}}_r\right) vec(\mathbf{\Omega}_w \odot \mathbf{H}_w) + vec(\mathbf{Z}). \end{aligned} \tag{8}$$

Defining

$$\mathbf{y} = vec(\mathbf{Y}) \in \mathbb{C}^{\tau M \times 1},\ \mathbf{z} = vec(\mathbf{Z}) \in \mathbb{C}^{\tau M \times 1}, \tag{9}$$
$$\mathbf{\Phi} \doteq \mathbf{S}^T \otimes \mathbf{I}_M \in \mathbb{C}^{\tau M \times NM}, \tag{10}$$
$$\mathbf{\Psi} \doteq \tilde{\mathbf{W}}_t \otimes \tilde{\mathbf{W}}_r \in \mathbb{C}^{NM \times NM}, \tag{11}$$
$$\mathbf{g} \doteq vec(\mathbf{\Omega}_w \odot \mathbf{H}_w) \in \mathbb{C}^{NM \times 1}, \tag{12}$$

allows one to compactly rewrite (8) as

$$\mathbf{y} = \mathbf{\Phi\Psi g} + \mathbf{z}, \tag{13}$$

which corresponds to a CS reconstruction model. Note that, in this model, the measurement matrix $\mathbf{\Phi}$ is constructed by the known training sequence matrix $\mathbf{S}$. Thus the number of measurements corresponds to the training sequence length, which is assumed to be shorter than the number of transmit antennas, i.e. $\tau < N$. The basis dictionary matrix is given by the compound transmit-receive spatial basis $\tilde{\mathbf{W}}_t$ and $\tilde{\mathbf{W}}_r$, herein assumed to be generated by DFT matrices. Moreover, in the CS model (13), the signal of interest to be recovered, represented by the vector $\mathbf{g}$, reveals the spatial structure of the VL-MIMO channel, which is likely to be sparse in practical propagation channels when very large transmit and receive arrays are used. The sparsity of $\mathbf{g}$ is determined by the number of non zero entries in the coupling matrix $\mathbf{\Omega}_w$.

From Eq. (1), we propose to solve the following optimization problem:

$$\hat{\mathbf{g}} = \arg\min ||\mathbf{g}||_1 \ \text{ s. t. } ||\mathbf{y} - \mathbf{\Phi\Psi g}||_2 < \epsilon. \tag{14}$$

There are some algorithms that solve (14). In this work, we choose the well known orthogonal matching pursuit (OMP) algorithm [11]. The advantage of this greedy algorithm is the reduced complexity compared to another class of algorithm based on l_1 relaxation, for example the Dantzig selector [17], [18]. The proposed CS-based channel estimation algorithm using OMP is described in the Table I.

From Eq. (7), a pertinent question is "how much training we need to recover the K-sparse vector $\mathbf{g}$"? Thanks to Theorem 1, it is possible to recover this vector almost surely if

$$\begin{aligned} K &\leq \mathcal{O}(\tau M/log(NM/\tau M)) \\ &\leq \mathcal{O}(\tau M/log(N/\tau)). \end{aligned} \tag{15}$$

Thus, from the above condition, it is possible to guide the choice of τ for a pre-specified number of transmit and receive antennas, and sparsity degree K of the VL-MIMO channel. Note that the term $\mathcal{O}(\tau M/log(N/\tau))$ increases linearly with

TABLE I

CS-BASED CHANNEL ESTIMATION ALGORITHM USING OMP

Step 1	Set parameters: the residual error $\mathbf{r}_0 = vec(\mathbf{Y})$; components of sparse representation $\hat{\mathbf{g}} = 0$, $\hat{\mathbf{g}} \in \mathbf{C}^{MN\text{x}1}$; $\mathbf{\Upsilon} = \mathbf{\Phi}\mathbf{\Psi} = [\mathbf{u}_1, \ldots, \mathbf{u}_{MN}]$ counter $t = 1$.
Step 2	Select the component in the dictionary that best match the residual error. $l = \arg\max_{i=1,\ldots,MN} \frac{\mathbf{u}_i^H \mathbf{r}_{t-1}}{\lvert\lvert\mathbf{u}_i\rvert\rvert}$
Step 3	Update the residual error $\mathbf{r}_t = \mathbf{r}_{t-1} - \frac{\mathbf{u}_l^H \mathbf{r}_{t-1}}{\lvert\lvert\mathbf{u}_l\rvert\rvert} \mathbf{u}_l$ Update the l-th component of $\hat{\mathbf{g}}$ $\hat{g}_l = \hat{\theta}_l + \frac{\mathbf{u}_i^H \mathbf{r}_{t-1}}{\lvert\lvert\mathbf{u}_i\rvert\rvert}$
Step 4	If $t < T_0$ and $\lvert\lvert\mathbf{r}_t\rvert\rvert_2 > \lvert\lvert vec(\mathbf{Y})\rvert\rvert_2$ the set $t = t + 1$ and go back to Step 2, otherwise go to Step 5
Step 5	Reconstruct the channel $vec(\mathbf{H}) = \mathbf{\Psi}\hat{\mathbf{g}}$.

M, so that the training sequence length can be reduced without compromising the reliability of the channel estimation. Note, however, that an increase in the number of antennas results in a increase in the dimensions of the "spatial" basis dictionary. This condition establishes a trade off between training sequence length and number of antennas.

V. Cramer-Rao Lower Bound

In the previous section, a estimator $\hat{\mathbf{g}}$ was derived for the vector $\mathbf{g}$. Thus, it is worthwhile to derive the CRLB. Assume that the Eq. 13 is rewritten as following:

$$\mathbf{y} = \sqrt{\alpha}\mathbf{s} + \mathbf{z}, \tag{16}$$

where, $\sqrt{\alpha}$ is the transmitted power that is known, meanwhile $\mathbf{s} = \mathbf{\Psi}\Phi\mathbf{g}$ and $\mathbf{g}$ are a unknown parameters. The vector $\mathbf{z}$ is a white Gaussian noise of zero mean and its variance is σ^2.

The dependence of signal $\mathbf{y}$ on $\mathbf{s}$ is explicit noted. It is easy to conclude that

$$p(\mathbf{y}; \mathbf{s}) = \frac{1}{(2\pi\sigma^2)^{\tau M/2}} e^{-\frac{1}{2\sigma^2}\sum_{n=0}^{\tau M}(y[n] - s[n;\mathbf{g}])^2}, \tag{17}$$

where, $y[n]$ and $s[n; \mathbf{g}]$ is the n-th element from the vectors $\mathbf{y}$ and $\mathbf{s}$, respectively.

CRLB is obtained from Fisher information matrix $\mathbf{\Gamma}_F(\mathbf{s})$, where (i, j)-th matrix element is given by

$$[\mathbf{\Gamma}_F((s))]_{ij} = -\mathbf{E}\{\frac{\partial^2 \ln p(\mathbf{y}; \mathbf{s})}{\partial s_i \partial s_j}\}. \tag{18}$$

From [20], the CRLB is given by

$$\begin{aligned} CRLB &= \frac{\partial \mathbf{f}(\mathbf{s})}{\partial \mathbf{s}} \mathbf{\Gamma}_F^{-1}(\mathbf{s}) \frac{\partial \mathbf{f}(\mathbf{s})^H}{\partial \mathbf{s}}. \\ &= (\mathbf{\Upsilon}^\dagger)(\mathbf{\Upsilon}^\dagger)^H. \end{aligned} \tag{19}$$

The covariance matrix of $\hat{\mathbf{g}}$ is $\mathbf{C}_{\hat{\mathbf{g}}}$ and it is related to the $CRLB$ as follows [20]

$$\mathbf{C}_{\hat{\mathbf{g}}} \geq (\mathbf{\Upsilon}^\dagger)(\mathbf{\Upsilon}^\dagger)^H \tag{20}$$

where, the operator $()^\dagger$ stands for pseudoinverse, $\Upsilon = \mathbf{\Phi}\mathbf{\Psi}$ and $\mathbf{f}(\mathbf{s}) = \Upsilon^\dagger \mathbf{s}$.

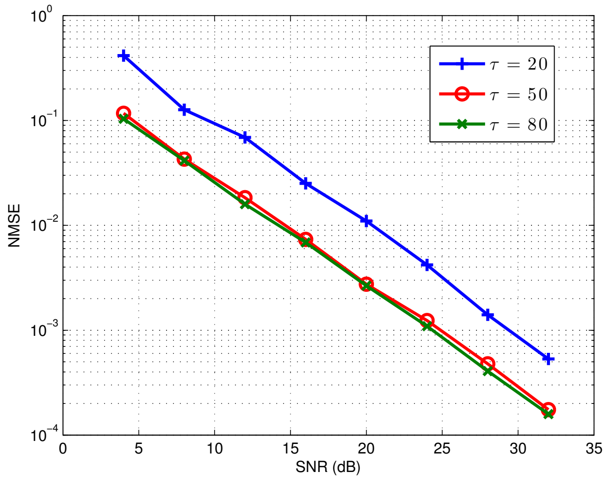

Fig. 2. NMSE of the channel for different lengths of training sequence, using the CS-estimator.

VI. Numerical Results

In this section, we simulate a system with $N = 100$ transmitters and $M = 30$ receivers antennas, using a training sequence under the constraint $\tau < 100$. To perform the simulation, a channel was synthesized using the model $\mathbf{H} = \tilde{\mathbf{W}}_r(\mathbf{\Omega} \odot \mathbf{H}_w)\tilde{\mathbf{W}}_t^H$. The coupling matrix $\mathbf{\Omega}_w$ is sparse and there are one hundred nonzero entries, all in the first column of the coupling matrix. In practice, this means that only one eigenmodes of the transmitter is connected to all other receivers eigenmodes. A more detailed description about this coupling matrix and others configuration can be found at [14]. The error matrices $\mathbf{\Delta}_r$ and $\mathbf{\Delta}_t$ depend on the field measurements. They vary with the environment and rely on the number of antennas, for this work we consider σ_r^2 and σ_t^2 on the order of 10^{-9}. The Fig. 2 represents the Normalized Mean Squares Error (NMSE) estimated of the channel for $\tau = 20$, 50, 80.

The results in Fig. 2 shows that is possible to estimate the channel even under the constraint $\tau << N$. As expected, the channel estimation for $\tau = 20$ was the poorest estimation meanwhile for $\tau = 50, 80$ the estimator presented the same performance. This way τ achieves its lower bound for the proposed estimator and any additional length to the training sequence means a waste of resources to the communication system.

Observe the Eq. (2), if the Least-Square approach was adopted the estimated channel would be

$$\hat{\mathbf{H}}^T = (\mathbf{S}^T)^\dagger \mathbf{Y}^T, \tag{21}$$

where $()^\dagger$ and $()^T$ mean pseudoinverse and transpose respectively. However, this strategy do not exploit the sparsity and it requires $\tau \geq N$ and resulting in a bad spectral efficiency. Therefore, using compressive sensing it is possible to design a VL-MIMO system reducing the penalty due to the large numbers of transmitter antennas on the spectral efficiency which would be caused adopting the Least Square Method.

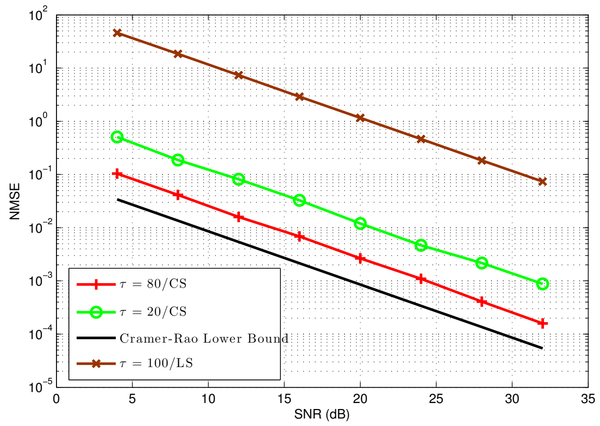

Fig. 3. Comparison among the CS-estimator, LS-estimator and CRLB.

The Fig. 3 compares the NMSE curve of LS-based estimator, CS-based estimator and Cramer-Rao Lower Bound. Note in Fig. 3 that the performance among CS-based estimator and LS-based estimator in a correlated channel model has huge gap. This is due to how CS-based estimator operates. Its functionality involves a iterative estimation, using the greedy algorithm OMP, where only the β most significant components of the basis $\mathbf{\Phi\Psi}$ expressed in Eq. 13 are estimated, where $\beta \geq K$ is the number of iterations set in greedy algorithm. On the other hand, LS is not an iterative process and estimates all the components. Because of this, components where there is no information the estimation is impaired by the noise. To obtain the results of Fig. 2 and Fig. 3 we use $\beta = 110$.

The Cramer-Rao lower bound is well known in the literature and widely use as a benchmark in problems of estimation. In the Fig. 3, we compare the two estimators with Cramer-Rao lower bound and due to the CS estimator take account a structured model of channel and exploit the inherent channel sparsity, this ones reaches a better performance than LS estimator, losing about 5 dB in performance compared to the CRLB. However, analysing table I, the process of searching the most significant component of the residual error $\mathbf{r}_t$ has a high computational cost. Thus, there is a trade off between the performance of the CS-estimator and the complexity of the algorithm OMP.

VII. Conclusion

This work verifies via simulation that the number of antennas increases than its eigenvector matrix converges to a DFT matrix as stated in [14]. This helps us to derive the codebook matrix to generate a sparse version of the channel to be estimated. Furthermore, the result of convergence give us a modified Weichselberger's model what was used to synthesize the channel for the simulations.

From the numerical results in this paper, we conclude that it is possible to estimate the channel even if the time of training sequence is less than the number of transmitter antennas. Moreover, the LS-based estimator and CRLB were compared with our approach. The numerical results show a huge gap between LS-based and CS-based estimator performances, indicating that the latter effectively exploits the sparsity of the channel. As future perspectives, we intend to extend this analysis for time-frequency selectivity channels, using the tensor model channel [21]. Moreover, the pilot design is a topic of future works.

Referências

[1] T. Marzetta, "Noncooperative cellular wireless with unlimited numbers of base station antennas," *IEEE Transactions on Wireless Communications*, vol. 9, pp. 3590 –3600, november 2010.

[2] D. Gesbert, M. Kountouris, R. Heath, C.-B. Chae, and T. Salzer, "Shifting the mimo paradigm," *Signal Processing Magazine, IEEE*, vol. 24, pp. 36 –46, sept. 2007.

[3] H. Q. Ngo, E. Larsson, and T. Marzetta, "Uplink power efficiency of multiuser mimo with very large antenna arrays," in *Communication, Control, and Computing (Allerton), 2011 49th Annual Allerton Conference on*, pp. 1272 –1279, sept. 2011.

[4] G. Caire and S. Shamai, "On achievable rates in a multi-antenna gaussian broadcast channel," in *IEEE International Symposium on Information Theory, 2001. Proceedings. 2001*, p. 147, 2001.

[5] T. Marzetta, "Signals, systems and computers, 2006. acssc '06. fortieth asilomar conference on," pp. 359 –363, 29 2006-nov. 1 2006.

[6] J. Jose, A. Ashikhmint, P. Whiting, and S. Vishwanath, "Scheduling and pre-conditioning in multi-user mimo tdd systems," in *IEEE International Conference on Communications, 2008. ICC '08.*, pp. 4100 –4105, may 2008.

[7] D. Donoho, "Compressed sensing," *IEEE Transactions on Information Theory*, vol. 52, pp. 1289 –1306, april 2006.

[8] E. Candes, J. Romberg, and T. Tao, "Robust uncertainty principles: exact signal reconstruction from highly incomplete frequency information," *IEEE Transactions on Information Theory*, vol. 52, pp. 489 – 509, feb. 2006.

[9] Y. Eldar, "Compressed sensing of analog signals in shift-invariant spaces," *IEEE Transactions on Signal Processing*, vol. 57, pp. 2986 –2997, aug. 2009.

[10] L. Shi, Z. Zhou, L. Tang, H. Yao, and J. Zhang, "Ultra-wideband channel estimation based on bayesian compressive sensing," in *2010 International Symposium on Communications and Information Technologies (ISCIT)*, pp. 779 –782, oct. 2010.

[11] J. Paredes, G. Arce, and Z. Wang, "Ultra-wideband compressed sensing: Channel estimation," *IEEE Journal of Selected Topics in Signal Processing*, vol. 1, pp. 383 –395, oct. 2007.

[12] C.-Y. Chen and P. Vaidyanathan, "Compressed sensing in mimo radar," in *Signals, Systems and Computers, 2008 42nd Asilomar Conference on*, pp. 41 –44, oct. 2008.

[13] A. Sayeed, "Deconstructing multiantenna fading channels," *IEEE Transactions on Signal Processing*, vol. 50, pp. 2563 – 2579, oct 2002.

[14] W. Weichselberger, M. Herdin, H. Ozcelik, and E. Bonek, "A stochastic mimo channel model with joint correlation of both link ends," *IEEE Transactions on Wireless Communications*, vol. 5, pp. 90 – 100, jan. 2006.

[15] J. Kermoal, L. Schumacher, K. Pedersen, P. Mogensen, and F. Frederiksen, "A stochastic mimo radio channel model with experimental validation," *IEEE Journal on Selected Areas in Communications*, vol. 20, pp. 1211 – 1226, aug 2002.

[16] M. Duarte and R. Baraniuk, "Kronecker compressive sensing," *IEEE Transactions on Image Processing*, vol. 21, pp. 494 –504, feb. 2012.

[17] W. Bajwa, J. Haupt, A. Sayeed, and R. Nowak, "Compressed channel sensing: A new approach to estimating sparse multipath channels," *Proceedings of the IEEE*, vol. 98, no. 6, pp. 1058–1076, 2010.

[18] M. Elad, *Sparse and Redundant Representations: From Theory to Applications in Signal and Image Processing*. Springer Publishing Company, Incorporated, 1st ed., 2010.

[19] E. Candes, "The restricted isometry property and its implications for compressed sensing," *Comptes Rendus Mathematique*, vol. 346, pp. 589–592, May 2008.

[20] S. Kay, *Fundamentals of Statistical Signal Processing: Estimation Theory*. No. v. 1 in Fundamentals of Statistical Signal Processing, Prentice-Hall PTR, 1998.

[21] N. Costa and S. Haykin, *Multiple-Input Multiple-Output Channel Models: Theory and Practice*. Adaptive and Learning Systems for Signal Processing, Communications and Control Series, Wiley, 2010.