Wideband Massive MIMO Channel Estimation via Sequential Atomic Norm Minimization

Stelios Stefanatos\textsuperscript{1} Mahdi Barzegar Khalilsarai\textsuperscript{2} Gerhard Wunder\textsuperscript{1}

\textsuperscript{1}Department of Mathematics and Computer Science
Freie Universität Berlin, Germany

\textsuperscript{2}Department of Telecommunication Systems
Technische Universität Berlin, Germany

GlobalSIP 2018

Freie Universität Berlin

TU Berlin
1. Introduction and System Model

2. Wideband Massive MIMO Channel Estimation via ANM
   - Algorithm Design
   - Performance Characterization

3. Numerical Results

4. Conclusion
Outline

1. Introduction and System Model

2. Wideband Massive MIMO Channel Estimation via ANM
   - Algorithm Design
   - Performance Characterization

3. Numerical Results

4. Conclusion
Introduction

Challenges of “Massive” Communications

- Extremely large number of UEs
- Short-length transmissions
- Extremely large signal space

Critical System Design Goal

Employ channel estimation procedures that

- provide reliable estimates
- are of low complexity
- require small training overhead

In this Work:

1. A low-complexity, ANM-based channel estimator for uplink wideband mMIMO is proposed
2. MSE performance characterized by tight lower bounds
3. Close to optimal for low-to-moderate number of propagation paths
Introduction

Challenges of “Massive” Communications

- Extremely large number of UEs
- Short-length transmissions
- Extremely large signal space

Critical System Design Goal

Employ channel estimation procedures that

- provide reliable estimates
- are of low complexity
- require small training overhead

In this Work:

1. A low-complexity, ANM-based channel estimator for uplink wideband mMIMO is proposed
2. MSE performance characterized by tight lower bounds
3. Close to optimal for low-to-moderate number of propagation paths
System Model (1/2)

- Single cell, uplink
- ULA of \( M \gg 1 \) antennas at BS
- Single-antenna UE
- OFDM signaling with \( N \gg 1 \) subcarriers
- Link characterized by the unknown space-frequency transfer matrix \( H \in \mathbb{C}^{M \times N} \)
- UE transmits pilot symbols over a set \( \mathcal{N}_p \subseteq \{0, 1, \ldots, N - 1\} \) of \( N_p \) subcarriers
- BS utilizes the observations from a set \( \mathcal{M}_p \subseteq \{0, 1, \ldots, M - 1\} \) of \( M_p \) antennas

Assumption

Sets \( \mathcal{N}_p, \mathcal{M}_p \) are selected randomly and uniformly from \( \{0, 1, \ldots, N - 1\} \), \( \{0, 1, \ldots, M - 1\} \), respectively

- motivated by compressive sensing theory
- results in an robust and multiuser-fair design
- allows for tractable analysis

- \( N_p, M_p \) are design parameters to be specified
System Model (2/2)

- Observed $M_p \times N_p$ signal at the BS (all-ones pilot symbols):
  \[ Y = S M_p H S_{N_p}^T + Z \]

  - $S M_p \in \mathbb{R}^{M_p \times M}$, $S_{N_p} \in \mathbb{R}^{N_p \times M}$: downsampling matrices
  - $Z \in \mathbb{C}^{M_p \times N_p}$: AWGN of variance $\sigma^2$

Receiver Task

Obtain a low-complexity and accurate estimate of the $MN$ elements of $H$ given the $M_p N_p < MN$ observations in $Y$

- Underdetermined linear system

Key concept

Exploit structural properties of the physical channel
**System Model (2/2)**

- Observed $M_p \times N_p$ signal at the BS (all-ones pilot symbols):
  \[
  Y = S_{M_p} H S_{N_p}^T + Z
  \]
  
  $S_{M_p} \in \mathbb{R}^{M_p \times M}$, $S_{N_p} \in \mathbb{R}^{N_p \times M}$: downsampling matrices
  
  $Z \in \mathbb{C}^{M_p \times N_p}$: AWGN of variance $\sigma^2$

---

**Receiver Task**

Obtain a low-complexity and accurate estimate of the $MN$ elements of $H$ given the $M_p N_p < MN$ observations in $Y$

- Underdetermined linear system

---

**Key concept**

Exploit structural properties of the physical channel
\[ H[n; m] = \sum_{l=0}^{L-1} c_l e^{-i2\pi m \theta_l} e^{-i2\pi n \tau_l}, \quad n \in [N], \quad m \in [M] \]

- \( L \): number of paths
- \( c_l \in \mathbb{C} \): gain of \( l \)th path
- \( \theta_l \in [0, 1] \): angle of arrival (AoA) of \( l \)th path (normalized)
- \( \tau_l \in [0, 1] \): delay of \( l \)th path (normalized)

- **Channel described by** 3L ≪ MN **path parameters** \( \{(\rho_l, \theta_l, \tau_l)\}_{l=0}^{L-1} \) **in the angle-delay domain**

- **Maximum Likelihood (ML) detection of path parameters:**
  \[
  \{(\hat{c}_l, \hat{\tau}_l, \hat{\theta}_l)\}_{l=0}^{L-1} = \arg \min_{\{(c_l, \tau_l, \theta_l)\}_{l=0}^{L-1}} \left\| Y - \mathbf{S}_M \mathbf{H} \mathbf{S}_N^T \right\|^2
  \]

NP-hard problem \( \Longrightarrow \) suboptimal solutions necessary
Parametric Wideband Massive MIMO Channel Model

\[ H[n; m] = \sum_{l=0}^{L-1} c_l e^{-i2\pi m n_1} e^{-i2\pi n_1 l}, n \in [N], m \in [M] \]

- \( L \): number of paths
- \( c_l \in \mathbb{C} \): gain of \( l \)th path
- \( \theta_l \in [0, 1] \): angle of arrival (AoA) of \( l \)th path (normalized)
- \( \tau_l \in [0, 1] \): delay of \( l \)th path (normalized)

- Channel described by \( 3L \ll MN \) path parameters \( \{(\rho_l, \theta_l, \tau_l)\}_{l=0}^{L-1} \) in the angle-delay domain

- Maximum Likelihood (ML) detection of path parameters:

\[
\{(\hat{c}_l, \hat{\tau}_l, \hat{\theta}_l)\}_{l=0}^{L-1} = \arg \min_{\{(c_l, \tau_l, \theta_l)\}_{l=0}^{L-1}} \| Y - S_{M_p} HS_{N_p} \|_2^2
\]

NP-hard problem \( \implies \) suboptimal solutions necessary
Outline

1 Introduction and System Model

2 Wideband Massive MIMO Channel Estimation via ANM
   - Algorithm Design
   - Performance Characterization

3 Numerical Results

4 Conclusion
Channel Estimation via Atomic Norm Minimization

Define the *atom set* (manifold)

\[ \mathcal{A} \triangleq \{ f_M(\theta)f_N^H(\tau) : (\theta, \tau) \in [0, 1] \times [0, 1] \} \]

- \[ f_M(\theta) \triangleq [1, e^{-i2\pi\theta}, \ldots, e^{-i2\pi\theta(M-1)}]^T \] and similarly for \( f_N(\tau) \)

Rationale for this set:

\[ H = \sum_{l=0}^{L-1} c_l f_M(\theta_l)f_N^H(\tau_l), \text{ i.e., } H \in \text{span}(\mathcal{A}) \]

### Definition (Atomic Norm)

The atomic norm of an arbitrary matrix \( X \in \mathbb{C}^{M \times N} \) w.r.t. \( \mathcal{A} \) is

\[
\|X\|_A \triangleq \inf_{c_l \in \mathbb{C}, \theta_l, \tau_l \in [0, 1]} \left\{ \sum_l |c_l| : \left\| X = \sum_l c_l f_M(\theta_l)f_N^H(\tau_l) \right\| \right\}
\]

- Extension of the standard \( \ell_1 \)-norm

Channel Estimation via Atomic Norm Minimization

\[ \hat{H} = \arg\min_{X \in \mathbb{C}^{M \times N}} \left\{ \left\| X \right\|_A : \left\| Y - S_MpXS_{Np}^T \right\| \leq \|\hat{Z}\| \right\} \]
Define the *atom set* (manifold)

\[ \mathcal{A} \triangleq \left\{ f_M(\theta)f_H^N(\tau) : (\theta, \tau) \in [0, 1] \times [0, 1] \right\} \]

- \( f_M(\theta) \triangleq [1, e^{-i2\pi\theta}, \ldots, e^{-i2\pi\theta(M-1)}]^T \) and similarly for \( f_N(\tau) \)
- Rationale for this set: \( H = \sum_{l=0}^{L-1} c_l f_M(\theta_l)f_H^N(\tau_l) \), i.e., \( H \in \text{span}(\mathcal{A}) \)

**Definition (Atomic Norm)**

The atomic norm of an arbitrary matrix \( X \in \mathbb{C}^{M \times N} \) w.r.t. \( \mathcal{A} \) is

\[ \|X\|_\mathcal{A} \triangleq \inf \left\{ \sum_l |c_l| : c_l \in \mathbb{C}, \theta_l, \tau_l \in [0, 1] \right\} \left\{ \sum_l |c_l| \right\} \left\{ \sum_l c_l f_M(\theta_l)f_H^N(\tau_l) \right\} \]

- Extension of the standard \( \ell_1 \)-norm

**Channel Estimation via Atomic Norm Minimization**

\[ \hat{H} = \arg\min_{X \in \mathbb{C}^{M \times N}} \left\{ \|X\|_\mathcal{A} : \|Y - S_{M_p}X S_{N_p}^T \| \leq \|\hat{Z}\| \right\} \]
**Theorem (Informal Statement)**

Under (a) noiseless conditions and (b) sufficiently large $N_p$, $M_p$, perfect recovery of $\mathbf{H}$ can be achieved with high probability as long as channel paths are sufficiently separated in the delay-angle domain, i.e.,

$$\min_{l \neq l'} \max \{|\theta_l - \theta_{l'}|, |\tau_l - \tau_{l'}|\} > d \approx 1 / \min \{M, N\}$$

(a) separable paths

(b) non-separable paths
Computation of $\| \cdot \|_{\mathcal{A}}$ can be formulated as an SDP problem, resulting in a convex program for obtaining $\hat{H}$:

$$
\begin{cases}
\text{minimize} & \frac{1}{2} \left( \text{tr} \{ T_{2D}(u) \} + t \right) \\
\hat{H} \in \mathbb{C}^{M \times N}, u \in \mathbb{C}^{MN}, t > 0 \\
\text{subject to} & \begin{pmatrix} T_{2D}(u) & \text{vec}(\hat{H}) \\ \text{vec}(\hat{H})^H & t \end{pmatrix} \succeq 0, \\
& \| Y - S_{M_p} \hat{H} S_{N_p}^T \| \leq \| \hat{Z} \|
\end{cases}
$$

- $T_{2D}(u) \in \mathbb{C}^{MN \times MN}$: block Toeplitz matrix
- Angle-delay pairs of paths can be estimated from the Vandermonde Decomposition of $T_{2D}(u)$
  - denoising gains when $L$ is known

Complexity of solution: $O(MN)$

Impractical when $M \gg 1$ and/or $N \gg 1 \implies$ Low-complexity alternatives needed
Proposed Approach (1/2)

Basic Idea

Decouple the spatial and frequency dimensions, treating them sequentially as Multiple Measurement Vectors (MMV) estimation problems and apply ANM-based estimation to each.

(a) observation

(b) interpolate over space

(c) interpolate over frequency
Proposed Approach (2/2)

1. Spatial dimension interpolation:
   - Rewrite the observation matrix as \( Y = S_M p H S_{N_p}^T + Z \)
   - \( \triangleq H_1 \)
   - Note that \( H_1 \) can be written as \( H_1 = \sum_{l=0}^{L-1} c_l f_M(\theta_l) b_{1,l}^H, b_{1,l} \triangleq S_{N_p} f_N(\tau_l) \)
   - By ignoring the structure of \( \{b_{1,l}\} \) and noise, an estimate of \( H_1 \) can be obtained as
     \[
     \hat{H}_1 = \arg\min_{X \in \mathbb{C}^{M \times N_p}} \left\{ \|X\|_{A_{MMV_1}} \mid Y = S_M p X \right\},
     \]
     where \( A_{MMV_1} \triangleq \{ f_M(\theta)b_{1}^H, \theta \in [0, 1], b_1 \in \mathbb{C}^M, \|b_1\|^2 = 1 \} \)
   - Denoise the estimate exploiting that there are \( L \) paths

2. Frequency dimension interpolation:
   - Repeat the same approach treating now \( \hat{H}_1 \in \mathbb{C}^{M \times N_p} \) as the partial observations of the complete channel matrix with structure \( H = \sum_{l=0}^{L-1} c_l b_{2,l} f_N^H(\tau_l) \)

SDP implementation with complexity order \( O(M + N) \ll O(MN) \)
Universal MSE Bound

- Exact characterization of ANM-based estimation performance extremely difficult
  - Resort to bounds

**Theorem (Universal Bound)**

The per-element MSE of any unbiased estimator of $\mathbf{H}$ is lower bounded as

$$\frac{1}{MN} \mathbb{E}(\|\hat{\mathbf{H}} - \mathbf{H}\|^2) \geq \frac{2L\sigma^2}{M_p N_p},$$

where the expectation is over the statistics of noise, $N_p$, $M_p$.

- Bound is looser than the CRLB, i.e., non-achievable, in general
- **Trade off** $N_p$ for $M_p \implies N_p \geq L$ is not required in massive MIMO
- Scales as $O(L)$
- Bound holds with no assumptions on path separability
MSE Bound for Proposed Algorithm

The following result can serve as an approximation of the MSE performance.

**Theorem**

Under the assumption that the error $\hat{H}_1 - H_1$ consists of i.i.d., zero mean, Gaussian elements, the per-element MSE of any unbiased estimator of $H$ from $\hat{H}_1$ that treats the rows of $H$ as MMV, is lower bounded as

$$\frac{1}{MN} \mathbb{E}(\|\hat{H} - H\|^2) \geq \frac{L^2 \sigma^2 (1 + 2N_p)(1 + 2M)}{4MM_pN_p^2} \approx \frac{L^2 \sigma^2}{M_pN_p} \text{ (for } N_p, M \gg 1).$$

- obtained under assumptions for the spatial-interpolation estimate that do not hold
- $L/2$ times greater than the universal bound
- Scales as $O(L^2)$ instead of $O(L)$
Outline

1 Introduction and System Model

2 Wideband Massive MIMO Channel Estimation via ANM
   - Algorithm Design
   - Performance Characterization

3 Numerical Results

4 Conclusion
System Setup

- $M = 100$ ULA elements with full antenna observations ($M_p = M$)
- $N = 100$ OFDM subcarriers
  - 2D ANM-based estimation practically infeasible

- i.i.d. paths with $\theta_l \sim U[0, 1], \tau_l \sim U[0, 1/4], c_l \sim \mathcal{CN}(0, 1/L)$
  - No restrictions on the path separability

- average SNR $= 1/\sigma^2 = 10$ dB

Compare MSE of proposed algorithm with:

1. naive LS with $N_p = N$ (MSE $= \sigma^2 = 10^{-1}$)
2. LMMSE interpolator
3. conventional (oversampled) $\ell_1$-norm minimization (BPDN)
4. $\mathcal{O}(N + M)$-complexity ANM-based approach (with path separability)$^1,^2$
5. universal bound

---


$^2$ J.-F. Cai, W. Xu, and Y. Yang, “Large scale 2D spectral compressed sensing in continuous domain,” ICASSP 2017
MSE dependence on number of pilot subcarriers

- Results averaging over $\mathbf{H}, N_p, Z$
- $L = 3$ (very sparse channel)
- Proposed performs very close to optimal and outperforms other approaches
- Massive MIMO offers potential for (huge) denoising and/or training overhead gains
MSE dependence on number of paths

- $N_p = N$ (full observations)
- analysis closely follows MSE of proposed algorithm
- MSE scaling as $O(L^2)$ eventually results in worse performance than BPDN
- For low-to-moderate $L$, proposed approach provides significant better performance
Outline

1 Introduction and System Model

2 Wideband Massive MIMO Channel Estimation via ANM
   - Algorithm Design
   - Performance Characterization

3 Numerical Results

4 Conclusion
Conclusion and Future Work

- An ANM-based algorithm for wideband massive MIMO channel estimator was proposed.

- Performs close to optimal for low-to-moderate number of paths w/o any assumptions on path separability.

- Possible extensions:
  - time-varying channels, multi-antenna UEs
  - multi-user, multi-cell setting