Large-MIMO: A Technology Whose Time Has Come

A. Chockalingam and B. Sundar Rajan

(achockal,bsrajan@ece.iisc.ernet.in)

April 2010

Department of Electrical Communication Engineering

(http://www.ece.iisc.ernet.in/)

Indian Institute of Science

Bangalore – 560 012. INDIA
A. Chockalingam and B. Sundar Rajan: Large-MIMO: A Technology Whose Time Has Come
Dept. of ECE, IISc, Bangalore, April 2010

1

MIMO System

Transmit Side

Receive Side

• (e.g., Base Station, Access Point, Set top box)
• (e.g., Set top box, Laptop, HDTV)

\[ N_t : \# \text{Transmit Antennas} \]
\[ N_r : \# \text{Receive Antennas} \]
Why Multiple Antennas?

$N_t$: No. of transmit antennas, $N_r$: No. of receive antennas

<table>
<thead>
<tr>
<th># Antennas</th>
<th>Error Probability ($P_e$)</th>
<th>Capacity ($C$), bps/Hz</th>
</tr>
</thead>
<tbody>
<tr>
<td>$N_t = N_r = 1$ (SISO)</td>
<td>$P_e \propto SNR^{-1}$</td>
<td>$C = \log(SNR)$</td>
</tr>
<tr>
<td>$N_t = 1$, $N_r &gt; 1$ (SIMO)</td>
<td>$P_e \propto SNR^{-N_r}$</td>
<td>$C = \log(SNR)$</td>
</tr>
<tr>
<td>$N_t &gt; 1$, $N_r &gt; 1$ (MIMO)</td>
<td>$P_e \propto SNR^{-N_tN_r}$</td>
<td>$C = \min(N_t, N_r) \log(SNR)$</td>
</tr>
</tbody>
</table>

$N_tN_r$ : Diversity Gain  
$\min(N_t, N_r)$ : Spatial Mux Gain

• Large $N_t$, $N_r \rightarrow$ increased spectral efficiency


Large-MIMO Approach

• Employ large number (several tens) of antennas at the Tx and Rx

• Achieve high spectral efficiencies (tens to hundreds of bps/Hz)
  – Data rate (bps) = Spectral efficiency (bps/Hz) × Bandwidth (Hz)
  – e.g., 100 bps/Hz \(\Rightarrow\) 1 Gbps rate in just 10 MHz bandwidth

• Limitation in current MIMO standards
  – spectral efficiency: \(\sim\) 10 bps/Hz only
  – 2 to 4 transmit antennas
  – e.g., 750 Mbps in 80 MHz in 802.11n using 4 Tx antennas
  – do not exploit the potential of large spatial dimensions
Technological Challenges in Realizing Large-MIMO

- Placement of large number of antennas in communication terminals
  - Feasible in moderately sized communication terminals (e.g., Set top boxes, Laptops, BS towers)
  - Use high carrier frequencies for small carrier wavelengths (e.g., 5 GHz, 60 GHz)
- RF technologies
  - Multiple IF/RF transmit and receive chains
- Large-MIMO detection
  - Need low-complexity detectors
- Channel estimation
  - Estimation of large number of channel coefficients
16-Antenna Channel Sounding for IEEE 802.11ac (5 GHz)

Example Antenna Configurations Used

8 Slot Antenna Array
Two V-H and two ±45° pairs
λ/2 separation between slot pairs

16 Linear Dipole Antenna Array
λ/2 separation between elements

64 × 64 MIMO Indoor Channel Sounding (5 GHz)

(a) 64-Antenna/RF hardware at 5 GHz

(b) LOS setup

Some Recent Wireless Products

• Can see the trend in packaging increasing number of antennas/RF chains in wireless products

Source: Internet
Detection Complexity: A Key Challenge in Large-MIMO

- Optimal detection has **exponential complexity** in # transmit antennas

- Need **low-complexity** algorithms that are **near-optimal**

- A possible approach to low-complexity solutions
  - Seek algorithms from machine learning (ML)
  - Large-dimension problems are routinely addressed in other areas (e.g., computer vision, web search) using ML algorithms
  - Communications area too has benefited from ML algorithms
    - **MUD** in CDMA (**large # users**), Turbo/LDPC decoding (**large frame sizes**)
Linear Vector Channels

• Several communication systems can be characterized by the following linear vector channel model

\[ y_c = H_c x_c + n_c \]

\[ x_c \in \mathbb{C}^{d_t}, \quad H_c \in \mathbb{C}^{d_r \times d_t}, \quad y_c \in \mathbb{C}^{d_r}, \quad n_c \in \mathbb{C}^{d_r} \]

• Examples

  – **MIMO**
    * \( d_t = N_t \), # Tx antennas; \( d_r = N_r \), # Rx antennas; \( x_c \): Tx symbol vector
    * \( H_c \): Channel gain matrix; \( y_c \): Rx signal vector; \( n_c \): Noise vector

  – **Coding**
    * \( d_t = k \), # Information bits; \( d_r = n \), # Coded bits; \( x_c \): Information bit vector
    * \( H_c \): Generator matrix; \( y_c \): Rx signal vector; \( n_c \): Noise vector

  – **CDMA**
    * \( d_t = d_r = K \), # users; \( x_c \): Tx. bit vector; \( H_c \): Cross correlation matrix,
    * \( y_c \): Rx signal vector; \( n_c \): Noise vector
Optimum Detection

- **Problem**
  - Obtain an estimate of $x_c$, given $y_c$ and $H_c$

- **Maximum likelihood (ML) solution**

\[
\mathbf{x}_{ML} = \arg\min_{\mathbf{x}_c \in \mathbb{A}^{d_t}} \left\| \mathbf{y}_c - H_c \mathbf{x}_c \right\|^2  \\
\triangleq \phi(x_c)
\]

$\mathbb{A}$: signaling alphabet; $\phi(x_c)$: ML cost

- ML cost: $\phi(x_c) = x_c^H H_c^H H_c x_c - 2\Re\left( y_c^H H_c x_c \right)$
Optimum Detection

- Let $\mathbb{A}$ be $M$-PAM or $M$-QAM (Two PAMs in quadrature)
  - $M$-PAM symbols take values from $\{A_m, m = 1, \cdots, M\}$, $A_m = (2m - 1 - M)$
  - $y_c = y_I + jy_Q$, $x_c = x_I + jx_Q$, $n_c = n_I + jn_Q$, $H_c = H_I + jH_Q$
- Convert (1) into a real-valued system model
  \[ y = Hx + n \]  
  $H \in \mathbb{R}^{2d_t \times 2d_t}$, $y \in \mathbb{R}^{2d_t}$, $x \in \mathbb{R}^{2d_t}$, $n \in \mathbb{R}^{2d_t}$

- ML solution
  \[ x_{ML} = \arg \min_{x \in \mathbb{S}} \|y - Hx\|^2 x^T H^T H x - 2y^T H x, \]  
  $\mathbb{S}$: $2d_t$-dimensional signal space (Cartesian product of $\mathbb{A}_1$ to $\mathbb{A}_{2d_t}$; $\mathbb{A}_i$: $M$-PAM signal set from which $x_i$ takes values, $i = 1, \cdots, 2d_t$).
  **ML Complexity**: Exponential in $d_t$
Optimum Detection

- Maximum a posteriori (MAP) solution
  - Consider square $M$-QAM
  - Each entry of $\mathbf{x}$ belongs to a $\sqrt{M}$-PAM constellation
  - Let $b_i^{(0)}, b_i^{(1)}, \ldots, b_i^{(q-1)}$ denote the $q = \log_2(\sqrt{M})$ constituent bits of $x_i$
  - $x_i$ can be written as
    \[ x_i = \sum_{j=0}^{q-1} 2^j b_i^{(j)}, \quad i = 0, 1, \ldots, 2d_t - 1 \]
  - Let the bit vector $\mathbf{b} \in \{\pm 1\}^{2qd_t}$ be written as
    \[ \mathbf{b} \triangleq \begin{bmatrix} b_0^{(0)} & \cdots & b_0^{(q-1)} & b_1^{(0)} & \cdots & b_1^{(q-1)} & \cdots & b_{2d_t-1}^{(0)} & \cdots & b_{2d_t-1}^{(q-1)} \end{bmatrix}^T \]
  - Defining $\mathbf{c} \triangleq [2^0 2^1 \cdots 2^{q-1}]$, $\mathbf{x}$ can be written as
    \[ \mathbf{x} = (\mathbf{I}_{2d_t} \otimes \mathbf{c}) \mathbf{b} \]
Optimum Detection

- Rx signal model can be written as

\[
\mathbf{y} = \mathbf{H}(\mathbf{I}_{2d_t} \otimes \mathbf{c})\mathbf{b} + \mathbf{n}
\]

\[
\triangleq \mathbf{H}' \in \mathbb{R}^{2d_r \times 2qd_t}
\]

- MAP estimate of \( b_i^{(j)} \), \( i = 0, \cdots, 2d_t - 1, \quad j = 0, \cdots, q - 1 \) is

\[
\hat{b}_i^{(j)} = \arg \max_{a \in \{\pm 1\}} p(b_i^{(j)} = a \mid \mathbf{y}, \mathbf{H}')
\]

- Complexity: Exponential in \( qd_t \)
Sub-optimum Solutions

- Matched filter (MF)

\[ x_{MF} = H^T y \]

- Zero-forcing (ZF) solution

\[ x_{ZF} = H^{-1} y \]

- Minimum mean square error (MMSE) solution

\[ x_{MMSE} = (H + \sigma^2 I)^{-1} y \]

- These suboptimum solution vectors can be used as initial vectors in search algorithms to improve performance further
Near-Optimal Algorithms for Large $d_t$

- **Near-ML algorithms**
  - Local neighborhood search based
  - Likelihood ascent search (LAS)
  - Reactive tabu search (RTS)

- **Near-MAP algorithms**
  - Message passing based
  - Belief propagation (BP)
  - Probabilistic association (PDA)
LAS Algorithm

- Search for good solution vectors in the local neighborhood

- Neighborhood definition
  - Neighbors that differ in one coordinate
    * e.g., Consider $\mathbb{A} = \{\pm 1\}; \quad \mathbf{x} = [-1, 1, 1, -1]$
    * 1-bit away neighbors of $\mathbf{x}$:
      $$\mathcal{N}_1(\mathbf{x}) = \left\{ [-1, 1, 1, 1], [-1, 1, -1, -1], [-1, -1, 1, -1], [1, 1, 1, -1] \right\}$$
  - Neighbors that differ in two coordinates
  - 2-bit away neighbors of $\mathbf{x}$:
    $$\mathcal{N}_2(\mathbf{x}) = \left\{ [-1, 1, -1, 1], [-1, -1, -1, -1], [1, -1, 1, -1], \right.$$\left.$$ [1, 1, 1, 1], [1, 1, 1, -1], [1, -1, 1, 1] \right\}$$

- Choose best neighbor based on ML cost:
  $$\phi(\tilde{\mathbf{x}}) = \tilde{\mathbf{x}}^T \mathbf{H}^T \mathbf{H} \tilde{\mathbf{x}} - 2\mathbf{y}^T \mathbf{H} \tilde{\mathbf{x}}$$
LAS Algorithm

START

Compute initial solution vector

Find the neighborhood of the solution vector

Find the best vector in the neighborhood

Does this vector have a better cost function than that of the current solution vector?

Yes

Make this neighbor as the current solution vector

No

END
An Illustration of LAS Search Path
Local minima trap
Large-Dimension Behavior of LAS in V-BLAST [5],[6]

* 1-LAS: 1-symbol away neighborhood
* BER improves with increasing $N_t$ (large-dimension effect)


Complexity of 1-LAS in V-BLAST

• Consider $N_t = N_r$

• Total complexity comprises of 3 main parts
  1. Computing initial vector (e.g., ZF, MMSE): $O(N_t^2)$ per symbol
  2. Computing $H^T H$: $O(N_t^2)$ per symbol
  3. Search operation: $O(N_t)$ per symbol (through simulations)

• So, overall average per-symbol complexity: $O(N_t^2)$

• This low-complexity allows detection of V-BLAST signals in hundreds of spatial dimensions
Large # Dimensions: The Key

- Observation
  - In V-BLAST, LAS algorithm achieves near-ML performance, but only when the number of antennas is in hundreds
  - hundreds of antennas may not be practical

- Note
  - LAS requires large # dimensions to perform well
  - but, all dimensions need not be in space alone

- Q1: Can large # dimensions be created with less # Tx antennas?

- A1: Yes. Use time dimension as well. Approach: Non-orthogonal STBCs

- Q2: Can LAS modified to work well for smaller (tens) dimensions?

- A2: Yes. Approach: Escape strategies from local minima
An Escape Strategy from Local Minima [7]

- **Multistage LAS (M-LAS)
  - Start the algorithm as 1-LAS
  - On reaching the local minima,
    - find 2-symbol away neighbors of the local minima
    - choose the best 2-symbol away neighbor if it has lesser cost than local minima
    - run 1-LAS from this best neighbor till a local minima is reached
  - Expect better performance. Complexity is increased a little, but not by an order
  - Escape strategy with 3-symbol away neighborhood on reaching local minima

- Another promising strategy is **reactive tabu search**

Performance of M-LAS

* 3-LAS performs better than 1-LAS

- (e) 3-LAS versus 1-LAS
- (f) 3-LAS
Space-Time Block Codes

- Provide redundancy across space and time
- Goal of space-time coding
  - Achieve the maximum Tx-diversity of $N_t$ (i.e., full-diversity), high rate, decoding at low-complexity
- An STBC is usually represented by a $p \times n_t$ matrix
  - rows: time slots; $p$: # time slots
  - columns: Tx. antennas; $n_t$: # Tx. antennas

$$
X = \begin{bmatrix}
    s_{11} & s_{12} & \cdots & s_{1n_t} \\
    s_{21} & s_{22} & \cdots & s_{2n_t} \\
     \vdots & \vdots & \ddots & \vdots \\
    s_{p1} & s_{42} & \cdots & s_{pn_t}
\end{bmatrix}
$$

- $s_{ij}$ denotes the complex number transmitted in the $i$th time slot on the $j$th Tx antenna
Space-Time Block Codes

- Rate of an STBC, \( r = \frac{k}{p} \)
  - \( k \): number of information symbols sent in one STBC
  - \( p \): number of time slots in one STBC
  * Higher rate means more information carried by the code

- A matrix \( X \) is said to be a Orthogonal STBC if
  \[
  X^H X = \left( |x_1|^2 + |x_2|^2 + \cdots + |x_k|^2 \right) I_{nt}
  \]
  - Elements of \( X \) are linear combinations of \( x_1, \cdots, x_k \) and their conjugates
  - \( x_1, x_2, \cdots, x_k \) are information symbols

- 2-Tx Antennas Codes (\( 2 \times 2 \) Alamouti Code)
  \[
  X = \begin{bmatrix}
  x_1 & x_2 \\
  -x_2^* & x_1^*
  \end{bmatrix}, \quad k = 2, p = 2, r = 1, \text{ orthogonal STBC}
  \]
Linear-Complexity Decoding of OSTBCs

- Consider Alamouti code with $n_t = 2$, $n_r = 1$
- Received signal in $i$th slot, $y_i$, $i = 1, 2$, is
  
  \begin{align*}
  y_1 &= h_1 x_1 + h_2 x_2 + n_1 \\
  y_2 &= -h_1 x_2^* + h_2 x_1^* + n_2 
  \end{align*}

- ML decoding amounts to
  - computing
    \begin{align*}
    \tilde{x}_1 &= y_1 h_1^* + y_2^* h_2 \\
    \tilde{x}_2 &= y_1 h_2^* - y_2^* h_1
    \end{align*}
  - decoding $x_1$ by finding the symbol in the constellation that is closest to $\tilde{x}_1$
  - and decoding $x_2$ by finding the symbol that is closest to $\tilde{x}_2$

- This decoding feature is called **Single-Symbol Decodability (SSD)**
Orthogonal vs Non-Orthogonal STBCs

- **Orthogonal STBCs** are more widely known
  - e.g., $2 \times 2$ Alamouti code (Rate-1; 2 symbols in 2 channel uses)
  - advantages
    * linear complexity ML decoding, full transmit diversity
  - major drawback
    * rate falls linearly with increasing number of transmit antennas

- **Non-orthogonal STBCs**: less widely known
  - e.g., $2 \times 2$ Golden code (Rate-2; 4 symbols in 2 chl uses; same as V-BLAST)
    * advantages
      · High-rate (same as V-BLAST, i.e., $N_t$ symbols/channel use)
      · Full Transmit diversity
      · best of both worlds (in terms of data rate and transmit diversity)
    * What is the catch
      · decoding complexity
Non-Orthogonal STBCs

- Golden code \[8\] (2 × 2 non-orthogonal STBC)

\[\mathbf{X} = \begin{bmatrix}
    x_1 + \tau x_2 & x_3 + \tau x_4 \\
    i(x_3 + \mu x_4) & x_1 + \mu x_2
\end{bmatrix}, \quad k = 4, p = 2, r = 2\]

where \(\tau = \frac{1+\sqrt{5}}{2}\) and \(\mu = \frac{1-\sqrt{5}}{2}\)

- Features
  - Information Losslessness (ILL)
  - Full Diversity (FD)
  - Coding Gain (CG)

- ‘Perfect codes’ \[9\] achieve all the above three features
  - Golden code is a perfect code

---


High-Rate Non-Orthogonal STBCs from CDA for any $N_t$

- High-rate non-orthogonal STBCs from Cyclic Division Algebras (CDA) for arbitrary # transmit antennas, $n$, is given by the $n \times n$ matrix [10]

$$X = \begin{bmatrix}
\sum_{i=0}^{n-1} x_{0,i} t^i & \delta \sum_{i=0}^{n-1} x_{1,i} \omega_n^i t^i & \delta \sum_{i=0}^{n-1} x_{2,i} \omega_n^i t^i & \cdots & \delta \sum_{i=0}^{n-1} x_{n-1,i} \omega_n^{(n-1)i} t^i \\
\sum_{i=0}^{n-1} x_{1,i} t^i & \sum_{i=0}^{n-1} x_{0,i} \omega_n^i t^i & \delta \sum_{i=0}^{n-1} x_{2,i} \omega_n^i t^i & \cdots & \delta \sum_{i=0}^{n-1} x_{n-1,i} \omega_n^{(n-1)i} t^i \\
\sum_{i=0}^{n-1} x_{2,i} t^i & \sum_{i=0}^{n-1} x_{1,i} \omega_n^i t^i & \sum_{i=0}^{n-1} x_{0,i} \omega_n^i t^i & \cdots & \delta \sum_{i=0}^{n-1} x_{n-2,i} \omega_n^{(n-1)i} t^i \\
\vdots & \vdots & \vdots & \ddots & \vdots \\
\sum_{i=0}^{n-1} x_{n-2,i} t^i & \sum_{i=0}^{n-1} x_{n-3,i} \omega_n^i t^i & \sum_{i=0}^{n-1} x_{n-4,i} \omega_n^i t^i & \cdots & \delta \sum_{i=0}^{n-1} x_{n-1,i} \omega_n^{(n-1)i} t^i \\
\sum_{i=0}^{n-1} x_{n-1,i} t^i & \sum_{i=0}^{n-1} x_{n-2,i} \omega_n^i t^i & \sum_{i=0}^{n-1} x_{n-3,i} \omega_n^i t^i & \cdots & \sum_{i=0}^{n-1} x_{0,i} \omega_n^{(n-1)i} t^i
\end{bmatrix}$$

- $\omega_n = e^{\frac{j2\pi}{n}}$, $j = \sqrt{-1}$, and $x_{u,v}, 0 \leq u, v \leq n - 1$ are the data symbols from a QAM alphabet

- $n^2$ complex data symbols in one STBC matrix (i.e., $n$ complex data symbols per channel use)

- $\delta = t = 1$: Information-lossless (ILL); $\delta = e^{\sqrt{5}j}$ and $t = e^j$: Full diversity and ILL

- **Ques:** Can large (e.g., $32 \times 32$) STBCs from CDA decoded? **Ans:** LAS algorithm can.

---

Linear Vector Channel Model for NO-STBC

- \((n, p, k)\) STBC is a matrix \(X_c \in \mathbb{C}^{n \times p}\), \(n\): number of time slots, \(p\): number of tx antennas, \(k\): number of data symbols in one STBC; \((n = p \text{ and } k = n^2)\) for NO-STBC from CDA

- Received space-time signal matrix

\[
Y_c = H_c X_c + N_c,
\]

- Consider linear dispersion STBCs where \(X_c\) can be written in the form

\[
X_c = \sum_{i=1}^{k} x_c^{(i)} A_c^{(i)}
\]

where \(A_c^{(i)} \in \mathbb{C}^{N_t \times p}\) is the weight matrix corresponding to data symbol \(x_c^{(i)}\)

- Applying \(\text{vec}(.)\) operation

\[
\text{vec}(Y_c) = \sum_{i=1}^{k} x_c^{(i)} \text{vec}(H_c A_c^{(i)}) + \text{vec}(N_c)
\]

\[
= \sum_{i=1}^{k} x_c^{(i)} (I_{p \times p} \otimes H_c) \text{vec}(A_c^{(i)}) + \text{vec}(N_c)
\]
Linear Vector Channel Model for NO-STBC

- Define $y_c \triangleq \text{vec}(Y_c) \in \mathbb{C}^{N_r \times p}$, $\hat{H}_c \triangleq (I \otimes H_c) \in \mathbb{C}^{N_r \times N_t \times p}$, $a_c^{(i)} \triangleq \text{vec}(A_c^{(i)}) \in \mathbb{C}^{N_t \times p}$, $n_c \triangleq \text{vec}(N_c) \in \mathbb{C}^{N_r \times p}$

- System model can then be written in vector form as

$$y_c = \sum_{i=1}^{k} x_c^{(i)} (\hat{H}_c a_c^{(i)}) + n_c$$

$$= \tilde{H}_c x_c + n_c$$

(4)

$\tilde{H}_c \in \mathbb{C}^{N_r \times p \times k}$, whose $i$th column is $\hat{H}_c a_c^{(i)}$, $i = 1, \cdots, k$

$x_c \in \mathbb{C}^k$, whose $i$th entry is the data symbol $x_c^{(i)}$

- Convert the complex system model in (4) into real system model as before

- Apply LAS algorithm on the resulting real system model
LAS Performance in Decoding NO-STBCs [11]

(g) Uncoded $8 \times 8$, $16 \times 16$, $32 \times 32$ NO-STBC, 4-QAM

(h) Turbo Coded $32 \times 32$ NO-STBC, 16-QAM

### Comparison with Other Architectures/Detectors [11]

<table>
<thead>
<tr>
<th>No.</th>
<th>MIMO Architecture/Detector Combinations</th>
<th>Complexity (in # real operations to achieve $5 \times 10^{-2}$ uncoded BER)</th>
<th>SNR required for all combinations</th>
</tr>
</thead>
<tbody>
<tr>
<td>i)</td>
<td>$16 \times 16$ ILL-only CDA STBC (rate-16), 4-QAM and 1-LAS detection (Proposed scheme [11])</td>
<td>$3.473 \times 10^3$</td>
<td>$6.8,\text{dB}$</td>
</tr>
<tr>
<td>ii)</td>
<td>$16 \times 16$ ILL-only CDA STBC (rate-16), 4-QAM and ISIC algorithm [Choi, Cioffi]</td>
<td>$1.187 \times 10^5$</td>
<td>$11.3,\text{dB}$</td>
</tr>
<tr>
<td>iii)</td>
<td>Four $4 \times 4$ stacked rate-1 QOSTBCs, 256-QAM and IC algorithm [Jafarkhani]</td>
<td>$5.54 \times 10^6$</td>
<td>$24,\text{dB}$</td>
</tr>
<tr>
<td>iv)</td>
<td>Eight $2 \times 2$ stacked rate-1 Alamouti codes, 16-QAM and IC algorithm [Jafarkhani]</td>
<td>$8.719 \times 10^3$</td>
<td>$17,\text{dB}$</td>
</tr>
<tr>
<td>v)</td>
<td>$16 \times 16$ V-BLAST (rate-16) scheme, 4-QAM and sphere decoding</td>
<td>$4.66 \times 10^4$</td>
<td>$7,\text{dB}$</td>
</tr>
<tr>
<td>vi)</td>
<td>$16 \times 16$ V-BLAST (rate-16) scheme, 4-QAM and V-BLAST detector (ZF-SIC)</td>
<td>$1.75 \times 10^4$</td>
<td>$13,\text{dB}$</td>
</tr>
</tbody>
</table>
Effect of Spatial Correlation?

- Spatially correlated MIMO fading channel model by Gesbert et al [12]

\[
\begin{align*}
\text{correlated channel matrix}: \quad H &= \frac{1}{\sqrt{S}} R_{\theta_r,d_r}^{1/2} G_r R_{\theta_S,2D_r/S}^{1/2} G_t R_{\theta_t,d_t}^{1/2}
\end{align*}
\]

Spatial Correlation Degrades Performance \cite{11}

![Graph showing BER performance of 1-LAS detector](image)

**Figure 2:** Uncoded/coded BER performance of 1-LAS detector \(i\) in i.i.d. fading, and \(ii\) in correlated MIMO fading in \cite{3} with \(f_c = 5\) GHz, \(R = 500\) m, \(S = 30, D_t = D_r = 20\) m, \(\theta_t = \theta_r = 90^\circ\), and \(d_t = d_r = 2\lambda/3 = 4\) cm. 16 \(\times\) 16 STBC, \(N_t = N_r = 16\), 16-QAM, rate-3/4 turbo code, 48 bps/Hz.

Spatial correlation degrades performance.

![Graph showing Bit Error Rate vs Average Received SNR for different configurations.](image)

**Figure**: Effect of $N_r > N_t$ in correlated MIMO fading in [3] keeping $N_r d_r$ constant and $d_t = d_r$. $N_r d_r = 72$ cm, $f_c = 5$ GHz, $R = 500$ m, $S = 30$, $D_t = D_r = 20$ m, $\theta_t = \theta_r = 90^\circ$, $12 \times 12$ ILL-only STBC, $N_t = 12$, $N_r = 12, 18$, 16-QAM, rate-3/4 turbo code, 36 bps/Hz. Increasing # receive dimensions alleviates the loss due to spatial correlation.
**Channel Estimation in Large-MIMO?**

- **Training based channel estimation** \([11]\)
  - Send 1 Pilot matrix followed by \(N_d\) data STBC matrices

\[
\begin{array}{c}
\text{1 Pilot Matrix} \\
\text{Space} \\
\text{Data STBCs} \\
\text{N_t} \\
\text{Nd \times N_t} \\
\text{1 Frame}
\end{array}
\]

\[
\begin{array}{c}
\text{Pilot Matrix} \\
\text{Data STBCs}
\end{array}
\]

\[
\text{time}
\]

* 1 frame length (in # of channel uses), \(T = (N_d + 1)N_t\) [coherence time]
* 1 pilot matrix length (in # of channel uses), \(\tau = N_t\)

- Obtain an MMSE estimate of the channel matrix during pilot phase
- Use estimated channel matrix to detect data matrices using LAS detection
- Iterate between detection and channel estimation
Hassibi-Hochwald (H-H) bound [13] on capacity with estimated CSIR:

\[
C \geq \frac{T - \tau}{T} \mathbb{E} \left[ \log \det \left( I_{N_t} + \frac{\gamma^2 \beta_d \beta_p \tau}{N_t (1 + \gamma \beta_d) + \gamma \beta_p \tau} \hat{H}_c \hat{H}_c^H N_t \sigma^2 H_c \right) \right]
\]

Figure: H-H capacity bound [13] for 1P+8D (\(T = 144, \tau = 16, \beta_p = \beta_d = 1\)) and 1P+1D (\(T = 32, \tau = 16, \beta_p = \beta_d = 1\)) training for a 16 × 16 MIMO channel.

How Much Training is Required?

Fig. 5. Capacity as a function of number of transmit antennas $M$ with $\rho = 18$ dB and $N_r = 12$ receive antennas. The solid line is optimized over $T_r$ for $\rho_T = \rho_d = \rho$ (see (40)), and the dashed line is optimized over the power allocation with $T_r$ (Theorem 3). The dash-dotted line is the capacity when the receiver knows the channel perfectly. The maximum throughput is attained at $M \approx 15$.

**Figur6:** Capacity as a function of $N_t$ with $\text{SNR} = 18$ dB and $N_r = 12$. For a given $N_r$, SNR ($\gamma$), and coherence time ($T$), there is an optimum $N_t$ [13].
Figure 6: Turbo coded BER performance of LAS detection and channel estimation as a function of coherence time, $T = 32, 144, 400, 784$ ($N_d = 1, 8, 24, 48$), for a given $N_t = N_r = 16$. 16 × 16 ILL-only STBC, 4-QAM, rate-3/4 turbo code. Spectral efficiency and BER performance with estimated CSIR approaches to those with perfect CSIR in slow fading (i.e., large $T$).
Other Promising Large-MIMO Detection Algorithms

- Reactive Tabu Search [14]
- Probabilistic Data Association [15]
- Belief Propagation [16],[17]

These algorithms exhibit large-dimension behavior; i.e., their bit error performance improves with increasing $N_t$.


Reactive Tabu Search

- Another iterative local search algorithm
  - A metaheuristics algorithm
  - cannot guarantee optimal solution, but generally gives near optimal solution
- Uses ‘tabu’ mechanism to escape from local minima or cycles
  - Certain vectors are prohibited (made tabu) from becoming solution vectors for certain number of iterations (called tabu period) depending on the search path
  - This is meant to ensure efficient exploration of the search space
- The reactive part adapts the tabu period
RTS Algorithm [14]

START

Compute initial solution vector

Find the neighborhood of the solution vector

Find the best vector in the neighborhood

Does this vector have the best cost function found so far?

Yes

Is the move to this vector tabu?

Yes

Exclude the vector from the neighborhood

No

No

Is any move non-tabu?

Yes

Make this neighbor as the current solution vector

No

Make the oldest move performed as non-tabu

Update tabu matrix to reflect current and past P moves

Check for repetition of the solution vector

Update tabu period P based on repetition

END

stopping criterion satisfied?

Yes
An Illustration of RTS Search Path
Global Minima
Performance of RTS in V-BLAST

(a) Convergence of RTS

(b) RTS versus LAS

**Performance/Complexity of RTS in V-BLAST**

* RTS performs better than LAS at the same order of LAS complexity

![Graph showing performance and complexity](image)

(c) Performance in $32 \times 32$ V-BLAST, $M$-QAM

(d) Complexity
Performance of RTS in NO-STBC [14]

* RTS performs better than LAS

(e) 4-QAM  
(f) 16-QAM
Probabilistic Data Association

- Originally developed for target tracking
- Used in digital communications recently
- PDA
  - A reduced complexity alternative to a posteriori probability (APP) detector/decoder/equalizer.
  - Has been applied in
    * Turbo equalization (Yin et al 2004)
PDA Based Large-MIMO Detection [15]

- Iterative algorithm
  - In each iteration, \(2qk\) statistic updates (one for each bit) are performed

- Likelihood ratio of bit \(b_i^{(j)}\) in an iteration is

\[
\Lambda_i^{(j)} \triangleq \frac{P(y | b_i^{(j)} = +1)}{P(y | b_i^{(j)} = -1)} \quad \frac{P(b_i^{(j)} = +1)}{P(b_i^{(j)} = -1)}
\]
\[
\triangleq \beta_i^{(j)} \quad \triangleq \alpha_i^{(j)}
\]

- Received signal vector \(y\) can be written as

\[
y = h_{qi+j} b_i^{(j)} + \sum_{l=0}^{2k-1} \sum_{m=0}^{q-1} h_{ql+m} b_i^{(m)} + \mathbf{n}
\]
\[
\triangleq \tilde{\mathbf{n}} \text{ (interference+noise vector)}
\]

\(h_t\): \(t\)th column of \(H\)
PDA Based Large-MIMO Detection [15]

• Define $p_i^{j+} \triangleq P(b_i^{(j)} = +1)$ and $p_i^{j-} \triangleq P(b_i^{(j)} = -1)$

• To compute $\beta_i^{(j)}$, approximate the distribution of $\tilde{n}$ to be Gaussian

• Mean of $y$

$$
\mu_i^{j+} \triangleq \mathbb{E}(y|b_i^{(j)} = +1) = h_{qi+j} + \sum_{l=0}^{2k-1} \sum_{m=0}^{q-1} h_{ql+m}(2p_l^{m+} - 1)
$$

$$
\mu_i^{j-} \triangleq \mathbb{E}(y|b_i^{(j)} = -1) = \mu_i^{j+} - 2h_{qi+j}
$$

• Covariance of $y$

$$
C_i^j = \sigma^2 I_{2N_r p} + \sum_{l=0}^{2k-1} \sum_{m=0}^{q-1} h_{ql+m} h_{ql+m}^T 4p_l^{m+}(1 - p_l^{m+})
$$
**PDA Based Large-MIMO Detection** [15]

- Using $\mu_{i}^{j\pm}$ and $C_{i}^{j}$, $P(y|b_{i}^{(j)} = \pm 1)$ can be written as
  \[
P(y|b_{i}^{(j)} = \pm 1) = \frac{e^{-(y-\mu_{i}^{\pm})^T(C_{i}^{j})^{-1}(y-\mu_{i}^{\pm})}}{(2\pi)^{N_r}p|C_{i}^{j}|^{\frac{1}{2}}}
  \]

- Using (5), $\beta_{i}^{j}$ can be written as
  \[
  \beta_{i}^{j} \beta_{i}^{j} = e^{-((y-\mu_{i}^{j+})^T(C_{i}^{j})^{-1}(y-\mu_{i}^{j+})-(y-\mu_{i}^{j-})^T(C_{i}^{j})^{-1}(y-\mu_{i}^{j-}))}
  \]

- Compute $\Lambda_{i}^{(j)}$ using $\alpha_{i}^{(j)}$ and $\beta_{i}^{(j)}$
- Update the statistics of $b_{i}^{(j)}$ as
  \[
P(b_{i}^{(j)} = +1|y) = \frac{\Lambda_{i}^{(j)}}{1 + \Lambda_{i}^{(j)}}, \quad P(b_{i}^{(j)} = -1|y) = \frac{1}{1 + \Lambda_{i}^{(j)}}
  \]
- This completes one iteration of the algorithm
PDA Based Large-MIMO Detection [15]

- Updated values of $P(b_i^{(j)} = +1|y)$ and $P(b_i^{(j)} = -1|y)$ for all $i, j$ are fed back as a priori probabilities to the next iteration.

- Algorithm terminates after a certain number of iterations.

- At the end of the last iteration,
  - decide $\hat{b}_i^{(j)}$ as $+1$ if $\Lambda_i^{(j)} \geq 1$, and $-1$ otherwise.

- In coded systems
  - feed $\Lambda_i^{(j)}$'s as soft inputs to the decoder.
Performance of PDA in V-BLAST

* PDA algorithm also exhibits large-dimension behavior

Figure 7: BER performance of PDA based detection of V-BLAST MIMO for $N_t = N_r = 8, 16, 32, 64, 96$, 4-QAM, and $m = 5$ iterations.
Performance of PDA in NO-STBC [15]

(a) 4-QAM

(b) 16-QAM
Large-MIMO Detection Based on Graphical Models

- Belief propagation (BP) is proven to work in cycle-free graphs
- BP is often successful in graphs with cycles as well
- MIMO graphical models are fully/densely connected
- Graphical models with certain simplifications/assumptions work successfully in large-MIMO detection

1. Use of **Pairwise Markov Random Field (MRF)** based graphical model in conjunction with message/belief damping [16],[17]

2. Use of **Factor Graph (FG)** based graphical model with **Gaussian Approximation of Interference (GAI)** [17]
Performance of Damped BP on Pairwise MRF [17]

(c) Performance as a function of damping factor

(d) BP on pairwise MRF exhibits large-dimension behavior

- Damping significantly improves performance
- Order of per-symbol complexity: $O(N_t^2)$
Large-MIMO Detection using BP on FGs [17]

- Each entry of the vector $y$ is treated as a function node (observation node)
- Each symbol, $x_i \in \{\pm 1\}$, is treated as a variable node
- Key ingredient: Gaussian approximation of the interference

$$y_i = h_{ik}x_k + \sum_{j=1, j\neq k}^{2N_t} h_{ij}x_j + n_i,$$

is modeled as $\mathcal{CN}(\mu_{z_{ik}}, \sigma^2_{z_{ik}})$ with $\mu_{z_{ik}} = \sum_{j=1, j\neq k}^{N_t} h_{ij}E(x_j)$, and

$$\sigma^2_{z_{ik}} = \sum_{j=1, j\neq k}^{2N_t} |h_{ij}|^2 \text{var}(x_j) + \frac{\sigma^2}{2},$$

where $h_{ij}$ is the $(i, j)$th element in $H$. 
Large-MIMO Detection using BP on FGs [17]

- With $x_i$'s $\in \{\pm 1\}$, the log-likelihood ratio (LLR) of $x_k$ at observation node $i$, denoted by $\Lambda^k_i$, is

$$\Lambda^k_i = \log \frac{p(y_i|H, x_k = 1)}{p(y_i|H, x_k = -1)} = \frac{2}{\sigma^2_{z_{ik}}} \Re(h^*_ik(y_i - \mu_{z_{ik}}))$$

- LLR values computed at observation nodes are passed to variable nodes.
- Using these LLRs, variable nodes compute the probabilities

$$p^{k+}_i \triangleq p_i(x_k = +1|y) = \frac{\exp(\sum_{l\neq i} \Lambda^k_l)}{1 + \exp(\sum_{l\neq i} \Lambda^k_l)}$$

and pass them back to the observation nodes.
- This message passing is carried out for a certain number of iterations.
- At the end, $x_k$ is detected as

$$\hat{x}_k = \text{sgn}\left(\sum_{i=1}^{2Nr} \Lambda^k_i\right)$$
Message Passing on Factor Graphs [17]

Figure 8: Message passing between variable nodes and observation nodes.
Performance of BP on FGs with GAI [17]

- BP with GAI achieves near-optimal performance for increasing $N_t = N_r$ with $O(N_t)$ per-symbol complexity
Large-MIMO Applications/Standardization

• Potential Applications
  – Fixed Wireless IPTV/HDTV distribution (e.g., in 5 GHz band)
    * potentially big markets in India
  – High-speed back haul connectivity between BSs/BSCs using high data rate large-MIMO links (e.g., in 5 GHz band)
  – Wireless mesh networks

• Large-MIMO in Wireless Standards?
  – Multi-Gigabit Rate LAN/PAN (e.g., in 5GHz / 60 GHz band)
    * Evolution of WiFi standards (IEEE 802.11ac and 802.11ad)
  – LTE-Advanced, WiMax (IEEE 802.16m)
  – Can consider $12 \times 12$, $16 \times 16$, $24 \times 24$, $32 \times 32$ MIMO systems
Concluding Remarks

• Low-complexity detection
  – critical enabling technology for large-MIMO
  – no more a bottleneck

• Large-MIMO systems can be implemented

• Large-MIMO approach scores high on spectral efficiency and operating SNR compared to other approaches (e.g., increasing QAM size)

• Standardization efforts can consider reaping the benefits of large-MIMO in their evolution
Thank You