Augmented Lattice Reduction for MIMO decoding

LAURA LUZZI

joint work with G. Rekaya-Ben Othman and J.-C. Belfiore
at Télécom-ParisTech

NANYANG TECHNOLOGICAL UNIVERSITY
SEPTEMBER 15, 2010

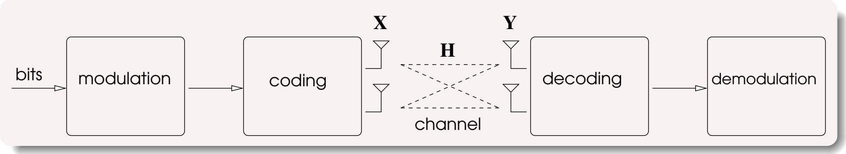
Motivation

- Multiple antenna systems allow to improve data rates and reliability.
- In order to increase data rates, both the number of antennas and the size of the signal set can be increased.
- This entails a high **decoding complexity** which is a real challenge for practical implementation.

Outline

1. MIMO systems
2. Decoding
   - Lattice decoding
   - Lattice reduction-aided decoding
3. Augmented lattice reduction
   - Method
   - Performance
   - Complexity
4. Open problems

MIMO systems



- multiplexing gain
  - send independent data on each antenna
  - improve the rate
- diversity gain
  - send independent copies of the same data
  - improve reliability

System model: spatial multiplexing

$$\underbrace{\mathbf{y}_{n\times 1}}_{\text{received signal}} = \underbrace{\mathbf{H}_{n\times m}}_{\text{channel}} \underbrace{\mathbf{x}_{m\times 1}}_{\text{codeword}} + \underbrace{\mathbf{w}_{n\times 1}}_{\text{noise}}$$

- m transmit antennas, n receive antennas
- $\mathbf{H}$, $\mathbf{w}$ are random matrices with i.i.d. complex Gaussian entries
- $x_i \in \mathcal{S} \subset \mathbb{Z}[i]$ is the signal transmitted by antenna i

System model: Space-Time Coding

$$\underset{\text{received signal}}{\mathbf{Y}_{n\times t}} = \underset{\text{channel}}{\mathbf{H}_{n\times m}} \; \underset{\text{codeword}}{\mathbf{X}_{m\times t}} + \underset{\text{noise}}{\mathbf{W}_{n\times t}}$$

- m transmit antennas, n receive antennas, t frame length
- codewords are represented by matrices or space-time blocks
- the matrix element $x_{i,j} \in \mathbb{C}$ represents the signal sent by antenna i at time j

Diversity order of a coding scheme

$$d = - \lim_{\mathrm{SNR}\to\infty} \frac{\log(P_e)}{\log(\mathrm{SNR})},$$

where P_e is the error probability, and SNR is the signal-to-noise ratio.

Maximum diversity order

- For spatial multiplexing:

$$d_{\max} = n \quad \text{(receive diversity)}$$

- For space-time coding:

$$d_{\max} = mn \quad \text{(transmit and receive diversity)}$$

Outline

1. MIMO systems
2. Decoding
   - Lattice decoding
   - Lattice reduction-aided decoding
3. Augmented lattice reduction
   - Method
   - Performance
   - Complexity
4. Open problems

Outline

1. MIMO systems
2. Decoding
   - Lattice decoding
   - Lattice reduction-aided decoding
3. Augmented lattice reduction
   - Method
   - Performance
   - Complexity
4. Open problems

Lattice point representation and decoding

Spatial multiplexing case

$$\mathbf{y} = \mathbf{H}\mathbf{x} + \mathbf{w}$$

The received vector is the translated version of a point in the lattice generated by $\mathbf{H}$.

Coded case

$$\mathbf{Y} = \mathbf{H}\mathbf{X} + \mathbf{W}$$

Equivalent lattice formulation after vectorizing matrices:

$$\mathbf{y} = \mathbf{H}_l \Phi \mathbf{s} + \mathbf{w}$$

- $\mathbf{H}_l$ linear map corresponding to left multiplication by $\mathbf{H}$
- Φ generator matrix of the code
- $\mathbf{s}$ vector of information signals

Decoding

Optimal decoding amounts to solving the Closest Vector Problem (CVP) in the lattice generated by $\mathbf{H}$:

$$\hat{\mathbf{x}} = \operatorname*{argmin}_{\mathbf{x}' \in \mathcal{S}} \left\| \mathbf{y} - \mathbf{H}\mathbf{x}' \right\|^2 \qquad \text{ML solution}$$

- ML decoders
  - Sphere Decoder, Schnorr-Euchner algorithm...
  - optimal performance but exponential complexity
- Suboptimal decoders
  - zero forcing (ZF), successive interference cancellation (SIC)...
  - polynomial complexity, but they don't attain maximal diversity

Channel preprocessing

Example: ZF decoding

$$\mathbf{y} = \mathbf{H}\mathbf{x} + \mathbf{w}$$
$$\hat{\mathbf{x}}_{\text{ZF}} = \left\lfloor \mathbf{H}^{-1}\mathbf{y} \right\rceil = \left\lfloor \mathbf{x} + \mathbf{H}^{-1}\mathbf{w} \right\rceil$$

- if $\mathbf{H}$ is orthogonal, ZF decoding is optimal
- if $\mathbf{H}$ is ill-conditioned, the noise $\mathbf{H}^{-1}\mathbf{w}$ is amplified

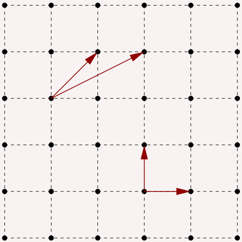
- **Solution:** channel preprocessing by lattice reduction improves the performance of suboptimal decoders

Outline

1. MIMO systems
2. Decoding
   - Lattice decoding
   - Lattice reduction-aided decoding
3. Augmented lattice reduction
   - Method
   - Performance
   - Complexity
4. Open problems

Lattice reduction

Two matrices $\mathbf{H}$ and $\mathbf{H}'$ generate the same lattice if $\mathbf{H}' = \mathbf{HT}$ with $\mathbf{T}$ unimodular.



Lattice reduction:
find a lattice basis formed by "short" and "nearly orthogonal" vectors

- the most popular lattice reduction algorithm is the LLL algorithm, thanks to its polynomial complexity

The LLL algorithm

Gram-Schmidt Orthogonalization (GSO)

$\mathbf{h}_1^* \leftarrow \mathbf{h}_1$
for $i = 2, \ldots, m$ **do**
  for $j = 1, \ldots, i-1$ **do**
    $\mu_{i,j} \leftarrow \frac{\langle \mathbf{h}_i, \mathbf{h}_j^* \rangle}{\left\| \mathbf{h}_j^* \right\|^2}$
  end
  $\mathbf{h}_i^* \leftarrow \mathbf{h}_i - \sum_{j=1}^{i-1} \mu_{i,j} \mathbf{h}_j^*$
end

LLL-reduced basis

$\mathbf{H}$ is LLL-reduced if its GSO satisfies the following properties:

- Size reduction: $|\mu_{k,l}| \leq \frac{1}{2}, \quad 1 \leq l < k \leq m,$
- Lovasz condition: $\left\| \mathbf{h}_k^* + \mu_{k,k-1} \mathbf{h}_{k-1}^* \right\|^2 \geq \frac{3}{4} \left\| \mathbf{h}_{k-1}^* \right\|^2, \quad 1 < k \leq m$

```
Compute GSO
k ← 2
while k ≤ m do
    Size reduction RED(k,k-1)
    if Lovasz(k,k-1) is not satisfied then
        swap h_k and h_{k-1}
        update GSO
        k ← max(k − 1, 2)
    end
    else
        for l = k − 2, . . . , 1 do
            Size reduction RED(k,l)
        end
        k ← k + 1
    end
end
```

Preprocessing using LLL reduction

m transmit antennas, n receive antennas

$$\underbrace{\mathbf{y}_{n\times 1}}_{\text{received signal}} = \underbrace{\mathbf{H}_{n\times m}}_{\text{channel}} \underbrace{\mathbf{x}_{m\times 1}}_{\text{codeword}} + \underbrace{\mathbf{w}_{n\times 1}}_{\text{noise}}$$

$$\mathbf{H}_{\text{red}} = \mathbf{H}\mathbf{U} \quad \text{LLL-reduced form}$$

Preprocessing using LLL reduction

m transmit antennas, n receive antennas

$$\underbrace{\mathbf{y}_{n\times 1}}_{\text{received signal}} = \underbrace{\mathbf{H}_{n\times m}}_{\text{channel}} \underbrace{\mathbf{x}_{m\times 1}}_{\text{codeword}} + \underbrace{\mathbf{w}_{n\times 1}}_{\text{noise}}$$

$$\mathbf{H}_{\text{red}} = \mathbf{HU} \quad \text{LLL-reduced form}$$

LLL-ZF decoder

- compute the pseudo-inverse $\mathbf{H}_{\text{red}}^{\dagger}$
- $\hat{\mathbf{x}}_{\text{LLL}-\text{ZF}} = \mathbf{U}\left(\left\lfloor \mathbf{H}_{\text{red}}^{\dagger}\mathbf{y}\right\rceil\right)$

Preprocessing using LLL reduction

m transmit antennas, n receive antennas

$$\underset{\text{received signal}}{\mathbf{y}_{n\times 1}} = \underset{\text{channel}}{\mathbf{H}_{n\times m}} \; \underset{\text{codeword}}{\mathbf{x}_{m\times 1}} + \underset{\text{noise}}{\mathbf{w}_{n\times 1}}$$

$$\mathbf{H}_{\text{red}} = \mathbf{HU} \quad \text{LLL-reduced form}$$

LLL-ZF decoder

- compute the pseudo-inverse $\mathbf{H}_{\text{red}}^{\dagger}$
- $\hat{\mathbf{x}}_{\text{LLL}-\text{ZF}} = \mathbf{U}\left(\left\lfloor \mathbf{H}_{\text{red}}^{\dagger}\mathbf{y} \right\rceil\right)$

LLL-SIC decoder

- QR decomposition of $\mathbf{H}_{\text{red}}$
- $\tilde{\mathbf{y}} = \mathbf{Q}^H\mathbf{y} = \mathbf{Rx} + \mathbf{Q}^H\mathbf{w}$
- recursively compute $\tilde{x}_m = \left\lfloor \frac{\tilde{y}_m}{r_{mm}} \right\rceil$,
  $\tilde{x}_i = \left\lfloor \frac{\tilde{y}_i - \sum_{j=i+1}^{m} r_{ij}\tilde{x}_j}{r_{ii}} \right\rceil, \quad i = m-1, \dots, 1$
- $\hat{\mathbf{x}}_{\text{LLL}-\text{SIC}} = \mathbf{U}\tilde{\mathbf{x}}$

Performance of LLL-aided decoding

Proposition [Taherzadeh, Mobasher, Khandani 2007]

LLL-ZF and LLL-SIC decoders attain the maximal receive diversity n.

- the average complexity of LLL-aided decoding is polynomial, which makes it an attractive technique
- however, LLL becomes less effective for high-dimensional lattices
- as a consequence, the performance gap with respect to ML decoding increases greatly when the number of antennas is large

Outline

1. MIMO systems
2. Decoding
   - Lattice decoding
   - Lattice reduction-aided decoding
3. Augmented lattice reduction
   - Method
   - Performance
   - Complexity
4. Open problems

Outline

Augmented lattice reduction - principle

- new technique which combines preprocessing and detection: lattice reduction is used directly to decode
- MIMO decoding amounts to solving the **closest vector problem** (CVP) in the lattice generated by the channel matrix
- reduce the CVP to the SVP (**shortest vector problem**)
- use **LLL reduction** to solve the SVP in polynomial time

Embedding method to reduce the CVP to the SVP

- Follow Kannan's approach (1987): **embed** the m-dimensional lattice generated by $\mathbf{H}$ into a suitable $(m+1)$-dimensional lattice

$$\widetilde{\mathbf{H}} = \begin{pmatrix} \mathbf{H} & -\mathbf{y} \\ \mathbf{0} & t \end{pmatrix}$$

augmented matrix

- $\mathbf{v} = \begin{pmatrix} \mathbf{Hx} - \mathbf{y} \\ t \end{pmatrix} = \begin{pmatrix} \mathbf{w} \\ t \end{pmatrix} = \widetilde{\mathbf{H}} \begin{pmatrix} \mathbf{x} \\ 1 \end{pmatrix}$

- **Strategy**: if $\|\mathbf{w}\|$ and t are "very small", $\mathbf{v}$ is the shortest vector in the augmented lattice. By finding $\mathbf{v}$, we recover $\mathbf{x}$.

Using LLL reduction to find the shortest lattice vector

- LLL-reduce $\widetilde{\mathbf{H}}$: $\boxed{\widetilde{\mathbf{H}}_{\text{red}} = \widetilde{\mathbf{H}}\widetilde{\mathbf{U}}}$

- **Problem:** in general, there's no guarantee that the shortest vector belongs to the reduced basis. However, the first column of $\widetilde{\mathbf{H}}_{\text{red}}$ satisfies

$$\left\|\widetilde{\mathbf{h}}_1^{\text{red}}\right\| \leq 2^{\frac{m}{2}} d_{\widetilde{\mathbf{H}}}$$

- $\mathbf{v}$ shortest vector in $\mathcal{L}$ is called α-unique if $\forall \mathbf{u} \in \mathcal{L}$,

$$\|\mathbf{u}\| \leq \alpha \|\mathbf{v}\| \Rightarrow \mathbf{u}, \mathbf{v} \quad \text{linearly dependent}$$

Exponential gap technique

$\mathbf{v}$ is $2^{\frac{m}{2}}$-unique $\Rightarrow$ $\pm\mathbf{v}$ is the first column of $\widetilde{\mathbf{H}}_{\text{red}}$

LLL reduction finds the shortest lattice vector

$$a(\mathbf{H}) \doteqdot \min_{1 \leq i \leq m} \|\mathbf{h}_i^*\| \quad \Rightarrow \quad \frac{d_{\mathbf{H}}}{2^{\frac{m-1}{2}}} \leq a(\mathbf{H}_{\text{red}}) \leq d_{\mathbf{H}}$$

Lemma

Let $t = \frac{a(\mathbf{H}_{\text{red}})}{2^{m+1}}$, and suppose that $\|\mathbf{w}\| \leq \frac{d_{\mathbf{H}}}{2^{m+1}}$.

Then $\mathbf{v} = \begin{pmatrix} \mathbf{Hx} - \mathbf{y} \\ t \end{pmatrix}$ is a $2^{\frac{m}{2}}$-unique shortest vector in $\mathcal{L}(\widetilde{\mathbf{H}})$.

Sketch of the proof:

Suppose $\exists \mathbf{u} = \begin{pmatrix} \mathbf{Hx}' - q\mathbf{y} \\ qt \end{pmatrix}$ such that $\|\mathbf{u}\| \leq 2^{\frac{m}{2}} \|\mathbf{v}\|$, and $\mathbf{u}, \mathbf{v}$ linearly independent

- $\|\mathbf{u}\| \geq |q|\, t \;\Rightarrow\; |q| \leq \frac{2^{\frac{m}{2}} \|\mathbf{v}\|}{t}$
- $\|\mathbf{H}(\mathbf{x}' - q\mathbf{x})\| \leq \|\mathbf{Hx}' - q\mathbf{y}\| + |q| \, \|\mathbf{y} - \mathbf{Hx}\| \leq 2^{\frac{m}{2}} \|\mathbf{v}\| + \frac{2^{\frac{m}{2}} \|\mathbf{v}\|}{t} \|\mathbf{w}\| \leq$
  $\leq 2^{\frac{m}{2}} \sqrt{\|\mathbf{w}\|^2 + t^2} \left(1 + \frac{\|\mathbf{w}\|}{t}\right) < d_{\mathbf{H}} \quad \Rightarrow \quad$ contradiction.

Augmented lattice reduction- Decoding

- Choose $t = \frac{a(\mathbf{H}_{\text{red}})}{2^{m+1}} \quad \Rightarrow \quad$ the augmented lattice has an exponential gap

- LLL-reduce $\widetilde{\mathbf{H}}$: $\quad \widetilde{\mathbf{H}}_{\text{red}} = \widetilde{\mathbf{H}}\widetilde{\mathbf{U}}$

- If $\|\mathbf{w}\| \leq \frac{d_{\mathbf{H}}}{2^{m+1}}$, $\quad \mathbf{v} = \widetilde{\mathbf{H}} \begin{pmatrix} \mathbf{x} \\ 1 \end{pmatrix}$ is the first column of $\widetilde{\mathbf{H}}_{\text{red}}$

- find the transmitted message $\mathbf{x}$ on the first column of $\widetilde{\mathbf{U}}$:

$$\boxed{\hat{\mathbf{x}} = \frac{1}{\widetilde{u}_{m+1,1}} (\widetilde{u}_{1,1}, \ldots, \widetilde{u}_{m,1})^T}$$

Outline

Performance analysis

Diversity gain

$$d = - \lim_{\mathrm{SNR} \to \infty} \frac{\log(P_e)}{\log(\mathrm{SNR})}$$

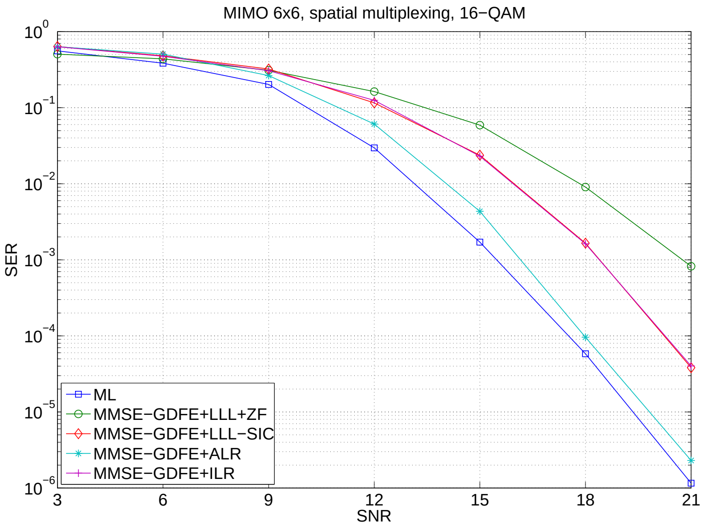
Proposition

If $t = \frac{a(\mathbf{H}_{\mathrm{red}})}{2^{m+1}}$, then augmented lattice reduction achieves the maximum receive diversity n.

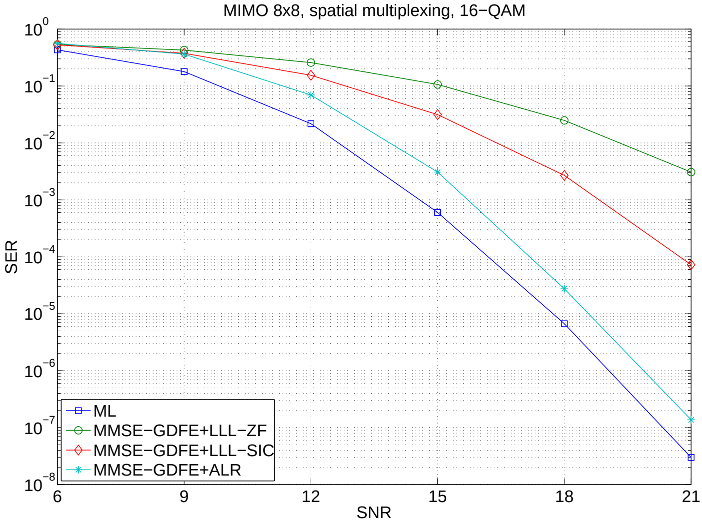
Sketch of the proof.

$$P_e(\mathbf{H}) \leq P\left\{ \|\mathbf{w}\| > \frac{d_{\mathbf{H}}}{2^{m+1}} \right\} \leq \frac{C(\log(\mathrm{SNR}))^{n+1}}{\mathrm{SNR}^n}$$

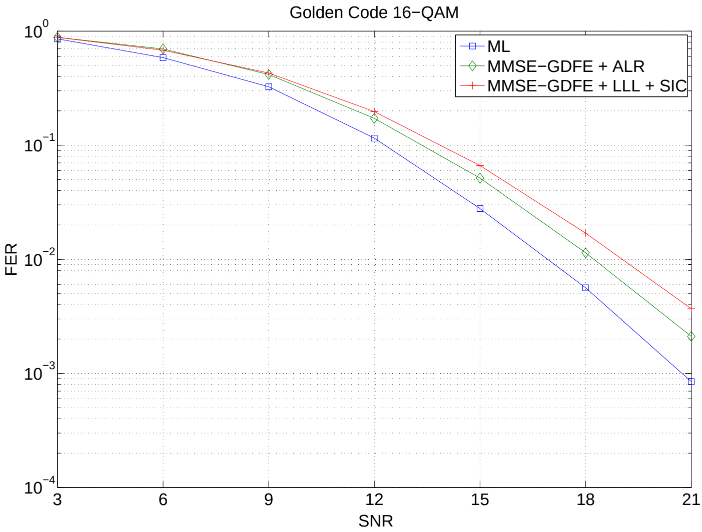
6×6 MIMO system, spatial multiplexing



8×8 MIMO system, spatial multiplexing



Coded case: 2×2 system, Golden Code



Comparison with Kim and Park's Method

[N. Kim, H. Park, "Improved lattice reduction aided detections for MIMO systems", *Vehicular Technology Conference* 2006]

- **Same form** of the augmented matrix $\widetilde{\mathbf{H}}$
- **No exponential gap technique:**
  - the parameter t is "big" ($t > \max |r_{ii}|$, where $\mathbf{H}_{\text{red}} = \mathbf{QR}$)
  - the solution is found on the last column of $\widetilde{\mathbf{U}}$
- no guarantee that LLL reduction will find the right vector. In fact, one can prove that the performance is the **same as LLL-SIC**.

Outline

Complexity analysis (for $n = m$)

- average number of iterations of the LLL algorithm:

$$\mathbb{E}[K(\mathbf{H})] \sim O\left(n^2 \log n\right) \qquad \text{[Jalden \textit{et al.} 2008]}$$

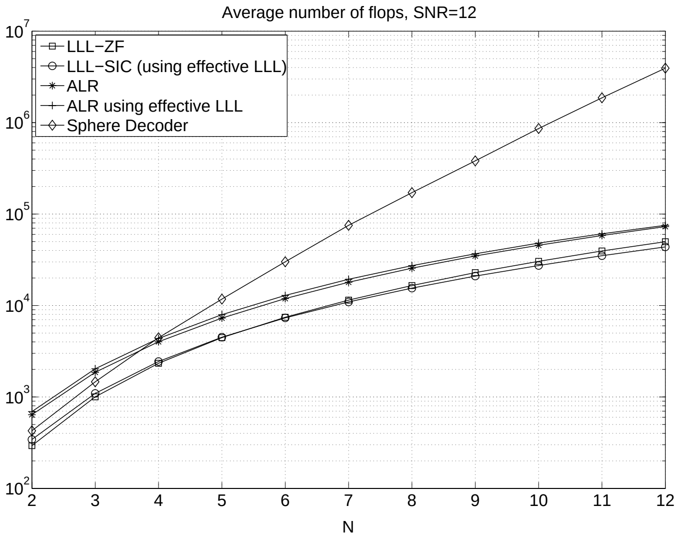
- each iteration requires $O(n^2)$ operations, which can be reduced to $O(n)$ for LLL-SIC [Ling, Howgrave-Graham 2007]

$\Rightarrow$ the total complexity of LLL-SIC is bounded by $O(n^3 \log n)$

Complexity bound for augmented lattice reduction

- $\mathbb{E}[K(\widetilde{\mathbf{H}})] \leq O(n^3)$
- the total complexity is bounded by $O(n^4)$.

Complexity: numerical simulations



Outline

1. MIMO systems
2. Decoding
   - Lattice decoding
   - Lattice reduction-aided decoding
3. Augmented lattice reduction
   - Method
   - Performance
   - Complexity
4. Open problems

Open problems

- explain *why* the performance of augmented lattice reduction is better than LLL-SIC
- find more realistic bounds on the complexity
- for high-dimensional space-time codes, the gap between ML decoding and augmented lattice reduction is still huge.
  How can we bridge this gap?

Thank you for listening!!