Channel Estimation Using Joint Dictionary Learning in FDD Massive MIMO Systems

Yacong Ding and Bhaskar Rao

Department of Electrical and Computer Engineering
University of California, San Diego

December 14, 2015

Outline

Introduction

Motivation

- Acquiring downlink channel information h^d at base station (CSIT):
  - TDD: easy, FDD: difficult.
  - BS has N antennas, user has 1 antennas: $h^d, h^u \in \mathbb{C}^{N\times 1}$

	Training	Feedback
TDD: $h^d = h^u$, uplink training (UE: $\phi \in \mathbb{C}^{1\times T^u}$, BS: $Y^u = h^u\phi + n^u$)	$T^u \geq 1$	No
FDD: $h^d \neq h^u$, downlink training (BS: $\Phi \in \mathbb{C}^{T^d\times N}$, UE: $Y^d = \Phi h^d + n^d$)	$T^d \geq N$	$\propto N$

Motivation

- Acquiring downlink channel information h^d at base station (CSIT):
  - TDD: easy, FDD: difficult.
  - BS has N antennas, user has 1 antennas: $h^d, h^u \in \mathbb{C}^{N\times 1}$

	Training	Feedback
TDD: $h^d = h^u$, uplink training (UE: $\phi \in \mathbb{C}^{1\times T^u}$, BS: $Y^u = h^u\phi + n^u$)	$T^u \geq 1$	No
FDD: $h^d \neq h^u$, downlink training (BS: $\Phi \in \mathbb{C}^{T^d\times N}$, UE: $Y^d = \Phi h^d + n^d$)	$T^d \geq N$	$\propto N$

- FDD is widely employed in existing communication systems:
  - Beneficial if directly adopt Massive MIMO to FDD.

Motivation

- Acquiring downlink channel information h^d at base station (CSIT):
  - TDD: easy, FDD: difficult.
  - BS has N antennas, user has 1 antennas: $h^d, h^u \in \mathbb{C}^{N \times 1}$

	Training	Feedback
TDD: $h^d = h^u$, uplink training (UE: $\phi \in \mathbb{C}^{1 \times T^u}$, BS: $Y^u = h^u \phi + n^u$)	$T^u \geq 1$	No
FDD: $h^d \neq h^u$, downlink training (BS: $\Phi \in \mathbb{C}^{T^d \times N}$, UE: $Y^d = \Phi h^d + n^d$)	$T^d \geq N$	$\propto N$

- FDD is widely employed in existing communication systems:
  - Beneficial if directly adopt Massive MIMO to FDD.
- FDD Downlink training: $Y^d = \Phi h^d + n^d, \Phi \in \mathbb{C}^{T^d \times N}$:
  - To be practical: T^d small.

Motivation

- Acquiring downlink channel information h^d at base station (CSIT):
  - TDD: easy, FDD: difficult.
  - BS has N antennas, user has 1 antennas: $h^d, h^u \in \mathbb{C}^{N \times 1}$

	Training	Feedback
TDD: $h^d = h^u$, uplink training (UE: $\phi \in \mathbb{C}^{1 \times T^u}$, BS: $Y^u = h^u \phi + n^u$)	$T^u \geq 1$	No
FDD: $h^d \neq h^u$, downlink training (BS: $\Phi \in \mathbb{C}^{T^d \times N}$, UE: $Y^d = \Phi h^d + n^d$)	$T^d \geq N$	$\propto N$

- FDD is widely employed in existing communication systems:
  - Beneficial if directly adopt Massive MIMO to FDD.
- FDD Downlink training: $Y^d = \Phi h^d + n^d, \Phi \in \mathbb{C}^{T^d \times N}$:
  - To be practical: T^d small.
  - $T^d < N$: underdetermined inverse problem, infinite solutions.

Motivation

- Acquiring downlink channel information h^d at base station (CSIT):
  - TDD: easy, FDD: difficult.
  - BS has N antennas, user has 1 antennas: $h^d, h^u \in \mathbb{C}^{N\times 1}$

	Training	Feedback
TDD: $h^d = h^u$, uplink training (UE: $\phi \in \mathbb{C}^{1\times T^u}$, BS: $Y^u = h^u\phi + n^u$)	$T^u \geq 1$	No
FDD: $h^d \neq h^u$, downlink training (BS: $\Phi \in \mathbb{C}^{T^d\times N}$, UE: $Y^d = \Phi h^d + n^d$)	$T^d \geq N$	$\propto N$

- FDD is widely employed in existing communication systems:
  - Beneficial if directly adopt Massive MIMO to FDD.
- FDD Downlink training: $Y^d = \Phi h^d + n^d, \Phi \in \mathbb{C}^{T^d\times N}$:
  - To be practical: T^d small.
  - $T^d < N$: underdetermined inverse problem, infinite solutions.
- Explore channel structure to regularize the problem?

Motivation

- Acquiring downlink channel information h^d at base station (CSIT):
  - TDD: easy, FDD: difficult.
  - BS has N antennas, user has 1 antennas: $h^d, h^u \in \mathbb{C}^{N \times 1}$

	Training	Feedback
TDD: $h^d = h^u$, uplink training (UE: $\phi \in \mathbb{C}^{1 \times T^u}$, BS: $Y^u = h^u \phi + n^u$)	$T^u \geq 1$	No
FDD: $h^d \neq h^u$, downlink training (BS: $\Phi \in \mathbb{C}^{T^d \times N}$, UE: $Y^d = \Phi h^d + n^d$)	$T^d \geq N$	$\propto N$

- FDD is widely employed in existing communication systems:
  - Beneficial if directly adopt Massive MIMO to FDD.
- FDD Downlink training: $Y^d = \Phi h^d + n^d$,$\Phi \in \mathbb{C}^{T^d \times N}$:
  - To be practical: T^d small.
  - $T^d < N$: underdetermined inverse problem, infinite solutions.
- Explore channel structure to regularize the problem?
  - **Sparse** channel structure: **compressive sensing**.

Compressed Channel Estimation

- Low dimensional representation of high dimensional signal:
  - Find a Ψ such that $h^d = \Psi \beta^d, \|\beta^d\|_0 < N$.

Compressed Channel Estimation

- Low dimensional representation of high dimensional signal:
  - Find a Ψ such that $h^d = \Psi\beta^d, \|\beta^d\|_0 < N$.
  - Downlink training: $Y^d = \Phi h^d + n^d = \Phi\Psi\beta^d + n^d$

Compressed Channel Estimation

- Low dimensional representation of high dimensional signal:
  - Find a Ψ such that $h^d = \Psi\beta^d, \|\beta^d\|_0 < N$.
  - Downlink training: $Y^d = \Phi h^d + n^d = \Phi\Psi\beta^d + n^d$
- Apply **compressive sensing** algorithm to estimate the sparse coefficient β^d.

Compressed Channel Estimation :

$$\hat{\beta}^d = \arg\min_{\beta^d} \|\beta^d\|_0 \text{ subject to } \|Y^d - \Phi\Psi\beta^d\|_2^2 \leq \sigma^2 \quad (1)$$

$$\hat{h}^d = \Psi\hat{\beta}^d$$

Compressed Channel Estimation

- Low dimensional representation of high dimensional signal:
  - Find a Ψ such that $h^d = \Psi\beta^d, \|\beta^d\|_0 < N$.
  - Downlink training: $Y^d = \Phi h^d + n^d = \Phi\Psi\beta^d + n^d$
- Apply **compressive sensing** algorithm to estimate the sparse coefficient β^d.

Compressed Channel Estimation :

$$\hat{\beta}^d = \arg\min_{\beta^d} \|\beta^d\|_0 \text{ subject to } \|Y^d - \Phi\Psi\beta^d\|_2^2 \leq \sigma^2 \qquad (1)$$

$$\hat{h}^d = \Psi\hat{\beta}^d$$

- Many practical algorithms. **Measurements**: $T^d \propto \|\beta^d\|_0 < N$

Compressed Channel Estimation

- Low dimensional representation of high dimensional signal:
  - Find a Ψ such that $h^d = \Psi\beta^d, \|\beta^d\|_0 < N$.
  - Downlink training: $Y^d = \Phi h^d + n^d = \Phi\Psi\beta^d + n^d$
- Apply **compressive sensing** algorithm to estimate the sparse coefficient β^d.

Compressed Channel Estimation :

$$\hat{\beta}^d = \arg\min_{\beta^d} \|\beta^d\|_0 \text{ subject to } \|Y^d - \Phi\Psi\beta^d\|_2^2 \leq \sigma^2 \qquad (1)$$

$$\hat{h}^d = \Psi\hat{\beta}^d$$

- Many practical algorithms. **Measurements**: $T^d \propto \|\beta^d\|_0 < N$
- Core requirement: find Ψ.

Sparse Channel Representation

Drawbacks of Existing Work

The existing work that applies compressed channel estimation use **orthogonal** DFT basis as Ψ:

- Agree with array manifold using ULA.
- **Infinite** number of antennas, **limited** scattering environment.

Drawbacks of Existing Work

The existing work that applies compressed channel estimation use **orthogonal** DFT basis as Ψ:

- Agree with array manifold using ULA.
- **Infinite** number of antennas, **limited** scattering environment.

For common channels models, such as 3GPP SCM channels:

- High $\|\beta^d\|_0$: $h^d = \Psi_{DFT}\beta^d$
- $T^d \propto \|\beta^d\|_0$: lose benefits of compressive sensing

Drawbacks of Existing Work

The existing work that applies compressed channel estimation use **orthogonal** DFT basis as Ψ:

- Agree with array manifold using ULA.
- **Infinite** number of antennas, **limited** scattering environment.

For common channels models, such as 3GPP SCM channels:

- High $\|\beta^d\|_0$: $h^d = \Psi_{DFT}\beta^d$
- $T^d \propto \|\beta^d\|_0$: lose benefits of compressive sensing

One easy better choice is **overcomplete** DFT matrix: redundancy in basis

- Only applicable to **ULA**.
- Can not adapt to **specific** cell characteristics: urban, rural, hills, etc.

Drawbacks of Existing Work

The existing work that applies compressed channel estimation use **orthogonal** DFT basis as Ψ:

- Agree with array manifold using ULA.
- **Infinite** number of antennas, **limited** scattering environment.

For common channels models, such as 3GPP SCM channels:

- High $\|\beta^d\|_0$: $h^d = \Psi_{DFT}\beta^d$
- $T^d \propto \|\beta^d\|_0$: lose benefits of compressive sensing

One easy better choice is **overcomplete** DFT matrix: redundancy in basis

- Only applicable to **ULA**.
- Can not adapt to **specific** cell characteristics: urban, rural, hills, etc.

How can we design a dictionary/basis such that:

Drawbacks of Existing Work

The existing work that applies compressed channel estimation use **orthogonal** DFT basis as Ψ:

- Agree with array manifold using ULA.
- **Infinite** number of antennas, **limited** scattering environment.

For common channels models, such as 3GPP SCM channels:

- High $\|\beta^d\|_0$: $h^d = \Psi_{DFT}\beta^d$
- $T^d \propto \|\beta^d\|_0$: lose benefits of compressive sensing

One easy better choice is **overcomplete** DFT matrix: redundancy in basis

- Only applicable to **ULA**.
- Can not adapt to **specific** cell characteristics: urban, rural, hills, etc.

How can we design a dictionary/basis such that:

- Adapt to **specific** cell properties (antenna, environment).

Drawbacks of Existing Work

The existing work that applies compressed channel estimation use **orthogonal** DFT basis as Ψ:

- Agree with array manifold using ULA.
- **Infinite** number of antennas, **limited** scattering environment.

For common channels models, such as 3GPP SCM channels:

- High $\|\beta^d\|_0$: $h^d = \Psi_{DFT}\beta^d$
- $T^d \propto \|\beta^d\|_0$: lose benefits of compressive sensing

One easy better choice is **overcomplete** DFT matrix: redundancy in basis

- Only applicable to **ULA**.
- Can not adapt to **specific** cell characteristics: urban, rural, hills, etc.

How can we design a dictionary/basis such that:

- Adapt to **specific** cell properties (antenna, environment).
- Lead to **sparse** representation.

Our Previous Work: Learning Good Representation

- Rather than using **predefined** dictionary/basis, **learn** *cell specific* D^d from data:

Our Previous Work: Learning Good Representation

- Rather than using **predefined** dictionary/basis, **learn** *cell specific* D^d from data:
  - **Overcomplete:** $D^d \in \mathbb{C}^{N \times M}, N < M$

Our Previous Work: Learning Good Representation

- Rather than using **predefined** dictionary/basis, **learn** *cell specific* D^d from data:
  - **Overcomplete:** $D^d \in \mathbb{C}^{N \times M}, N < M$
  - **Fit** model to data: $h_i^d \approx D^d \beta_i^d, i = 1, \ldots, L$

Our Previous Work: Learning Good Representation

- Rather than using **predefined** dictionary/basis, **learn** *cell specific* D^d from data:
  - **Overcomplete:** $D^d \in \mathbb{C}^{N \times M}, N < M$
  - **Fit** model to data: $h_i^d \approx D^d \beta_i^d, i = 1, \ldots, L$
  - **Encourage** sparsity: $\|\beta_i^d\|_0 \ll M, \forall i.$

Our Previous Work: Learning Good Representation

- Rather than using **predefined** dictionary/basis, **learn** *cell specific* D^d from data:
  - **Overcomplete:** $D^d \in \mathbb{C}^{N \times M}, N < M$
  - **Fit** model to data: $h_i^d \approx D^d \beta_i^d, i = 1, \ldots, L$
  - **Encourage** sparsity: $\|\beta_i^d\|_0 \ll M, \forall i.$

- What data can be utilized?

Our Previous Work: Learning Good Representation

- Rather than using **predefined** dictionary/basis, **learn** *cell specific* D^d from data:
  - **Overcomplete:** $D^d \in \mathbb{C}^{N \times M}, N < M$
  - **Fit** model to data: $h_i^d \approx D^d \beta_i^d, i = 1, \ldots, L$
  - **Encourage** sparsity: $\|\beta_i^d\|_0 \ll M, \forall i.$

- What data can be utilized?
  - Channel measurements: collected within a **specific** cell.

Our Previous Work: Learning Good Representation

- Rather than using **predefined** dictionary/basis, **learn** *cell specific* D^d from data:
  - **Overcomplete:** $D^d \in \mathbb{C}^{N \times M}, N < M$
  - **Fit** model to data: $h_i^d \approx D^d \beta_i^d, i = 1, \ldots, L$
  - **Encourage** sparsity: $\|\beta_i^d\|_0 \ll M, \forall i.$

- What data can be utilized?
  - Channel measurements: collected within a **specific** cell.
  - Effect of **environment** on the transmitted electromagnetic waves represented at **antennas**.

Our Previous Work: Learning Good Representation

- Rather than using **predefined** dictionary/basis, **learn** *cell specific* D^d from data:
  - **Overcomplete:** $D^d \in \mathbb{C}^{N \times M}, N < M$
  - **Fit** model to data: $h_i^d \approx D^d \beta_i^d, i = 1, \ldots, L$
  - **Encourage** sparsity: $\|\beta_i^d\|_0 \ll M, \forall i$.
- What data can be utilized?
  - Channel measurements: collected within a **specific** cell.
  - Effect of **environment** on the transmitted electromagnetic waves represented at **antennas**.
  - **Big data** paradigm in wireless communication.

Dictionary Learning

- Combine data fitting and sparsity encouragement, dictionary learning can be formulated:

$$\min_{D^d, \beta_1^d, \dots, \beta_L^d} \lambda \| H^d - D^d B^d \|_F^2 + \sum_{i=1}^{L} \| \beta_i^d \|_0 \qquad (2)$$

where $H^d = [h_1^d, \dots, h_L^d] \quad B^d = [\beta_1^d, \dots, \beta_L^d]$.

Dictionary Learning

- Combine data fitting and sparsity encouragement, dictionary learning can be formulated:

$$\min_{D^d, \beta_1^d, \ldots, \beta_L^d} \lambda \| H^d - D^d B^d \|_F^2 + \sum_{i=1}^{L} \| \beta_i^d \|_0 \tag{2}$$

where $H^d = [h_1^d, \ldots, h_L^d] \quad B^d = [\beta_1^d, \ldots, \beta_L^d]$.

- Benefits of dictionary learning and compressed channel estimation:
  - Applicable to **any antenna configurations**: no assumed structure.

Dictionary Learning

- Combine data fitting and sparsity encouragement, dictionary learning can be formulated:

$$\min_{D^d,\beta_1^d,\ldots,\beta_L^d} \lambda\|H^d - D^d B^d\|_F^2 + \sum_{i=1}^{L} \|\beta_i^d\|_0 \tag{2}$$

where $H^d = [h_1^d, \ldots, h_L^d] \quad B^d = [\beta_1^d, \ldots, \beta_L^d]$.

- Benefits of dictionary learning and compressed channel estimation:
  - Applicable to **any antenna configurations**: no assumed structure.
  - Robust to any **irregularities**: mismatched antennas, non-plane wave.

Dictionary Learning

- Combine data fitting and sparsity encouragement, dictionary learning can be formulated:

$$\min_{D^d,\beta_1^d,\ldots,\beta_L^d} \lambda\|H^d - D^d B^d\|_F^2 + \sum_{i=1}^{L} \|\beta_i^d\|_0 \tag{2}$$

where $H^d = [h_1^d, \ldots, h_L^d] \quad B^d = [\beta_1^d, \ldots, \beta_L^d]$.

- Benefits of dictionary learning and compressed channel estimation:
  - Applicable to **any antenna configurations**: no assumed structure.
  - Robust to any **irregularities**: mismatched antennas, non-plane wave.
  - Training and feedback overhead: proportional to **channel sparsity** S.

Joint Uplink/Downlink Dictionary Learning and Compressed Channel Estimation

Utilizing Uplink Channel Information

In compressive sensing, more measurements are always better:

- More information about the underlying sparse coefficients.
- Better recovery performance.

Utilizing Uplink Channel Information

In compressive sensing, more measurements are always better:

- More information about the underlying sparse coefficients.
- Better recovery performance.

In Massive MIMO, it implies T^d to be larger:

- Larger training and feedback overhead.
- Waste of resources.

Utilizing Uplink Channel Information

In compressive sensing, more measurements are always better:

- More information about the underlying sparse coefficients.
- Better recovery performance.

In Massive MIMO, it implies T^d to be larger:

- Larger training and feedback overhead.
- Waste of resources.

Is it possible to have more information about the underlying sparse coefficient, but without need of larger T^d?

Utilizing Uplink Channel Information

In compressive sensing, more measurements are always better:

- More information about the underlying sparse coefficients.
- Better recovery performance.

In Massive MIMO, it implies T^d to be larger:

- Larger training and feedback overhead.
- Waste of resources.

Is it possible to have more information about the underlying sparse coefficient, but without need of larger T^d?

From uplink channel h^u:

Utilizing Uplink Channel Information

In compressive sensing, more measurements are always better:

- More information about the underlying sparse coefficients.
- Better recovery performance.

In Massive MIMO, it implies T^d to be larger:

- Larger training and feedback overhead.
- Waste of resources.

Is it possible to have more information about the underlying sparse coefficient, but without need of larger T^d?

From uplink channel h^u:

- Easy to obtain: $T^u \geq 1$.
- Common sparse channel structure between h^d and h^u.

Joint Uplink/Downlink Channel Representation

- Similar to $h^d = D^d \beta^d$: $h^u = D^u \beta^u$.

Joint Uplink/Downlink Channel Representation

- Similar to $h^d = D^d \beta^d$: $h^u = D^u \beta^u$.
- Duplex distance not large: **similar scattering effect** for uplink and downlink transmission.

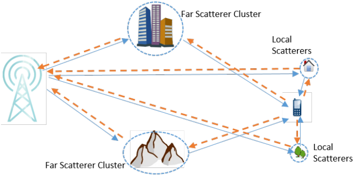

Figure 1: Uplink/Downlink Channel Model

Joint Uplink/Downlink Channel Representation

- Similar to $h^d = D^d \beta^d$: $h^u = D^u \beta^u$.
- Duplex distance not large: **similar scattering effect** for uplink and downlink transmission.

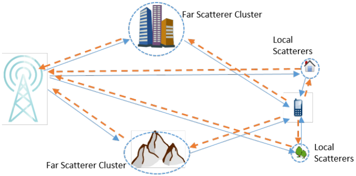

Figure 1: Uplink/Downlink Channel Model

- In our model, equivalently to assume $\chi(\beta^u) = \chi(\beta^d)$, where $\chi(\beta) = \{i | \beta(i) \neq 0\}$ denotes the set of locations of nonzero entries in β.

Joint Uplink/Downlink Dictionary Learning

$$\begin{aligned} \min_{D^u,B^u,D^d,B^d} \quad & \|H^u - D^u B^u\|_F^2 + \|H^d - D^d B^d\|_F^2 \\ \textit{subject to} \quad & \|\beta_i^u\|_0 = \|\beta_i^d\|_0 \leq T_0, \;\; \chi(\beta_i^u) = \chi(\beta_i^d) \;\; \forall i \end{aligned} \tag{3}$$

Joint UL/DL Dictionary Learning

Joint Uplink/Downlink Dictionary Learning

$$\begin{aligned} \min_{D^u, B^u, D^d, B^d} \quad & \|H^u - D^u B^u\|_F^2 + \|H^d - D^d B^d\|_F^2 \\ \textit{subject to} \quad & \|\beta_i^u\|_0 = \|\beta_i^d\|_0 \leq T_0, \;\; \chi(\beta_i^u) = \chi(\beta_i^d) \;\; \forall i \end{aligned} \tag{3}$$

Joint Uplink/Downlink Sparse Representation

$$\begin{aligned} & h^u \approx D^u \beta^u, h^d \approx D^d \beta^d \\ & \|\beta_i^u\|_0 = \|\beta_i^d\|_0 \leq T_0, \;\; \chi(\beta_i^u) = \chi(\beta_i^d) \;\; \forall i \end{aligned} \tag{4}$$

Joint UL/DL Dictionary Learning

Joint Uplink/Downlink Dictionary Learning

$$\begin{aligned} \min_{D^u,B^u,D^d,B^d} \quad & \|H^u - D^u B^u\|_F^2 + \|H^d - D^d B^d\|_F^2 \\ \textit{subject to} \quad & \|\beta_i^u\|_0 = \|\beta_i^d\|_0 \leq T_0, \;\; \chi(\beta_i^u) = \chi(\beta_i^d) \;\; \forall i \end{aligned} \tag{3}$$

Joint Uplink/Downlink Sparse Representation

$$\begin{aligned} & h^u \approx D^u \beta^u, h^d \approx D^d \beta^d \\ & \|\beta_i^u\|_0 = \|\beta_i^d\|_0 \leq T_0, \;\; \chi(\beta_i^u) = \chi(\beta_i^d) \;\; \forall i \end{aligned} \tag{4}$$

Joint Uplink/Downlink Compressed Channel Estimation :

$$\begin{aligned} & \arg\min_{\beta^u,\beta^d} \|Y^u - \phi^u D^u \beta^u\|_2^2 + \|Y^d - \Phi^d D^d \beta^d\|_2^2 \\ & \textit{subject to} \;\; \chi(\beta^u) = \chi(\beta^d), \;\; \|\beta^u\|_0 = \|\beta^d\|_0 \leq T_0 \end{aligned} \tag{5}$$

Benefits of Joint Sparse Framework

- Joint dictionary learning:
  - Regularize the learning process.
  - Better performance when underlying generative model satisfies joint sparsity.

Benefits of Joint Sparse Framework

- Joint dictionary learning:
  - Regularize the learning process.
  - Better performance when underlying generative model satisfies joint sparsity.
- Joint channel estimation:
  - Better recovery: additional measurements from uplink training.

Benefits of Joint Sparse Framework

- Joint dictionary learning:
  - Regularize the learning process.
  - Better performance when underlying generative model satisfies joint sparsity.
- Joint channel estimation:
  - Better recovery: additional measurements from uplink training.
- In other words, we can further decrease downlink training duration T^d with the same performance.

Numerical Results

Simulation Settings

- Apply 3GPP SCM: far scatterer clusters and local scatterer clusters.

Simulation Settings

- Apply 3GPP SCM: far scatterer clusters and local scatterer clusters.
- Each channel snapshot:
  - 4 local SC: locations change with user.
  - 2 far SC: fixed locations.

Simulation Settings

- Apply 3GPP SCM: far scatterer clusters and local scatterer clusters.
- Each channel snapshot:
  - 4 local SC: locations change with user.
  - 2 far SC: fixed locations.
- Training samples: 50000 channel snapshots uniformly sampled in the cell.

Simulation Settings

- Apply 3GPP SCM: far scatterer clusters and local scatterer clusters.
- Each channel snapshot:
    - 4 local SC: locations change with user.
    - 2 far SC: fixed locations.
- Training samples: 50000 channel snapshots uniformly sampled in the cell.
- 100 antennas at base station and 1 antenna at user. Apply uniform linear array.

Simulation Settings

- Apply 3GPP SCM: far scatterer clusters and local scatterer clusters.
- Each channel snapshot:
  - 4 local SC: locations change with user.
  - 2 far SC: fixed locations.
- Training samples: 50000 channel snapshots uniformly sampled in the cell.
- 100 antennas at base station and 1 antenna at user. Apply uniform linear array.
- Pair of uplink/downlink channel: same angles, different amplitudes and phases.

Low Dimension Representation

Dictionary Learning in Channel Representation

Constrain T_0 atoms to be used. Compare MSE($E\|h^d - \hat{h}^d\|_2^2$) between h^d and $\hat{h}^d = D^d \beta^d$. $\|h^d\|_2 = 1$.

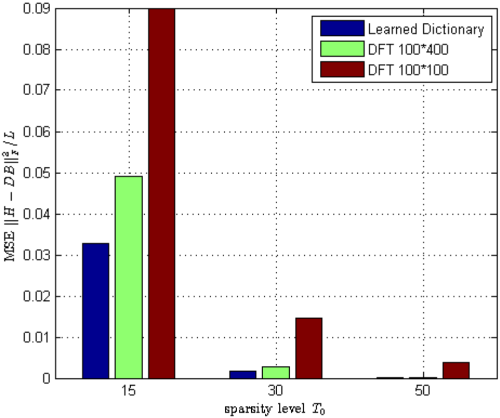

Figure 2: MSE comparison.

Compressed Channel Estimation

Dictionary Learning in Joint UL/DL Channel Estimation

Compare MSE between h^d and $\hat{h}^d = D^d \hat{\beta}^d$. $\hat{\beta}^d = \text{OMP}(Y^d, \Phi, D^d)$, or $\hat{\beta}^d = \text{jointOMP}(Y^d, \Phi, D^d; Y^u, \phi, D^u)$. D^d, D^u: learned dictionary.

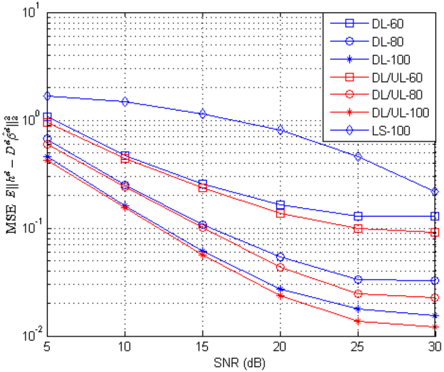

Figure 3: MSE comparison.

Conclusion

- In this work we propose a novel downlink channel estimation algorithm in FDD Massive MIMO systems.

Conclusion

- In this work we propose a novel downlink channel estimation algorithm in FDD Massive MIMO systems.
- Joint uplink/downlink dictionary learning can explore similar scattering effect between the uplink and downlink channel, leading to a joint sparse representation.

Conclusion

- In this work we propose a novel downlink channel estimation algorithm in FDD Massive MIMO systems.
- Joint uplink/downlink dictionary learning can explore similar scattering effect between the uplink and downlink channel, leading to a joint sparse representation.
- Joint compressed channel estimation can further improve the recovery performance by utilizing uplink training information.