Massive MIMO Systems: Signal Processing Challenges and Research Trends

Rodrigo C. de Lamare
CETUC, PUC-Rio, Brazil
Communications Research Group,
Department of Electronics, University of York, U.K.
delamare@cetuc.puc-rio.br
Outline

• Introduction

• Application Scenarios and Signal Models
  – Satellite, cellular and local area networks
  – Downlink and uplink Models

• Transmit Processing Strategies
  – TDD operation
  – Pilot contamination
  – Resource allocation
  – Precoding

• Receive Processing Techniques
  – Parameter estimation
  – Detection algorithms
  – Error control coding
  – Mitigation of RF impairment

• Future Trends and Emerging Topics

• Concluding Remarks
Introduction (1/2)

• Wireless networks are experiencing a huge increase in the delivered amount of data due to emerging applications: M2M, video, etc.

• Key problems: High data rates require extra spectrum and energy which are very scarce, scalability of devices and signal processing algorithms.

• Future networks will require techniques that can substantially increase the capacity (bits/Hz) whilst not requiring extra spectrum or extra energy.

• Massive MIMO is a potential solution to these problems:
  – Very large arrays with an order of magnitude higher number of sensors.
  – Deployment of devices (access points, mobile phones and tables) with a large number of antenna elements.
  – Huge multiplexing gains allowing an order of magnitude higher data rates.

Massive MIMO networks will be structured around the following elements:

- **Antennas:**
  - Reduction of RF chains and costs,
  - Compact antennas and mitigation of coupling effects.

- **Electronic components:**
  - Low-cost components such as power amplifiers and RF components.
  - Flexibility for different air interfaces and replacement of coaxial cables.

- **Network architectures:**
  - Heterogeneous networks, small cells and network MIMO,
  - Cloud radio access networks to help different devices.

- **Protocols:**
  - Scheduling and medium-access protocols for numerous heterogeneous users.

- **Signal processing:**
  - Transmit and receive processing,
  - Scalability and hardware implementation,
  - Integration between signal processing and RF devices to deal with impairments.
Application Scenarios: Satellite Networks

• Multi-beam satellite systems:
  – Clear and well-defined scenario for massive MIMO.
  – Coverage region is served by multiple spot beams intended for the users.
  – Beams are shaped by the antenna feeds forming part of the payload.

• Research problem:
  – Interference caused by multiple adjacent spot beams that share the same frequency band.

• Interference mitigation:
  – Precoding for the forward link that needs CSI.
  – Detection algorithms for the reverse link.

Application Scenarios: Mobile Cellular Networks

• 5G mobile cellular networks:
  – Base stations: large arrays placed on rooftops or distributed antennas systems.
  – User terminals: phones, tablets with a significant number of antenna elements.
  – Compact antennas: mutual coupling
  – Coordination between cells.
  – Operation in TDD Mode.

• Research problems:
  – Uplink channel estimation: use of non-orthogonal pilots, the existence of adjacent cells and the coherence time of the channel require the system to reuse the pilots.
  – Pilot contamination: occurs when the CSI at the base station in one cell is affected by users from other cells.

• Topics for investigation:
  – Design of channel estimation strategies that avoid pilot contamination.
  – Design of precoding and detection algorithms for Network MIMO with large arrays.

Application Scenarios: Local Area Networks

• Future wireless local area networks:
  – Tremendous increase in the last years with the proliferation of access points (APs) in hot spots and home users.
  – IEEE 802.11ac: MIMO-OFDM with 20 or 40 MHz, up to 8 antennas at APs, 64 subcarriers which are not all used.
  – Massive MIMO: compact antennas, planar array geometries, etc.

• Research problems:
  – Coupling effects, spatial correlation and I/Q imbalances.
  – Physical dimensions of APs and user devices.

• Topics for investigation:
  – Coupling must be mitigated by DSP or smart RF solutions.
  – Design of efficient precoding and detection algorithms.
  – Decoding algorithms with reduced delay.

Signal Models: Downlink Case

• Multiuser massive MIMO system:
  – \( N_A \) antennas at the transmitter (satellite gateway, base station or WLAN AP).
  – \( K \) users and each is equipped with \( N_U \) antennas.

• Fundamental massive MIMO scenarios:
  – When \( N_A >> K N_U \) --> excess degrees of freedom leverages array gain
  – When \( N_A ~ K N_U \) --> absence of extra degrees of freedom

• Precoded data:
  – \( K N_U \times 1 \) data vector \( s[i] = [s_1^T[i] \ldots s_k^T[i] \ldots s_K^T[i]]^T \)
  – \( N_A \times 1 \) precoded data \( x_k[i] = P(s_k[i]) \),
  where each symbol \( s_i \in A = \{a_1, \ldots, a_N\} \), has zero mean and variance \( \sigma_s^2 \).

• Received data:
  – \( N_U \times 1 \) received vector
    \[
    r_k[i] = \sum_{k=1}^{K} H_k x_k[i] + n_k[i]
    \]
  where \( H_k \) is the \( N_U \times N_A \) channel matrix and \( n_k \) is the \( N_U \times 1 \) noise vector.
Signal Models: Uplink Case

- Multiuser massive MIMO system:
  - \( N_A \) antennas at the receiver (satellite gateway, base station with centralized or distributed antennas or WLAN AP).
  - \( K \) users and each is equipped with \( N_U \) antennas.

- Fundamental massive MIMO scenarios:
  - When \( N_A \gg KN_U \) -> excess degrees of freedom leverages array gain
  - When \( N_A \sim KN_U \) -> absence of extra degrees of freedom

- Received data with sufficient statistics:
  - \( N_A \times 1 \) received vector

\[
\mathbf{r}[i] = \sum_{k=1}^{K} \mathbf{H}_k \mathbf{s}_k[i] + \mathbf{n}[i]
\]

where \( \mathbf{H}_k \) is the \( N_A \times N_U \) channel matrix and \( \mathbf{n} \) is the \( N_A \times 1 \) noise vector.
Transmit Processing

• Optimal transmit strategy:
  – Requires CSI, obtained either by feedback channels or reciprocity.
  – Dirty paper coding.
  – Implicit scheduling and power allocation.
  – Very costly and impractical.

• Practical strategies:
  – TDD mode.
  – Pilot contamination.
  – Resource allocation.
  – Precoding techniques.


Operation in TDD Mode

- Does not require feedback channels to acquire CSI
- Rely on reciprocity to obtain CSI at the either the Tx or the Rx.
  - In Massive MIMO CSI is obtained at the base station or AP.
  - Problem with the amplifiers and filters that are different.
- Independent from the number of antennas $N_A$ at the base station or AP.
  - In FDD the CSI feedback requirements are proportional to the number of antennas.
  - In Massive MIMO it is more likely the use of TDD to eliminate the need for CSI feedback.
- Research problems:
  - Calibration and measurement techniques.

Pilot Contamination

- Adoption of TDD, network MIMO and uplink training -> phenomenon called Pilot Contamination.
- In multi-cell scenarios, it is difficult to employ orthogonal pilot sequences because their duration depends on the number of cells.
- The duration of the sequences is limited by the channel coherence time.
- Therefore, non-orthogonal pilot sequences are likely to be employed and this affects the CSI accuracy employed at the transmitter.
- Specifically, CSI is contaminated by a linear combination of channels of other users that share the same pilot.
- Consequently, the precoders and resource allocation will be highly affected by the contaminated CSI.
- Research problems:
  - Design of innovative training schemes
  - Design of precoders and resource allocation algorithms that can deal with pilot contamination.

Resource Allocation

• Key resources such as antennas, users and power must be allocated based on the instantaneous CSI of users and a metric.
• In massive MIMO systems, the spatial signatures of the users to be scheduled play a fundamental role -> they are quasi orthogonal.
• The multiuser diversity along with high array gains might be exploited by resource allocation algorithms along with timely CSI.
• Problem of user selection: scheduling is a combinatorial problem equivalent to the combination of $K$ choosing $Q$.
• When $K$ in the system is reasonably large: we need cost-effective user selection algorithms.
• Research problems:
  – Strategies based on greedy and discrete optimization methods.
  – Chunk strategies: to groups users/streams to reduce cost.

Precoding and Related Techniques (1/2)

- Main goals:
  - Mitigation of the multiuser interference.
  - Increase in the achievable sum-rates.

- Transmit matched filter:
  \[ x[i] = H^H s[i] \]
  the \( N_A \times K \) \( N_U \) matrix \( H \) contains the parameters of all the channels and the \( N_A \times 1 \) vector \( x[i] \) represents the data processed.

- Linear precoding:
  \[ x[i] = W_k s_k[i] + \sum_{l=1,l\neq k}^{K} W_l s_l[i], \]
  where the \( N_A \times NU \) precoding matrix \( W_k \) is a function of the channels.

- Block diagonalization precoding:
  - Improved BER and sum-rate performance over linear MMSE or ZF precoding.
  - Computational complexity of original BD is high for large systems.

Precoding and Related Techniques (2/2)

- Tomlinson-Harashima precoding:

\[ x[i] = F \tilde{x}[i] \]

where \( F \) is the \( N_A \times K N_U \) feedforward matrix obtained by an LQ decomposition of the channel matrix \( H \) and the input is computed element-by-element by

\[ \tilde{x}_l[i] = \text{mod}\{s_l[i] - \sum_{q=1}^{l-1} b_{lq} x_q[i]\}, \quad l = 1, \ldots, KN_U \]

where \( b_{lq} \) are the elements of the \( KN_U \times KN_U \) lower triangular matrix \( B \) that can also be obtained by an LQ decomposition.

- Vector perturbation precoding:

\[ p[i] = \arg \min_{p'[i] \in A C Z^K} \|W (s[i] + p'[i])\|^2 \]

where \( W \) is a precoder such that \( \text{Tr}(W^H W) \leq P \), the scalar \( A \) depends on the constellation size and \( CZ^K \) is the \( K \)-dimensional complex lattice.

Simulation Results

- Sum-rate performance against SNR of precoding algorithms
- Scenario: $N_A = 128$, $K = 8$ users and $N_U = 8$ antenna elements
- Channels are uncorrelated with coefficients drawn from complex Gaussian random variables with zero mean and unit variance.
Receive Processing

• Parameter estimation techniques:
  – Channel estimation.
  – Estimation of receive filter parameters.
• Detection strategies:
  – ML detection.
  – Suboptimal detection.
• Error control coding:
  – Channel codes.
  – Iterative detection and decoding.
• Mitigation of RF impairments
  – I/Q imbalances in the RF chains of large arrays.
  – Spatial correlation and condition number of channels
  – Mutual coupling between antenna elements in compact arrays
Parameter Estimation Techniques

• Channel estimation:
  – Semi-blind techniques to deal with pilot contamination.
  – Superimposed training techniques.

• Receive filter parameter estimation:
  – Reduced-rank techniques.
  – Sparsity-aware and compressive sensing algorithms.
  – Low-complexity adaptive techniques for estimating and tracking parameters in the presence of mobility.

Detection Techniques (1/2)

- Main goal: to separate the data streams of the users at the receiver
- Optimal ML detector:
  \[
  \hat{s}_{\text{ML}}[i] = \arg\min_{s[i]} ||r[i] - Hs[i]||^2
  \]
  where the $KN_U \times 1$ data vector $s[i]$ contains the symbols of all users.
- The complexity of the ML detector grows exponentially with the constellation size and $KN_U$.
- Sphere decoders are appealing in MIMO systems with small dimensions but are unlikely to be used in massive MIMO systems.

Detection Techniques (2/2)

- Linear detectors:
  \[ \hat{s}[i] = Q(W^H r[i]) \]
  
  where \( W \) is the \( N_A \times KN_U \) matrix receive filter applied to the received data.

- Decision feedback detectors:
  \[ \hat{s} = Q(W^H r[i] - F^H \hat{s}_o[i]) \]
  
  where the receive filters \( W \) and \( F \) can be computed using various design criteria and optimization algorithms.


Simulation Results

- BER performance against SNR of detection algorithms.
- Scenario: $N_A = 128$, $K = 8$ users and $N_U = 8$ antenna elements.
- Channels are uncorrelated with coefficients drawn from complex Gaussian random variables with zero mean and unit variance.
Error Control Coding and Iterative Processing

- Choice of channel coding
  - Convolutional codes.
  - Turbo codes.
  - LDPC codes.
- Iterative detection and decoding

- Research problems
  - Reducing the number of iterations.
  - Improving the exchange of soft information.


Mitigation of RF Impairments

- **Coupling effects:**
  - Reduction of the physical size of antennas induces coupling.
  - Coupling reduces the multiplexing gain and the degrades the performance.

- **I/Q imbalances:**
  - Phase offsets degrade the performance.
  - Originate non-circular data.

- **Failures of antenna elements:**
  - Cheaper components may lead to more frequent failures.
  - Reduction of the degrees of freedom.
  - Signal processing algorithms should be able to cope with them.

---


Future Trends and Emerging Topics

• Transmit processing
  – Cost-effective scheduling techniques: greedy techniques and discrete optimisation tools.
  – Calibration techniques.
  – Precoders with scalability in terms of complexity: use of transmit matched filters as the front-end.

• Receive processing
  – Cost-effective detection algorithms: use of the receive matched filter as the front-end.
  – Reduced-delay decoding algorithms and IDD schemes.
  – Mitigation of impairments: estimation of I/Q imbalances and use of widely-linear processing.
Concluding Remarks

• A tutorial on massive MIMO systems focusing on signal processing challenges and future trends in this exciting research topic has been given.
• Key application scenarios which include multibeam satellite, cellular and local area networks have been examined along with several operational requirements of massive MIMO networks.
• Transmit and receive processing tasks have been discussed and signal processing needs for future massive MIMO networks have been identified.
• Numerical results have illustrated some of the discussions on transmit and receive processing functions and future trends have been highlighted.
• Massive MIMO technology is likely to be incorporated into applications on a gradual basis by the increase in the number of antenna elements.
Further Reading and References
