Leveraging Sparsity into Massive MIMO Channel Estimation with the A-LASSO

G. Destino, M. Juntti and S. Nagarj

CWC CENTRE FOR WIRELESS COMMUNICATIONS
cwc.oulu.fi University of Oulu

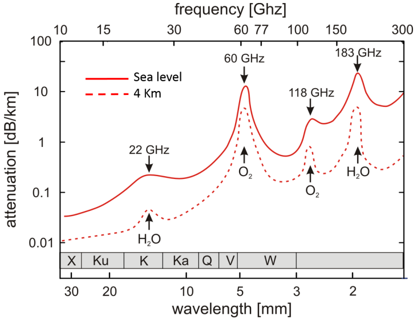
NOKIA

December 12, 2015

Going to higher carrier frequencies

Penetration loss

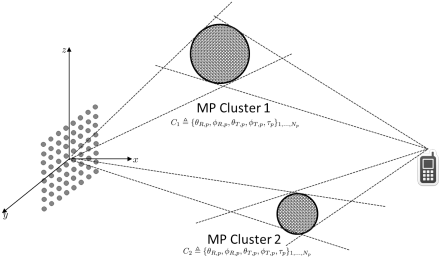
Freq.	Loss (dB)	Material
1.8 GHz	13.5	Wall
2.3 GHz	12.8	Wall
28 GHz	28	Wall



The channel at higher frequencies is characterized by

- Less reflections
- Less spatial diversity
- More attenuation

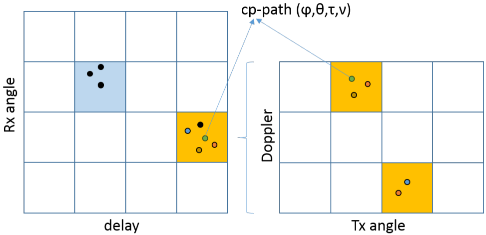
Physical channel model (frequency domain)



$$\underbrace{\mathbf{H}(kT_0, n\delta_f)}_{\substack{M\times N \\ \text{matrix}}} = \sum_{c=1}^{C}\sum_{p=1}^{P_c} b_{cp} \underbrace{e^{-j\frac{2\pi}{\lambda}\mathbf{P}_r^{\mathsf{T}}\mathbf{u}_{r,cp}}}_{\substack{\text{RX steering} \\ \text{vector}}} \underbrace{e^{j\frac{2\pi}{\lambda}\mathbf{u}_{t,cp}^{\mathsf{T}}\mathbf{P}_t}}_{\substack{\text{TX steering} \\ \text{vector}}} e^{-j2\pi\tau_{cp}n\delta_f} e^{j2\pi\nu_{cp}kT_0}$$

- C and P_c, number of clusters and paths per cluster
- β_{cp}, path coefficient
- τ_{cp}, path delay
- ν_{cp}, doppler frequency of the path
- $\mathbf{P}_r \in \mathbb{R}^{3\times M}$ and $\mathbf{P}_t \in \mathbb{R}^{3\times N}$, location of the RX and TX antenna elements
- $\mathbf{u} = f(\phi, \theta)$ is the unit vector of the wave-direction
- ϕ and θ, azimuth and elevation

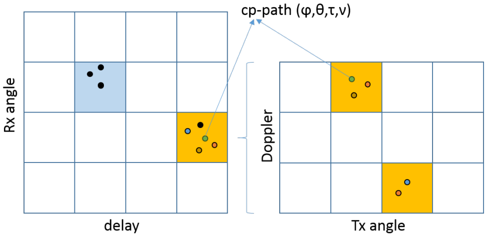
Full-dimensional channel model



$$\underbrace{\mathcal{H}}_{4D-tensor} = \sum_{c=1}^{C}\sum_{p=1}^{P_c} b_{cp}\,\underbrace{\mathbf{v}_r\left(\frac{\mathbf{P}_r^{\mathrm{T}}\mathbf{u}_{r,cp}}{\lambda}; M\right)}_{\text{RX spatial frequency}} \circ \underbrace{\mathbf{v}_t\left(\frac{\mathbf{P}_t^{\mathrm{T}}\mathbf{u}_{t,cp}}{\lambda}; N\right)^*}_{\text{TX spatial frequency}} \circ \underbrace{\mathbf{v}_f\left(\frac{\tau_{cp}W}{N_{\mathrm{FFT}}}; N_{\mathrm{FFT}}\right)}_{\text{DFT-frequency}} \circ \underbrace{\mathbf{v}_d\left(\frac{\nu_{cp}T_0}{T}; M_d\right)}_{\text{Doppler frequency}}$$

- Channel is characterized can be decomposed in multiple dimensions
- Generally, by increasing the dimensions paths (clusters) can be better separated
- Each dimension can be modeled with a discrete Fourier-based dictionary,
- $\mathbf{v}(x; L) = \left[e^{-j2\pi l x}\right]_{l=0,\ldots,L-1}$ and $x \in [0,1)$, L-size discrete frequency vector

Full-dimensional sparse channel model

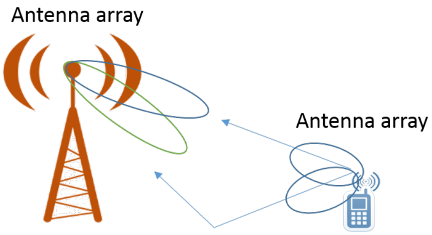


$$\underbrace{\mathcal{H}}_{4D-tensor} = \sum_{c=1}^{C}\sum_{p=1}^{P_c} b_{cp} \underbrace{\left(\mathbf{D}_r \mathbf{z}_{r,cp} \circ \mathbf{D}_t \mathbf{z}_{t,cp} \circ \mathbf{D}_f \mathbf{z}_{f,cp} \circ \mathbf{D}_d \mathbf{z}_{d,cp}\right)}_{\mathcal{H}_{cp},\ \text{path tensor}}$$

$$\text{vec}(\mathcal{H}) = \sum_{c=1}^{C}\sum_{p=1}^{P_c} b_{cp} \mathbf{D}_d \mathbf{z}_{d,cp} \otimes \mathbf{D}_f \mathbf{z}_{f,cp} \otimes \mathbf{D}_r \mathbf{z}_{t,cp} \otimes \mathbf{D}_t \mathbf{z}_{r,cp}$$

$$= \sum_{c=1}^{C}\sum_{p=1}^{P_c} b_{cp} \left(\mathbf{D}_d \otimes \mathbf{D}_f \otimes \mathbf{D}_t \otimes \mathbf{D}_r\right) \underbrace{\left(\mathbf{z}_{d,cp} \otimes \mathbf{z}_{f,cp} \otimes \mathbf{z}_{t,cp} \otimes \mathbf{z}_{r,cp}\right)}_{\text{sparse vector}}$$

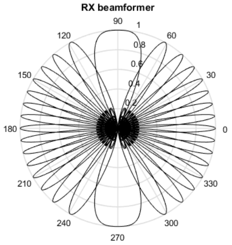
Estimation model in the angular domains

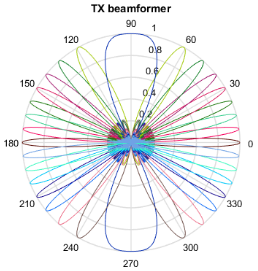


$$\mathrm{vec}(\mathbf{Y}) \quad = \quad (\mathbf{X}^{\mathsf{T}}\mathbf{P}^{\mathsf{T}} \otimes \mathbf{W}^{\mathsf{H}})\mathrm{vec}(\mathbf{H}) + (\mathbf{I}_N \otimes \mathbf{W}^{\mathsf{H}})\mathrm{vec}(\mathbf{N})$$

- $\mathbf{X} \in \mathbf{I}^{N \times N_t}$ pilot matrix
- $\mathbf{P} \in \mathbb{C}^{N \times N}$, TX precoder
- $\mathbf{W} \in \mathbb{C}^{M \times M_t}$, receive combiner
- Option 1) $\mathbf{P}, \mathbf{W}$ are dictionaries (fixed-beams).
- Option 2) $\mathrm{vec}(\mathbf{H}) = (\mathbf{D}_t \otimes \mathbf{D}_r)(\mathbf{z}_t \otimes \mathbf{z}_r)$

Sequential (TX) beam-switch





$$\text{vec}(\mathbf{Y}) \quad = \quad (\mathbf{X}^{\mathsf{T}}\mathbf{P}^{\mathsf{T}} \otimes \mathbf{W}^{\mathsf{H}})\text{vec}(\mathbf{H}) + (\mathbf{I}_N \otimes \mathbf{W}^{\mathsf{H}})\text{vec}(\mathbf{N})$$

- $\mathbf{P} = \mathbf{D}_t$, DFT-based TX beamformer
- $\mathbf{W} = \mathbf{D}_r$, DFT-based RX beamformer
- y_i is the sparse coefficient

Estimation algorithm with adaptive dictionary

Classic channel estimation

$$\hat{\mathbf{z}} = \arg\min_{\mathbf{z}\in\mathbb{C}^L} \lambda\|\mathbf{z}\|_1 + \frac{1}{2}\|\mathbf{y} - \underbrace{(\mathbf{X}^\mathsf{T}\mathbf{P}^\mathsf{T} \otimes \mathbf{W}^\mathsf{H})}_{\mathbf{\Phi}}\mathbf{\Psi}\mathbf{z}\|_2^2,$$

- $\mathbf{\Psi}\hat{\mathbf{z}} = \hat{\mathbf{h}}$, channel estimate
- $\mathbf{\Psi} = (\mathbf{D}_t \otimes \mathbf{D}_r)(\mathbf{z}_t \otimes \mathbf{z}_r)$, fixed

Adaptive-LASSO

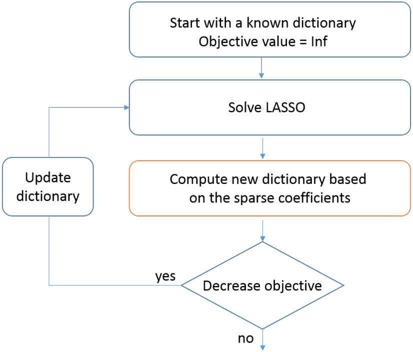
$$\begin{aligned}\left(\hat{\mathbf{z}}, \mathbf{\Psi}(\hat{\mathbf{U}})\right) \quad = \quad & \arg\min_{\substack{\mathbf{z}\in\mathbb{C}^L \\ \mathbf{U}\in\mathbb{C}^{L\times 2}}} \lambda\|\mathbf{z}\|_1 + \frac{1}{2}\|\mathbf{y} - (\mathbf{X}^\mathsf{T}\mathbf{P}^\mathsf{T} \otimes \mathbf{W}^\mathsf{H})\mathbf{\Psi}(\mathbf{U})\mathbf{z}\|_2^2, \\ \text{s.t.} \quad & \psi_i(\mathbf{u}_i) = \mathbf{v}_r(u_{1i}) \otimes \mathbf{v}_t(u_{2i}) \\ & 0 \leq u_{ij} \leq 1, \ \forall ij\end{aligned}$$

- Parameterize the dictionary
- Allow the dictionary to change during the optimization
- "Rule-of-thumb", search for dictionary vector closer to the channel spatial frequencies

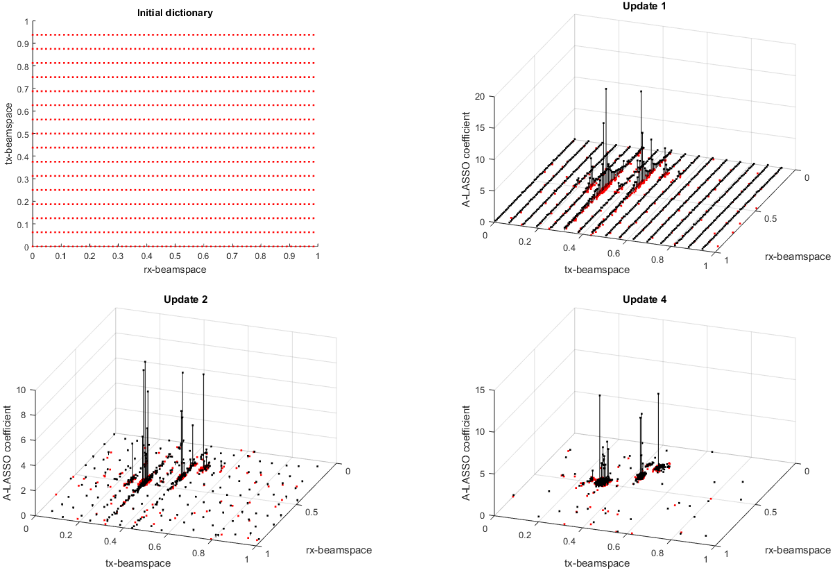
Key points of the optimization method

Dictionary adaptation

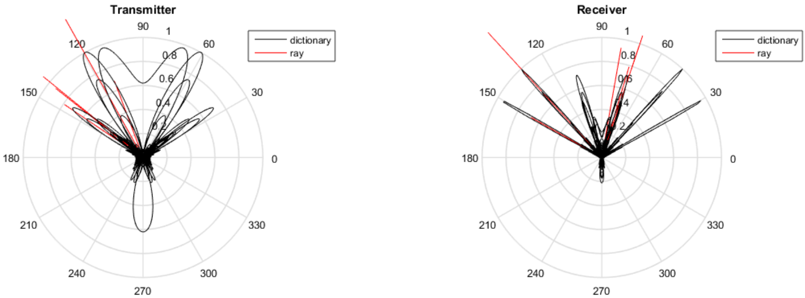
- Sparse coefficients are considered as importance weight for the corresponding dictionary vector
- Resample the dictionary vectors based on their weight
- Using the vector parameters, replace each resampled vector with one with a close parameter
- (Practical) Use a finer grid for the selection of the new vector parameter



Learning the beamspace



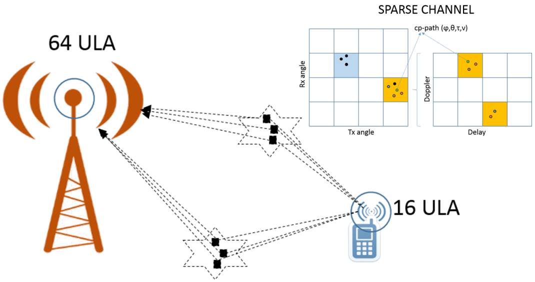
A-LASSO decomposition



- Accurate estimation of the DoA and AoD

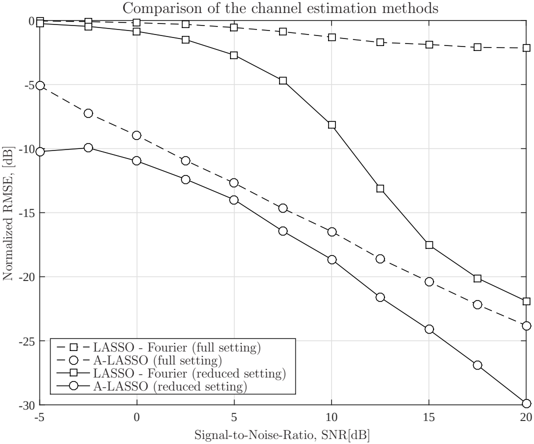
Performance evaluation

Synthetic channel

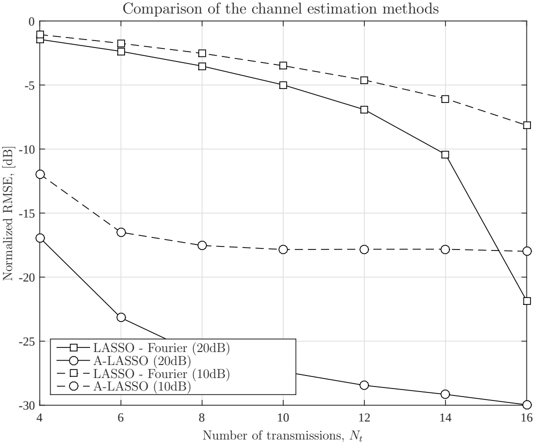


- 64x16 MIMO, ULA model
- (typical) 3 clusters, 2 paths per cluster
- $\mathbf{P} = \mathbf{I}$
- $\mathbf{W} = \mathbf{I}$
- reduce setting: 32x8 active antennas

Performance evaluation (SNR)



Performance evaluation (training)



THANK YOU
End