Learned Chester: A Multilevel-Greedy and Bayesian Compressive Channel Estimator for Frequency-Selective Hybrid mmWave MIMO Systems

Sameera Bharadwaja*, Marios Kountouris†, Christo Kurisummoottil†, Rakesh Mundlamuri†, Chandra R. Murthy*, and Sai Subramanyam Thoota*

December 16, 2020

*Indian Institute of Science, Bangalore, India.
†EURECOM, Sophia-Antipolis, France.

Problem Statement

System Model

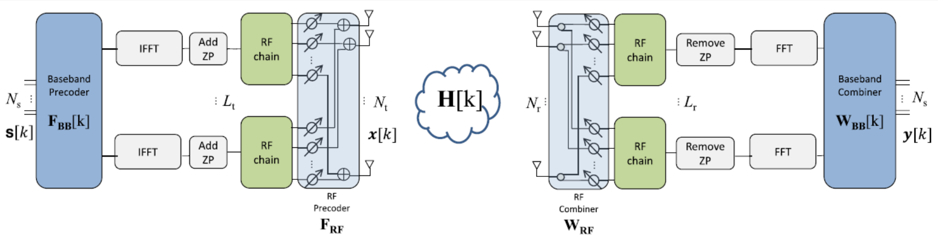


Fig. 1. Millimeter wave MIMO system based on a hybrid architecture [1]

- $N_t = 16$ transmit antennas, $N_r = 64$ receive antennas
- $L_t = 2$ transmit RF chains, $L_r = 4$ receive RF chains
- $N_s = 2$ data streams
- $K = 256$ subcarriers

Channel Model

- **d**$^{\text{th}}$ delay tap of the **Extended virtual channel model**:

$$\mathbf{H}_d \approx \tilde{\mathbf{A}}_{\mathrm{R}} \boldsymbol{\Delta}_d^{v} \tilde{\mathbf{A}}_{\mathrm{T}}^{*}, \qquad 0 \le d \le N_c - 1,$$

  where

  - $\boldsymbol{\Delta}_d^{v} \in \mathbb{C}^{G_r \times G_t}$ is a **sparse matrix** containing the path gains of the quantized spatial frequencies
  - $\tilde{\mathbf{A}}_{\mathrm{R}} \in \mathbb{C}^{N_r \times G_r}$ and $\tilde{\mathbf{A}}_{\mathrm{T}} \in \mathbb{C}^{N_t \times G_t}$ are the array steering matrices

- **Frequency domain representation** of subcarrier k:

$$\mathbf{H}[k] \approx \tilde{\mathbf{A}}_{\mathrm{R}} \left(\sum_{d=0}^{N_{\mathrm{c}}-1} \boldsymbol{\Delta}_d^{v} e^{-\mathrm{j}\frac{2\pi k}{K} d} \right) \tilde{\mathbf{A}}_{\mathrm{T}}^{*} \approx \tilde{\mathbf{A}}_{\mathrm{R}} \boldsymbol{\Delta}^{\mathrm{v}}[k] \tilde{\mathbf{A}}_{\mathrm{T}}^{*}$$

- Concatenated rx signal of **M training frames** in subcarrier k:

$$\underbrace{\begin{bmatrix} \mathbf{y}^{(1)}[k] \\ \vdots \\ \mathbf{y}^{(M)}[k] \end{bmatrix}}_{\mathbf{y}[k]} = \underbrace{\begin{bmatrix} \boldsymbol{\Phi}^{(1)} \\ \vdots \\ \boldsymbol{\Phi}^{(M)} \end{bmatrix}}_{\boldsymbol{\Phi}} \boldsymbol{\Psi} \, \mathbf{h}^{\mathrm{v}}[k] + \underbrace{\begin{bmatrix} \mathbf{n}^{(1)}[k] \\ \vdots \\ \mathbf{n}^{(M)}[k] \end{bmatrix}}_{\mathbf{n}[k]},$$

$\mathbf{h}^{\mathrm{v}}[k]$ → $\mathrm{vec}(\boldsymbol{\Delta}^{\mathrm{v}}[k])$

$\mathbf{n}[k]$ → Colored Noise

where

$$\boldsymbol{\Phi}^{(m)} = \mathbf{q}^{(m)T} \; \mathbf{F}_{\mathrm{tr}}^{(m)T} \otimes \mathbf{W}_{\mathrm{tr}}^{(m)*}, \qquad \boldsymbol{\Psi} = \bar{\tilde{\mathbf{A}}}_{\mathrm{T}} \otimes \tilde{\mathbf{A}}_{\mathrm{R}}$$

$\mathbf{q}^{(m)T}$ → Pilot signal

$\mathbf{F}_{\mathrm{tr}}^{(m)T}$ → RF Precoder

$\mathbf{W}_{\mathrm{tr}}^{(m)*}$ → RF Combiner

$\boldsymbol{\Psi}$ → Sparsifying Dictionary

Noise Whitening

- Noise covariance matrix:

$$\mathbf{C}_\mathrm{w} = \mathbb{E}\left[\mathbf{n}[k]\mathbf{n}^*[k]\right] = \mathrm{blkdiag}\{\overbrace{\mathbf{W}_\mathrm{tr}^{(1)^*}\mathbf{W}_\mathrm{tr}^{(1)}, \ldots, \mathbf{W}_\mathrm{tr}^{(M)^*}\mathbf{W}_\mathrm{tr}^{(M)}}^{\textbf{Known apriori at the Rx}}\}$$

- Cholesky decomposition of $\mathbf{C}_\mathrm{w} = \mathbf{D}_\mathrm{w}^*\mathbf{D}_\mathrm{w}$
- Noise whitened rx signal:

$$\begin{aligned}\mathbf{y}_\mathrm{w}[k] = \mathbf{D}_\mathrm{w}^{-*}\mathbf{y}[k] &= \mathbf{D}_\mathrm{w}^{-*}\boldsymbol{\Phi}\boldsymbol{\Psi}\mathbf{h}^\mathrm{v}[k] + \mathbf{D}_\mathrm{w}^{-*}\mathbf{n}[k] \\ &= \boldsymbol{\Phi}_\mathrm{w}\boldsymbol{\Psi}\mathbf{h}^\mathrm{v}[k] + \mathbf{n}_\mathrm{w}[k]\end{aligned}$$

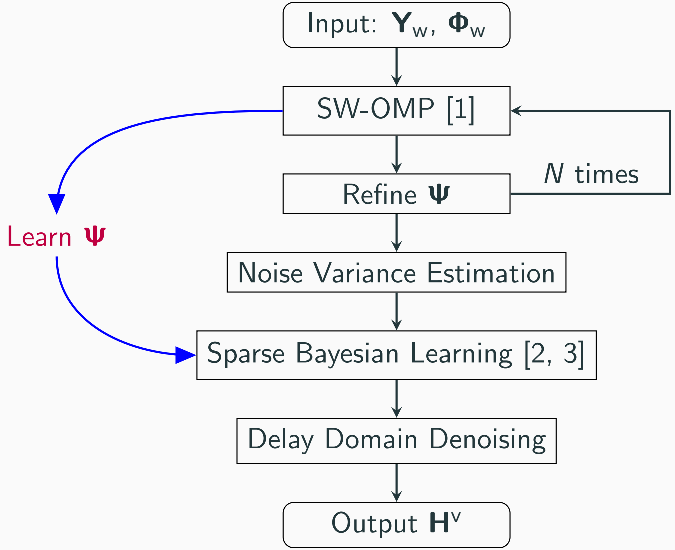
- Concatenated Rx signal of K subcarriers:

$$\begin{aligned}\mathbf{Y}_{\mathrm{w}} &= \begin{bmatrix}\mathbf{y}_{\mathrm{w}}[1] & \dots & \mathbf{y}_{\mathrm{w}}[K]\end{bmatrix} \\ &= \mathbf{\Phi}_{\mathrm{w}}\mathbf{\Psi}\begin{bmatrix}\mathbf{h}^{\mathrm{v}}[1] & \dots & \mathbf{h}^{\mathrm{v}}[K]\end{bmatrix} + \begin{bmatrix}\mathbf{n}_{\mathrm{w}}[1] & \dots & \mathbf{n}_{\mathrm{w}}[K]\end{bmatrix} \\ &= \mathbf{\Phi}_{\mathrm{w}}\mathbf{\Psi}\mathbf{H}^{\mathrm{v}} + \mathbf{N}_{\mathrm{w}}.\end{aligned}$$

- $\mathbf{H}^{\mathrm{v}}$ is **joint row sparse**
- **GOAL**: Estimate $\mathbf{\Psi}\mathbf{H}^{\mathrm{v}}$ given $\mathbf{Y}_{\mathrm{w}}$ and $\mathbf{\Phi}_{\mathrm{w}}$
- How do we initialize $\mathbf{\Psi}$?
  - Row truncated DFT matrices used as array steering matrices to generate $\mathbf{\Psi}$
    - Appropriate for uniform linear arrays
  - G_r and G_t set to 256

Proposed Solution

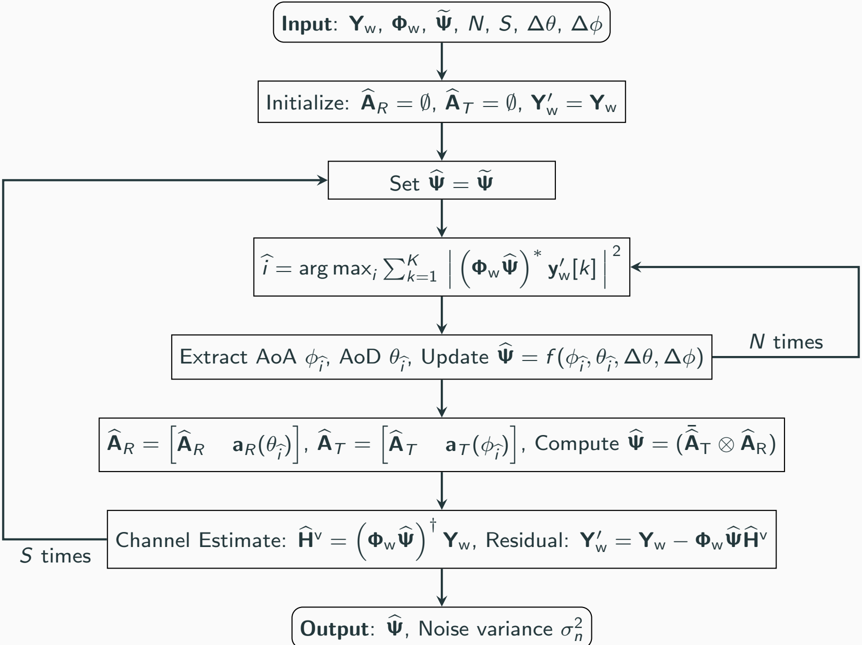
Outline of the Proposed Solution



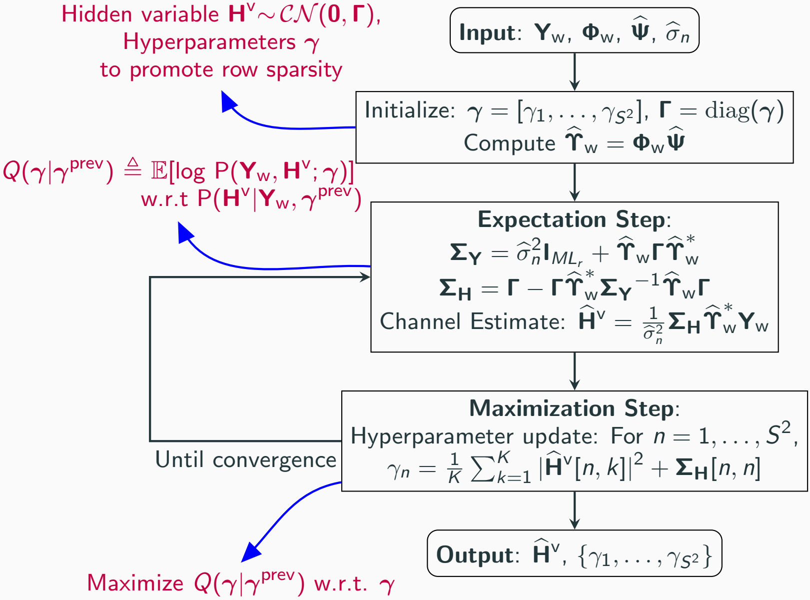
Important Aspects of the Proposed Solution

- **Fast** greedy search integrated with **high performing** sparse Bayesian learning (SBL)
  - **Dimensionality reduction** using SWOMP
  - **Expectation maximization** (EM) algorithm for estimation
  - **Delay domain thresholding** for channel denoising
- **Off-grid effects modelling** to recover AoD/AoD outside the quantized beamspace (spatial frequencies)
  - **Scope for further research** to provide mathematical guarantees
- **Guaranteed** convergence
- Exploits sparsity in both **beamspace** and **delay** domains

Stage 1: Multi-level Greedy Search (MLGS)



Stage 2: Sparse Bayesian Learning



NMSE Results on Test Datasets

Normalized Mean-Squared Error Table

SNR (dB)	[−20, −11)	[−11, −6)	[−6, 0]
Pilot Frames: 20	**−7.66 dB**	**−10.97 dB**	**−12.34 dB**
Pilot Frames: 40	−11.87 dB	−12.79 dB	−14.20 dB
Pilot Frames: 80	−13.62 dB	−16.23 dB	−20.08 dB

FINAL PERFORMANCE SCORE: −9.16 dB

References

References

[1] J. Rodríguez-Fernández, N. González-Prelcic, K. Venugopal, and R. W. Heath, "Frequency-domain compressive channel estimation for frequency-selective hybrid millimeter wave MIMO systems," *IEEE Trans. Wireless Commun.*, vol. 17, no. 5, pp. 2946–2960, 2018.

[2] M. E. Tipping, "Sparse Bayesian learning and the relevance vector machine," *J. Mach. Learn. Res.*, vol. 1, 2001.

[3] D. P. Wipf and B. D. Rao, "Sparse Bayesian learning for basis selection," *IEEE Trans. Signal Process.*, vol. 52, no. 8, Aug. 2004.

[4] https://github.com/ITU-AI-ML-in-5G-Challenge/ITU-ML5G-PS-025-Learned_Chester.

Thank You!

sameerah@iisc.ac.in, marios.kountouris@eurecom.fr,
christo.kurisummoottil@eurecom.fr,
rakesh.mundlamuri@eurecom.fr, cmurthy@iisc.ac.in,
thoota@iisc.ac.in