Subspace Estimation and Decomposition in Hybrid mmWave MIMO

H. Ghauch, T. Kim, M. Bengtsson, M. Skoglund

ACCESS Linnaeus Center, Royal Insititue of Technology
Department of Electronic Engineering, City University of Hong Kong

Automatic Control, June 4, 2015
# Table of contents

1. Overview of (mmWave) MIMO
2. Subspace Estimation
3. Subspace Decomposition
   - Multi-dimensional SD
   - The beamforming case
4. Subspace Estimation in Hybrid Architecture
   - Motivation
   - Echoing in the Hybrid Architecture
   - Proposed Algorithm

Based in parts on H. Ghauch, T. Kim, M. Bengtsson, M. Skoglund “Subspace Estimation and Decomposition for Large Millimeter-Wave MIMO systems”, submitted to J-STSP
Conventional MIMO systems

- $M \times N$ MIMO system can send $d$ data streams in parallel $d \leq \min(M, N)$
- Precoder $G \in \mathbb{C}^{M \times d}$, Combiner $U \in \mathbb{C}^{N \times d}$, Channel $H \in \mathbb{C}^{N \times M}$
- Rewrite SVD of $H$ as, $H = \Phi_1 \Sigma_1 \Gamma_1^\dagger + \Phi_2 \Sigma_2 \Gamma_2^\dagger$ where
  - $\Sigma_1$, diagonal, with $d$ largest right singular vectors
  - $\Gamma_1 \in \mathbb{C}^{M \times d} = d$ largest right singular vectors
  - $\Phi_1 \in \mathbb{C}^{N \times d} = d$ largest left singular vectors
Conventional MIMO systems

⋆ Transmission rate $R(G, U) = \log_2 |I + (U^\dagger HGG^\dagger H^\dagger U)(\sigma^2 U^\dagger U)^{-1}|$

MIMO channel capacity

Maximum transmission rate (well known)

$$(G^*, U^*) \triangleq \max R(G, U)$$

is such that $G^* = \Gamma_1 D$ and $U^* = \Phi_1$

$D$ diagonal, uniquely determined by $\Sigma_1$ (waterfilling).

⋆ Channel capacity only depends on $d$ dominant channel direction.

Tx needs to have $\Gamma_1$ and $\Sigma_1$

Rx needs to have $\Phi_1$
Conventional MIMO systems

- RF chain: frequency downconverter + analog-to-digital converter (ADC) + baseband processing
- Each transmit/receiver antenna is connected to one RF chain
- Obviously, larger M, N imply more data streams.

Issues with scaling up MIMO?

- **Size**: Minimum array size is inversely proportional to operating frequency. E.g. a 64 antenna array operating at 2 GHz, is $\sim 2.5\text{m} \times 2.5\text{m}$ (too large!)
- **Power consumption**: ADCs consume a lot of power ($\Rightarrow$ high operating costs for mobile operators).
- **Complexity** of baseband processing

- **massive MIMO**: MIMO system with large $M$.
- **millimeter-wave (mmWave) MIMO** circumvents the problem.
conversation between researcher and operator...

Why can’t we have 64 antennas at every Tx?

- 64 antenna array is too large (∼ 2.5m × 2.5)
- Power consumption is prohibitively large (since we need one RF chain per antenna)

No problem! Increase carrier frequency to 60 GHz, and reduce number of RF chains to say 4: array size reduced ∼ 30 times, power consumption reduced ∼ 16 times. But...
A Thought Experiment (A la Einstein)

Path loss is severe at 60 GHz
→ compensate by having lots of antennas (array gain)
→ no problem since antenna spacing small

Precoding cannot be fully digital (less RF chains than antennas)
→ offload part of the precoding to analog domain
→ hybrid precoding: precoder (resp. combiner) is a cascade of analog and digital filters

Gives rise to the idea of hybrid precoding for millimeter-wave MIMO systems
Pros / Cons

**Advantages**
- huge amounts of unused spectrum (~ 200 more than current cellular systems)
- most of the spectrum is unlicensed
- size of antenna array relatively small

**Challenges**
- channel statistics virtually non-existent
- no results on optimal transmission -
- previous channel estimation techniques, not applicable
- channel estimation should take into account sparsity of eigenmodes
System Model

- TDD $M \times N$ MIMO system, with $r$ RF chains, $d$ independent data streams, $d \leq r \leq \min(M, N)$
- Let $\mathcal{S}_{M,r} = \left\{ \mathbf{X} \in \mathbb{C}^{M \times r} \mid |X_{ik}| = 1/\sqrt{M}, \forall (i, k) \in \{1, \ldots, M\} \times \{1, \ldots, r\} \right\}$
- Analog precoder $\mathbf{F} \in \mathbb{C}^{M \times r}$, $\mathbf{F} \in \mathcal{S}_{M,r}$
- Analog combiner $\mathbf{W} \in \mathbb{C}^{N \times r}$, $\mathbf{W} \in \mathcal{S}_{N,r}$
- Channel $\mathbf{H} \in \mathbb{C}^{N \times M}$ has only $L$ paths ($L \ll \min(M, N)$)
- Rewrite SVD of $\mathbf{H}$ as, $\mathbf{H} = \Phi_1 \Sigma_1 \Gamma_1^\dagger + \Phi_2 \Sigma_2 \Gamma_2^\dagger$ where $\Gamma_1 \in \mathbb{C}^{M \times d}$ is $d$ largest right singular vectors, and $\Phi_1 \in \mathbb{C}^{N \times d}$ is $d$ largest left singular vectors
Motivation

It can be shown that $FG = \Gamma_1$ and $WU = \Phi_1$ maximize the transmission rate $R$. Similar to conventional MIMO, → optimal performance is determined by $\Gamma_1, \Phi_1$ (disregard $\Sigma_1$ for simplicity)
→ Tx needs to know $\Gamma_1$, Rx needs to know $\Phi_1$

Roadmap

I - Subspace Estimation (SE): since no a priori CSI assumed, need to estimate $\Gamma_1$ at Tx, and $\Phi_1$ at Rx (where $\Gamma_1 \triangleq v_{1:d}[H^\dagger H]$, $\Phi_1 \triangleq v_{1:d}[HH^\dagger]$).
II - Subspace Decomposition (SD): find a way to approximate $\Gamma_1$ by $FG$, and $\Phi_1$ by $WU$. 
The Arnoldi Iteration

We present **subspace estimation** (SE) method in context of conventional MIMO (focusing on estimating $\Gamma_1$ only)

Intimate relation between **eigenvalue algorithms** (numerical analysis) and SE in TDD MIMO.

→ **Power Method / Subspace Iteration** were used to devise algorithms for subspace estimation in MIMO

We use the well-known Arnoldi Iteration

For given $A \in \mathbb{C}^{M \times M}$, **Arnoldi Iteration** finds eigenpairs of $A$

→ iteratively finds basis $Q_m \triangleq [q_1, \ldots, q_m] \in \mathbb{C}^{M \times m}$ ($m < M$) ($q_1, \ldots, q_m$ are defined recursively) such that

$$Q_m^\dagger A Q_m = T_m, \quad Q_m^\dagger Q_m = I_m,$$

→ eigenpairs of $A$ by finding eigenpairs of $T_m$ (upper Hessenberg)
Goal is to use Arnoldi Iteration to find the eigenpairs of $A = H^\dagger H$ (which are nothing but $\Gamma_1$), at the Tx, in a fully distributed manner.

It is easily verified that the latter requires Tx to have $\{H^\dagger Hq_l\}_{l=1}^m$. Tx has $\{q_l\}_{l=1}^m$. How to compute $\{H^\dagger Hq_l\}_{l=1}^m$ with no CSIT? using transmitter-initiated echoing

For each vector $\{q_l\}_{l=1}^m$:

// Tx sends $q_l$ in the downlink (DL): $s_l = Hq_l + n_i^{(r)}$

// Rx does amplify-and-forward (A-F) of the received signal and sends signal back in the uplink (UL).

$$p_l = H^\dagger s_l + n_i^{(t)} = H^\dagger Hq_l + H^\dagger n_i^{(r)} + n_i^{(t)}$$

Now, Tx has a noisy estimate $p_l$ of $H^\dagger Hq_l$, $l = 1, \ldots, m$
Subspace Estimation using Arnoldi Iteration (SE-ARN)

★ SE-ARN estimate of $\Gamma_1$ at the Tx (using transmitter-initiated echoing), and $\Phi_1$ at Rx, in a fully distributed way.
Problem Formulation

Once $\Gamma_1$ is estimated, we need a method to approximate (decompose) $\Gamma_1$, into $FG$, at the Tx, i.e., $^1$

$$
\begin{cases}
\min_{F,G} h_0(F,G) = \|\tilde{\Gamma}_1 - FG\|_F^2 \\
\text{s. t. } h_1(F,G) = \|FG\|_F^2 \leq d, \quad F \in S_{M,d}
\end{cases}
$$

(1)

Use **Block Coordinate Descent (BCD)** approach (due to coupling)
$\rightarrow$ fix $G_k$ and optimize $F_k, \forall k$ (and vice versa) iteratively
$\rightarrow$ power constraint is implicitly enforced and can be omitted

$$(J1) \quad F_{k+1} \triangleq \min_F h_0(F) = \|\tilde{\Gamma}_1 - FG_k\|_F^2 \quad \text{s. t. } F \in S_{M,d}$$

$$(J2) \quad G_{k+1} \triangleq \min_G h_0(G) = \|\tilde{\Gamma}_1 - F_{k+1}G\|_F^2$$

$^1$Similarly, $\Phi_1$ needs to be decomposed into $WU$ at Rx.
BCD for Subspace Decomposition

(J1) : Relax constraint $F \in S_{M,d}$, solve resulting convex problem, and project solution on $S_{M,d}$.

$$F_{k+1} = \Pi_S \left[ \tilde{\Gamma}_1 G_l^\dagger (G_k G_k^\dagger)^{-1} \right]$$  \hspace{1cm} (2)

where $\Pi_S[X] \triangleq \arg\min_{U \in S_{M,d}} \| U - X \|_F^2$ is the Euclidean projection on $S_{M,d}$

(J2) : Unconstrained convex problem,

$$G_{k+1} = (F_{k+1}^\dagger F_{k+1})^{-1} F_{k+1}^\dagger \tilde{\Gamma}_1$$  \hspace{1cm} (3)

Block Coordinate Descent for Subspace Decomposition (BCD-SD)

Start with arbitrary $G_0$

for $k = 0, 1, 2, \ldots$ do

$F_{k+1} \leftarrow \Pi_S \left[ \tilde{\Gamma}_1 G_k^\dagger (G_k G_k^\dagger)^{-1} \right]$; $G_{l+1} \leftarrow F_{k+1}^\dagger F_{k+1}^{-1} F_{k+1}^\dagger \tilde{\Gamma}_1$

end for
The beamforming case

When \( d = 1 \), the SD problem is,

\[
\begin{align*}
\min_{\mathbf{f}, \mathbf{g}} h_o(\mathbf{f}, \mathbf{g}) &= \|\mathbf{f}\|_2^2 g^2 - 2g \Re(\mathbf{f}^\dagger \gamma_1) \\
\text{s. t. } [\mathbf{f}]_i &= 1/\sqrt{M} \ e^{i\phi_i}, \forall i
\end{align*}
\]

(4)

where \( g \in \mathbb{R}_+ \) and \([\gamma_1]_i = r_i e^{i\theta_i} \). The problem admits a globally optimum solution given by,

\[
[\mathbf{f}^*]_i = 1/\sqrt{M} \ e^{i\theta_i}, \forall i \text{ and } g^* = \|\tilde{\gamma}_1\|_1/\sqrt{M}
\]

Decomposing a vector \( \gamma_1 \) is extremely simple:

- set phase of \( \mathbf{f} \) as \( \phi_i = \arg([\gamma_1]_i), \forall i \)
- set \( g = \|\tilde{\gamma}_1\|_1/\sqrt{M} \)
Recall that the sequence $\{H^\dagger Hq_i\}_{i=1}^m$ is all that is needed at Tx to use Arnoldi Iteration for subspace estimation (in SE-ARN estimates of the latter were obtained using transmitter-initiated echoing).

Echoing $\iff$ Amplify-and-Forward at receiver.
This cannot be done in the hybrid architecture because the received signal at antennas cannot be digitally processed (as in conventional MIMO).

A 'blind' application of the original echoing mechanism to estimate $\{H^\dagger Hq_i\}_{i=1}^m$ yields a very poor estimate,

$$\{F_i^\dagger H^\dagger W_i W_i^\dagger Hq_i\}_{i=1}^m$$

Need to suppress the effects of processing with $W_i$ and $F_i$,
DL phase of echoing mechanism

- decompose $q_l$ at Tx using SD: $q_l = f_l g_l + e_l^{(t)}$ ($e_l^{(r)}$ is the decomposition error)
- send $f_l g_l$ in the DL, $K_r$ times where $K_r \triangleq N/r$
- process each received signal with $\{W_{l,k}\}_{k=1}^{K_r}$ to obtain digital samples $\{s_{l,k} = W_{l,k}^\dagger H f_l g_l\}_{k=1}^{K_r}$

$\{W_{l,k}\}_{k=1}^{K_r}$ are taken from the columns of a Discrete Fourier Transform (DFT) matrix, $D_r: [W_{l,1}, \ldots, W_{l,K_r}] = D_r$

- use the same sequence of analog filters, $\{W_{l,k}\}_{k=1}^{K_r}$, to process $\{s_{l,k}\}_{k=1}^{K_r}$: $\tilde{s}_l \triangleq \sum_{k=1}^{K_r} W_{l,k} s_{l,k}$
Rewrite received signal \( \tilde{s}_l \) as,

\[
\tilde{s}_l \triangleq \sum_{k=1}^{K_r} W_{l,k} s_{l,k} = \left( \sum_{k=1}^{K_r} W_{l,k} W_{l,k}^\dagger \right) H f_{l,g_l} = H(q_l - e^{(r)}_l)
\]

This echoing method removes the effect of analog combiner \( W_l \) by using repetition to obtain multiple measurements of the received signal by carefully choosing \( \{W_{l,k}\}_{k=1}^{K_r} \) as columns of DFT matrices.

The same repetition-based method removes the effect of \( F_l \):

**Repetition-Aided (RAID) echoing**

// DL phase
\[
q_l = f_{l,g_l} + e^{(t)}_l
\]
\[
s_{l,k} = W_{l,k}^\dagger H f_{l,g_l}, \ \forall k \in \{K_r \triangleq N/r\}
\]
\[
\tilde{s}_l = \sum_{k=1}^{K_r} W_{l,k} s_{l,k}
\]

// UL phase
\[
z_{l,m} = F_{l,m}^\dagger H^\dagger w_{l,u_l}, \ \forall m \in \{K_t \triangleq M/r\}
\]
\[
p_l = \sum_{m=1}^{K_t} F_{l,m} z_{l,m}
\]
Repetition-Aided (RAID) echoing

$$q_l \approx (f \cdot g)_l$$

Combine
Compute $$\tilde{s}_l$$

$$K_r$$ times

Combine
Compute $$p_l$$

$$K_t$$ times

Subspace Estimation for Hybrid Architecture

Replace echoing procedure in SE-ARN with RAID echoing
Subspace Estimation and Decomposition (SED) for Hybrid Architecture

// Estimate $\Gamma_1$ and $\Phi_1$
$\Gamma_1 = \text{SE-ARN} (H, m)$
$\Phi_1 = \text{SE-ARN} (H^\dagger, m)$

// Decompose $\Gamma_1$ and $\Phi_1$
$[F, G] = \text{BCD-SD} (\Gamma_1, \rho)$
$[W, U] = \text{BCD-SD} (\Phi_1, \rho)$

where $m$ is number of iteration of Arnoldi iteration.

Total communication overhead required by the algorithm is
$\Omega = 2m(M + N)/r$ channel uses,
Simulation Setup

Realistic propagation using SCM channels (3GPP), $M = 64$, $N = 32$, $r = 8$, $m = 2d$, for several values of $d$. Performance metric:

$$R = \log_2 \left| I_d + \frac{1}{\sigma_f^2} U^\dagger W^\dagger H F G G^\dagger F^\dagger H^\dagger W U (U^\dagger W^\dagger W U)^{-1} \right|$$

We use Adaptive Channel Estimation from [Alkhateeb’14] as benchmark.

- Estimate the mmWave MIMO channel, one path at a time

We compare against an **ideal upper bound**, $R^*$

- The capacity of an equivalent conventional MIMO system
- Not known if this $R^*$ is achievable
Significant gains over the benchmark, especially in low-SNR region.
Investigate scalability and performance by scaling up $M, N$, i.e., $N = M/2$, $r = M/8$, $d = 2$, $m = 6$, $\Omega = 144$.

Performance very close to fully digital case, but with $\sim 8$ to $\sim 16$ times less RF chains.
Recap

- overview of mmWave MIMO systems
- motivated the problem of subspace estimation in hybrid mmWave MIMO systems
- presented a Krylov Subspace method to estimate the left / right singular subspace, in a fully distributed manner
- motivated the subspace decomposition problem and presented an iterative method for that

Hybrid architecture has similar performance at its fully digital counterpart, however, with a drastically reduced number of RF chains (≈ 8 to ≈ 16 times less).
Thank you! Questions?