Learning to detect: on site-specific channel estimation with hybrid MIMO architectures.

Joan Palacios †⋆, Dolores Garcia †⋆, Joerg Widmer†

† IMDEA Networks Institute, Madrid, Spain

⋆ Universidad Carlos III de Madrid, Madrid, Spain

E-mails: †{firstname.lastname}@imdea.org

***Abstract*—Acquiring channel information to establish a millimeter wave MIMO link considering a hybrid MIMO architecture is one of the most difficult problems in the 5G physical layer. In this paper we present the design of a signal processing algorithm that can learn some priors from the provided training data set to provide high accuracy channel estimates with low training overhead during the testing phase.**

***Index Terms*—machine learning, channel estimation, 5G**

I. The challenge

We aim to estimate the channel of a frequency selective hybrid millimeter wave MIMO-OFDM system at low SNR from received pilots as described in [1]. The system operates with uniform linear arrays (ULAs) at both ends, the transmitter and receiver are equipped with $N_{\mathrm{t}} = 16$ and $N_{\mathrm{r}} = 64$ antennas respectively. A hybrid MIMO architecture is considered, with $L_{\mathrm{t}} = 2$ and $L_{\mathrm{t}} = 4$ RF-chains. The MIMO-OFDM system operates with $K = 256$ subcarriers. The mmWave channel is assumed to be frequency selective.

II. Notation

We will be obviate basic vectorization properties with the matrix, Kronecker and scalar products such as

$$v(\mathbf{a}\mathbf{b}^{\mathrm{T}}) = \mathbf{b} \otimes \mathbf{a}$$

$$v(\mathbf{A}^{\mathrm{H}}\mathbf{B}\mathbf{C}) = (\mathbf{C}^{*} \otimes \mathbf{A})^{\mathrm{H}} v(\mathbf{B})$$

$$< \mathbf{a} \otimes \mathbf{b}, \mathbf{c} \otimes \mathbf{d} > = < \mathbf{a}, \mathbf{c} >< \mathbf{b}, \mathbf{d} >$$

III. Problem formulation

The geometric channel model for subcarrier f can be written as:

$$\mathbf{H}_f = \sum_{l}^{L} \alpha_l e^{2\pi i \tau_l f} \mathbf{a}_{\mathrm{RX}}(\phi_l) \mathbf{a}_{\mathrm{TX}}^{\mathrm{H}}(\theta_l)$$

Where l is the l-th path, α_l is the complex gain of the l-th path , L is the total number of paths, τ_l is the delay of the l-th path, ϕ_l and θ_l are the angles of arrival and departure of the l-th path respectively and $\mathbf{a}_{\mathrm{RX}}(\phi_l)$ and $\mathbf{a}_{\mathrm{TX}}^{\mathrm{H}}(\theta_l)$ are the array steering vectors of the receive and transmit antennas respectively.

This vectorizes into

$$\mathbf{h}_f = \sum_{l}^{L} \alpha_l e^{2\pi i \tau_l f} (\mathbf{a}_{\mathrm{TX}}^{*}(\theta_l) \otimes \mathbf{a}_{\mathrm{RX}}(\phi_l))$$

and can be horizontally stacked into the matrix

$$\hat{\mathbf{H}} = \sum_{l}^{L} \alpha_l (\mathbf{a}_{\mathrm{TX}}^{*}(\theta_l) \otimes \mathbf{a}_{\mathrm{RX}}(\phi_l)) \mathbf{a}_{\mathrm{F}}^{\mathrm{T}}(\tau_l).$$

Then the measurements are computed as

$$\mathbf{M} = \mathbf{\Phi}^{\mathrm{H}} \hat{\mathbf{H}} + \mathbf{N}$$

where $\mathbf{N}$ may have correlation between the noise components and $\mathbf{M}$ is the vector of stacked received samples. To obtain the whittened measurement matrix in which $\mathbf{N}$ is a matrix with identically independently distributed entries of complex white noise with unknown standard deviation σ, we calculate the noise covariance matrix as in [1]. The covariance matrix is

$$C = \mathrm{blkdiag}(\mathrm{W}_{\mathrm{tr}}^{(1)*}\mathrm{W}_{\mathrm{tr}}^{(1)}, \ldots, \mathrm{W}_{\mathrm{tr}}^{(\mathrm{M})*}\mathrm{W}_{\mathrm{tr}}^{(\mathrm{M})})$$

since the noise at a given sample m is given by $\mathbb{E}\{n^{(i)}, n^{(j)}\} = W_{\mathrm{tr}}^{(i)*}\sigma^2\delta[i-j]W_{\mathrm{tr}}^{(j)}$, where $W_{\mathrm{tr}}^{(i)}$ is the combiner for a given sample i. Using the Cholesky factorization $C = D_w^{*} D_w$. Therefore, by taking $\Phi_w = D_w^{-1}\Phi$ and $\hat{H}_w = D_w^{-1}\hat{H}$, we obtain the whittened measurement matrix:

$$\mathbf{M_w} = \mathbf{\Phi_w}^{\mathrm{H}} \hat{\mathbf{H}}_{\mathbf{w}} + \mathbf{N_w}$$

where $\mathbf{N_w}$ is a matrix with identically independently distributed entries of complex white noise. From this point we will drop the sub index but assume the noise is white.

To ease the notation we will define the spatial component of a path as

$$\mathbf{a}_{\mathrm{RX-TX}}(\phi_l, \theta_l) = \mathbf{a}_{\mathrm{TX}}^{*}(\theta_l) \otimes \mathbf{a}_{\mathrm{RX}}(\phi_l)$$

and the channel component of a path as

$$\mathbf{a}_{\mathrm{RX-TX-F}}(\phi_l, \theta_l, \tau_l) = \mathbf{a}_{\mathrm{F}}(\tau_l) \otimes \mathbf{a}_{\mathrm{RX-TX}}(\theta_l, \phi_l)$$

IV. Approach

We base our strategy in the fact that $\hat{\mathbf{H}}$ is a very sparse matrix in the sense that the amount of paths L is much smaller than the maximum matrix rank $\min(N_{\mathrm{r}}N_{\mathrm{t}}, K)$. This make the fact that each matrix component $\mathbf{a}_{\mathrm{F}}(\tau_l) \otimes \mathbf{a}_{\mathrm{TX}}^{*}(\theta_l) \otimes \mathbf{a}_{\mathrm{RX}}(\phi_l)$ is close to be orthogonal to each other, this is

$$< \mathbf{a}_{\mathrm{RX-TX-F}}(\phi_l, \theta_l, \tau_l), \mathbf{a}_{\mathrm{RX-TX-F}}(\phi_{l'}, \theta_{l'}, \tau_{l'}) > \simeq 0$$

for all $l \neq l'$.

Our algorithm will consist on extracting the path parameters $(\phi_l, \theta_l, \tau_l)$ one by one and then subtracting their contribution in an orthogonal matching pursuit algorithm (OMP). The key to our algorithm rises from the way we search the path parameters by projections and the way we detect a new path.

At each step of the OMP we want to solve the best matching channel component, this is

$$\max_{\phi_l,\theta_l,\tau_l} |< v(\mathbf{\Phi}^{\mathrm{H}}\mathbf{a}_{RX-TX}(\phi_l,\theta_l)\mathbf{a}_F^{\mathrm{T}}(\tau_l)), v(\mathbf{M}) >|$$

which can be simplified into

$$\max_{\phi_l,\theta_l,\tau_l} |\mathbf{a}_{RX-TX}(\phi_l,\theta_l)^{\mathrm{H}}\mathbf{\Phi}\mathbf{M}\,\mathbf{a}_F^{\mathrm{T}}(\tau_l))|.$$

We start by considering a small set of values for the path parameters equally spaced in their domains, we choose to have a resolution of $4K$ values for τ_l and of $N_\mathrm{r}/2$ ($N_\mathrm{t}/2$) values for ϕ_l (θ_l). Since we are considering for the angles a smaller resolution than the required one for them to cover the whole angle spectrum we substitute each $\mathbf{a}_{\mathrm{RX}}(\phi_l)$ by a sector beam-pattern $\hat{\mathbf{a}}_{\mathrm{RX}}(\phi_l)$ of width $4\pi/N_\mathrm{r}$. The same trick applies to the transmitter values and thus we can define $\hat{\mathbf{a}}_{RX-TX}(\phi_l,\theta_l) = \hat{\mathbf{a}}_{TX}(\theta_l)^* \otimes \hat{\mathbf{a}}_{RX}(\phi_l)$. The sector beam-pattern we are considering is the one defined in [2]. With this definition we can extract a coarse version of the path parameters $(\phi_l, \theta_l, \tau_l)$ maximizing

$$d_{\phi_l,\theta_l,\tau_l} = |\hat{\mathbf{a}}_{RX-TX}(\phi_l,\theta_l)^{\mathrm{H}}\mathbf{\Phi}\mathbf{M}\,\mathbf{a}_F^{\mathrm{T}}(\tau_l))|.$$

A. Detection

Now we want to know if those parameters can be considered as a path detection, to sort this, we take into account the null hypothesis of $\mathbf{H} = 0$, in that case all elements of $\mathbf{M}$ are independent white noise and thus $\hat{\mathbf{a}}_{RX-TX}(\phi_l,\theta_l)^{\mathrm{H}}\mathbf{\Phi}\mathbf{M}\,\mathbf{a}_F^{\mathrm{T}}(\tau_l))$ is too. Consequently $d_{\phi_l,\theta_l,\tau_l}$ follows a Rayleigh distribution. The Rayleigh cumulative distribution function is

$$F(d_{\phi_l,\theta_l,\tau_l}) = 1 - e^{-\frac{d^2_{\phi_l,\theta_l,\tau_l}}{2\sigma^2}}.$$

Since only for a very few path parameters there should be a channel contribution we have that the median of all computed values of $d_{\phi_l,\theta_l,\tau_l}$ should be close to that of the Rayleigh distribution $\sigma\sqrt{2\ln(2)}$. Then σ can be approximated as

$$\sigma \simeq \mu(d_{\phi_l,\theta_l,\tau_l})/\sqrt{2\ln(2)}.$$

In this case, using that the cumulative function of $\max x_k$ ($F_{\max}(x)$) can be computed as $F_{\max}(x) = \prod F_k(x)$ we have that

$$F_{\max}(\max(d_{\phi_l,\theta_l,\tau_l})) = (1 - e^{-\max(d_{\phi_l,\theta_l,\tau_l})^2/(2\sigma^2)})^{N_\mathrm{r}N_\mathrm{t}K}.$$

Knowing this we can compute a threshold of θ confidence

$$\mu(d_{\phi_l,\theta_l,\tau_l})\sqrt{-\log_2(1-(\theta)^{\frac{1}{N_\mathrm{r}N_\mathrm{t}K}})}$$

to decide where to continue computing paths or stop the OMP iterations.

B. Refinement

Once a path has been detected we proceed to refine the path components by iterative projections. What we are doing here is to freeze two of the variables and increase the resolution of the other one.

1) First steps: We are going to first need to adapt our estimation to be able to handle a higher resolution due to the trick we did with the angle resolution.

We start by increasing the time resolution by computing the maximum of $|\hat{\mathbf{a}}_{RX-TX}(\phi_l,\theta_l)^{\mathrm{H}}\mathbf{\Phi}\mathbf{M}\,\mathbf{a}_F^{\mathrm{T}}(\tau_l))|$ for fixed (ϕ_l, θ_l) and τ_l with a resolution of $32K$ equally spaced points.

Then by fixing τ_l we can simplify the expression to

$$\hat{\mathbf{a}}_{\mathrm{RX}}^{\mathrm{H}}(\phi_l)\overline{\mathbf{H}}(\tau_l)\hat{\mathbf{a}}_{\mathrm{TX}}(\theta_l)$$

with $\overline{\mathbf{H}}(\tau_l)$ such that $v(\overline{\mathbf{H}}(\tau_l)) = \mathbf{\Phi}\mathbf{M}\,\mathbf{a}_F^{\mathrm{T}}(\tau_l)$.

Now we proceed to refine the angle components with the highest amount of antennas, for simplicity, let's assume that $N_\mathrm{t} > N_\mathrm{r}$. By increasing the resolution of θ_l to $32N_\mathrm{t}$ equally spaced points we don't need to use the sector beam-pattern trick, thus we can simply maximize

$$\hat{\mathbf{a}}_{\mathrm{RX}}^{\mathrm{H}}(\phi_l)\overline{\mathbf{H}}(\tau_l)\mathbf{a}_{\mathrm{TX}}(\theta_l)$$

over θ_l while the other path parameters are fixed.

Finally we refine the expression respect to the remaining angle. Again, the trick is not required and we can maximize

$$\mathbf{a}_{\mathrm{RX}}^{\mathrm{H}}(\phi_l)\overline{\mathbf{H}}(\tau_l)\mathbf{a}_{\mathrm{TX}}(\theta_l)$$

over θ_l while the other path parameters are fixed.

2) Iteration steps: Now that with the first steps we removed the angle uncertainty caused by the sector beam-pattern we can proceed to repeat the same steps in an iterative way by substituting all sector beam-patterns $\hat{\mathbf{a}}$ by array responses $\mathbf{a}$.

V. Training

The value we need to fit in order to improve the results is the detection threshold θ. We create a specific optimization method for the structure of our approach. This optimization method is focused on reducing the optimization time while being able to perform a high resolution grid search of the parameter.

We base our training algorithm on the fact that our approach is a greedy algorithm with a stop condition. This means that we can predict when a change on the solution will happen based on the selected threshold. Known this, we create a modified version of our approach that saves all channel iterations together with the computed threshold required for them to pass up to a minimum threshold value, in our case 0.7. Once outside the function, we can evaluate these channel estimations and compute the error as a step-wise function of the threshold. Then we apply the average operation to the error step-wise functions for different scenarios to get a better and smoother result of the error behavior over different threshold values.

The obtained results are the following.

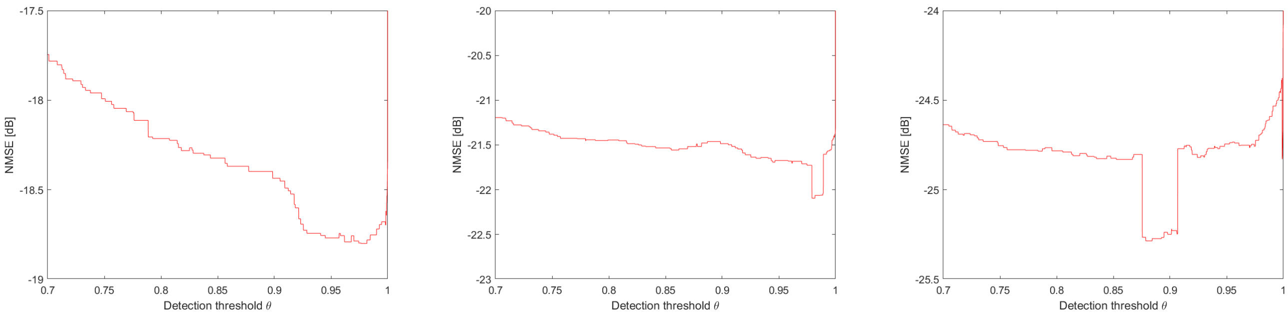

Fig. 1: NMSE behavior over the decision threshold of the 3 train datasets.

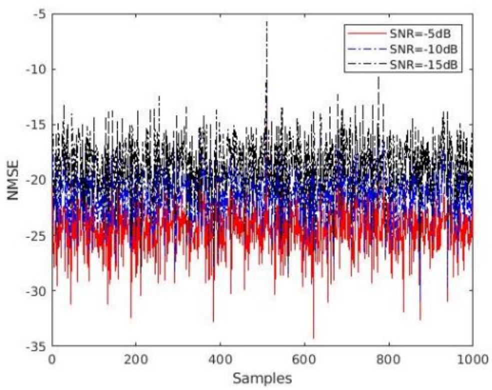

Fig. 2: NMSE of the first 1000 samples of the train dataset.

VI. Numerical Results

The NMSE of 1000 of the train channels is presented in Figure 2 for the different datasets, corresponding to SNR=$\{-15, -10, -5\}$. Indicative NMSE obtained for the first 1000 channels are:$\{-18.7046, -21.4987, -24.4888\}$.

References

[1] Javier Rodríguez-Fernández, Nuria González-Prelcic, Kiran Venugopal, and Robert W Heath. Frequency-domain compressive channel estimation for frequency-selective hybrid millimeter wave mimo systems. *IEEE Transactions on Wireless Communications*, 17(5):2946–2960, 2018.

[2] J. Palacios, D. De Donno, and J. Widmer. Lightweight and effective sector beam pattern synthesis with uniform linear antenna arrays. *IEEE Antennas and Wireless Propagation Letters*, 16:605–608, 2017.