Low-complexity hybrid QRD-MCMC MIMO detection

Rong-Hui Peng, Koon Hoo Teo, Jinyun Zhang and Rong-Rong Chen

Mitsubishi Electric Research Laboratories
Department of Electrical and Computer Engineering
University of Utah

2011/7/16

Outline

- Channel model
- Review of MIMO detection
- QRD-M and MCMC detector
- Hybrid MIMO detector
- Complexity analysis
- Application to IEEE802.16e system
- Conclusion

Channel model

- Consider a MIMO Rayleigh fast fading channels.

$$\mathbf{y} = \mathbf{H}\mathbf{d} + \mathbf{n}$$

where

$\mathbf{d} \in \mathbf{C}^{N_t}$ is a vector of transmit symbols

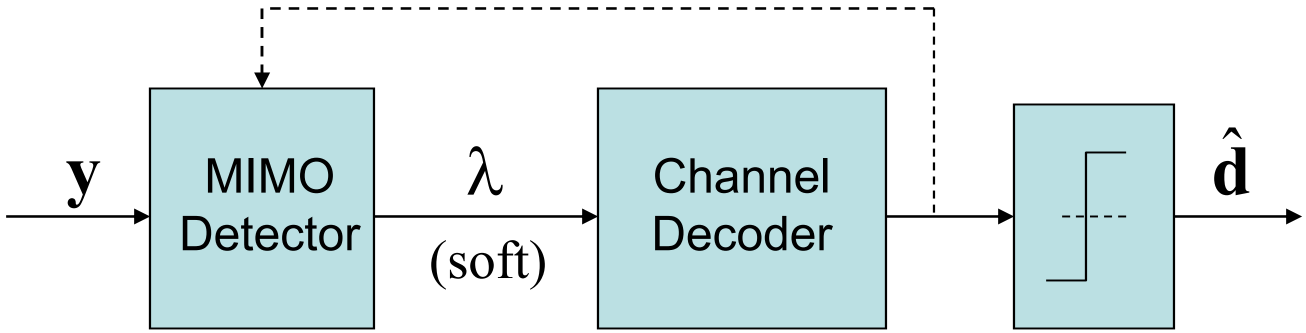
$\mathbf{y} \in \mathbf{C}^{N_r}$ is a vector of received signal

$\mathbf{H} \in \mathbf{C}^{N_r \times N_t}$ is the channel gain matrix

$\mathbf{n} \in \mathbf{C}^{N_r}$ is an additive noise vector

N_t, N_r is the number of tx and rx antennas

Receiver structure



• Separate detection and decoding (SDD) : no feedback from channel decoder

• Joint detection and decoding (JDD) : exchange soft information between detector and decoder

MIMO detection

- Maximal likelihood (ML) detection or Maximum *a posteriori* (MAP) is optimal
- Optimal detection usually has exponential complexity and is computation infeasible for practical system
- Low complexity sub-optimal detectors
  - ZF, MMSE, VBLAST …
- Approximate optimal detectors
  - Tree search based (sphere decoding, QRD-M)
  - Markov chain Monte Carlo (MCMC) based

Optimal MAP detection

- Soft output detector generate soft message, usually log likelihood ratio (LLR). It will be used by soft channel decoder
  - MAP:

$$\lambda_k = \ln \frac{P(b_k = +1 \mid \mathbf{y})}{P(b_k = -1 \mid \mathbf{y})} = \ln \frac{\sum_{\mathbf{b}_{-k}} P(b_k = +1, \mathbf{b}_{-k} \mid \mathbf{y})}{\sum_{\mathbf{b}_{-k}} P(b_k = -1, \mathbf{b}_{-k} \mid \mathbf{y})} \approx \ln \frac{\max_{\mathbf{b}_{-k}} P(b_k = +1, \mathbf{b}_{-k} \mid \mathbf{y})}{\max_{\mathbf{b}_{-k}} P(b_k = -1, \mathbf{b}_{-k} \mid \mathbf{y})}$$

where

$$\mathbf{d} = (b_1, b_2, \Lambda, b_{N_t M_c}); \quad \mathbf{b}_{-k} = (b_1, \Lambda\, b_{k-1}, b_{k+1}, \Lambda, b_{N_t M_c}); \quad b_i \in \{-1, 1\}$$

Low complexity detection

- Approximate optimal detectors

$$\lambda_k = \ln \frac{P(b_k = +1 \mid \mathbf{y})}{P(b_k = -1 \mid \mathbf{y})} = \ln \frac{\sum_{\mathbf{b}_{-k}} P(b_k = +1, \mathbf{b}_{-k} \mid \mathbf{y})}{\sum_{\mathbf{b}_{-k}} P(b_k = -1, \mathbf{b}_{-k} \mid \mathbf{y})}$$

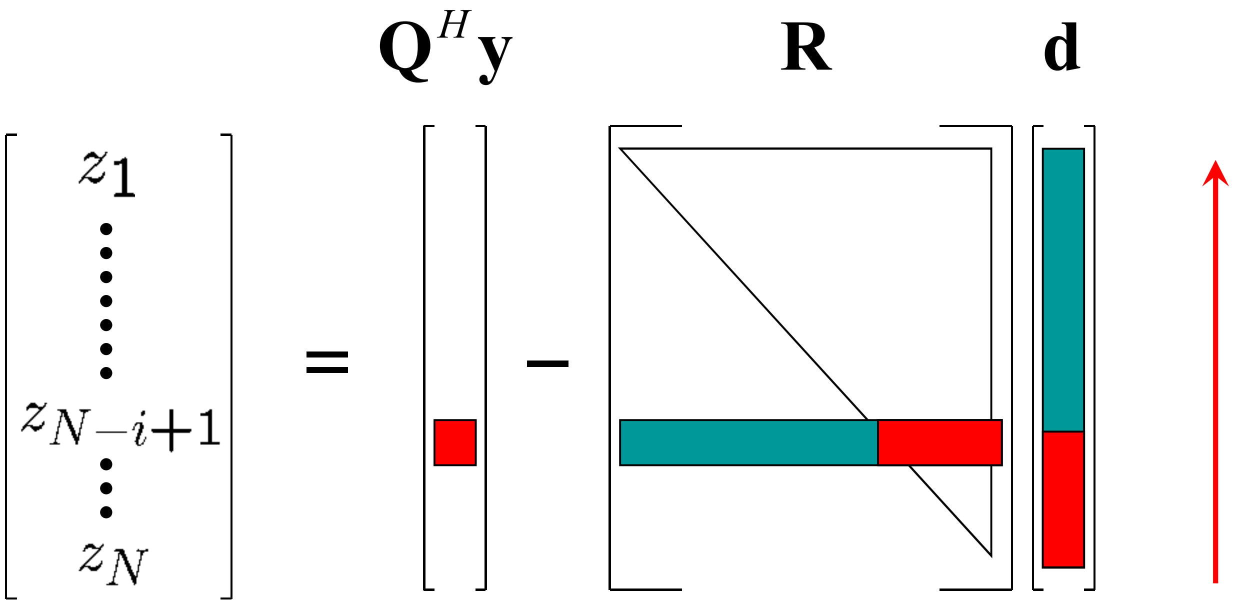
$$\approx \ln \frac{\max_{\mathtt{I}_{-k}} P(b_k = +1, \mathbf{b}_{-k} \mid \mathbf{y})}{\max_{\mathtt{I}_{-k}} P(b_k = -1, \mathbf{b}_{-k} \mid \mathbf{y})}$$

$$\text{where } \mathtt{I}_{-k} \subset \mathbf{b}_{-k}$$

The searching is performed over a small subset (*Important set*) instead of a large full set.

Tree search based detection

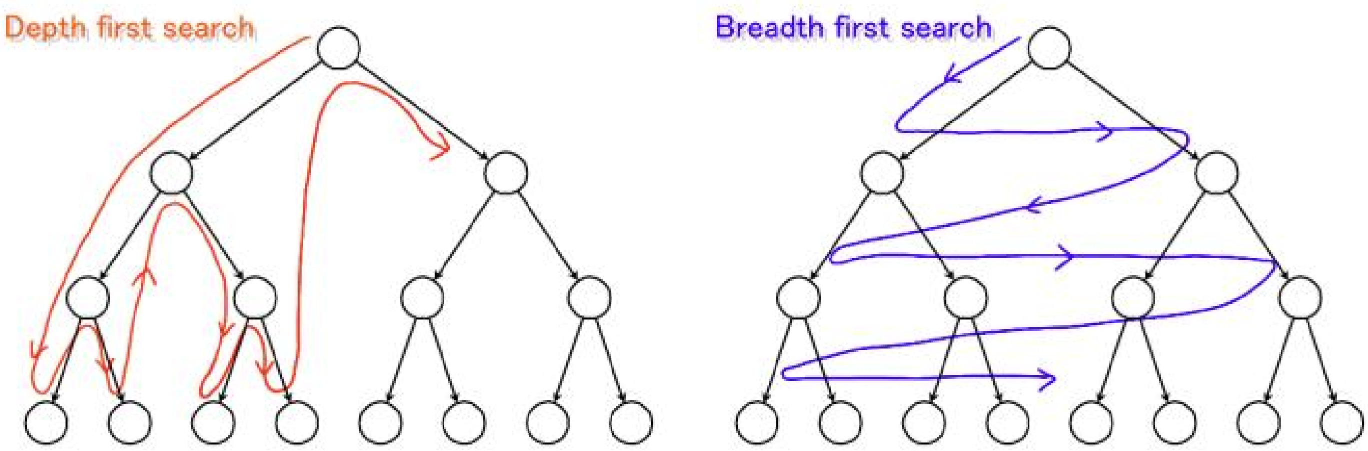
$$r^2 = \|\mathbf{y} - \mathbf{Hd}\|^2 = \|\mathbf{y} - \mathbf{QRd}\|^2 = \|\mathbf{Q}^H\mathbf{y} - \mathbf{Rd}\|^2 \quad (N_t \le N_r)$$



$$\lambda_i = |z_{N-i+1}|^2$$

Sphere decoding VS. QRD-M

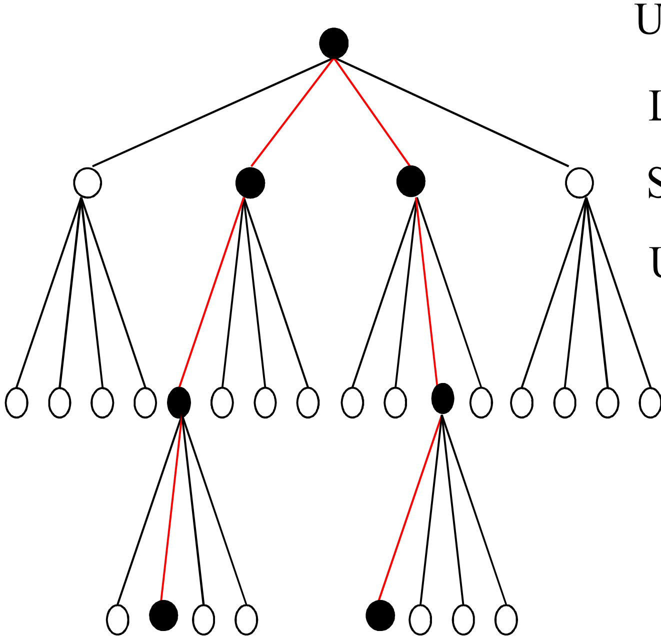
- Sphere decoding [1]: Depth-First Search (DFS)
  - Search the tree inside the sphere
  - Variable throughput with average polynomial complexity but exponential at low SNR
- QRD-M [2]: Breadth-First Search (BFS)
  - At each layer, select *M* minimal paths
  - Fixed number of visited nodes, constant throughput



[1] B. M. Hochwald and S. ten Brink, “Achieving near-capacity on a multiple antenna channel,” *IEEE Trans. Commun.*, vol. 51, no. 3, pp. 389–399, Mar. 2003.

[2] K. J. Kim and J. Yue, “Joint channel estimation and data detection algorithms for MIMO-OFDM systems,” in *Thirty-Sixth Asilomar Conference on Signals, Systems and Computers*, 2002, pp. 1857–1861.

QRD-M



$$\text{L1}: \lambda_4 = \left| r_4 - t_{4,4} d_4 \right|^2 - 2\sigma^2 p(d_4)$$

Sorting λ_4, select 2 out of 4

Update $r_k = r_k - t_{k,4}\hat{d}_4, k = 1,2,3$

$$\text{L2}: \lambda_3 = \left| r_3 - t_{3,3} d_3 \right|^2 - 2\sigma^2 p(d_3) + \lambda_4$$

Sorting λ_3, select 2 of 8

Update $r_k = r_k - t_{k,3}\hat{d}_3, k = 1,2$

$$\text{L3}: \lambda_2 = \left| r_2 - t_{2,2} d_2 \right|^2 - 2\sigma^2 p(d_2) + \lambda_3$$

Sorting λ_2, select 2 of 8

Update $r_k = r_k - t_{k,2}\hat{d}_2, k = 1$

$$\text{L4}: \lambda_1 = \left| r_1 - t_{1,1} d_1 \right|^2 - 2\sigma^2 p(d_1) + \lambda_2$$

Sorting λ_1, select 2 of 8

QRD-M in a 4x4 QPSK system

Complexity of QRD-M

- Need preprocessing QR decomposition, $O(N_t^3)$
- At each layer (except root layer), MM_c square euclidian distance calculations are needed to find M minimal distance from MM_c path where M_c is the number of symbols in a constellation
- The complexity of QRD-M depends on the parameter M,
- If M is too large, the complexity is very high; if M is too small, promising candidates have been discarded before the process proceeds to the lowest layer

MCMC

- MCMC detector [3] finds important sets using Markov chain Monte Carlo
  - Create Markov chain with state space: **d** and stationary distribution P(**d**|**Y**)
  - Run Markov chain using Gibbs sampler
  - After Markov chain converge, the samples are generated according to P(**d**|**Y**). Those samples with large P(**d**|**Y**) generated with high probabilities

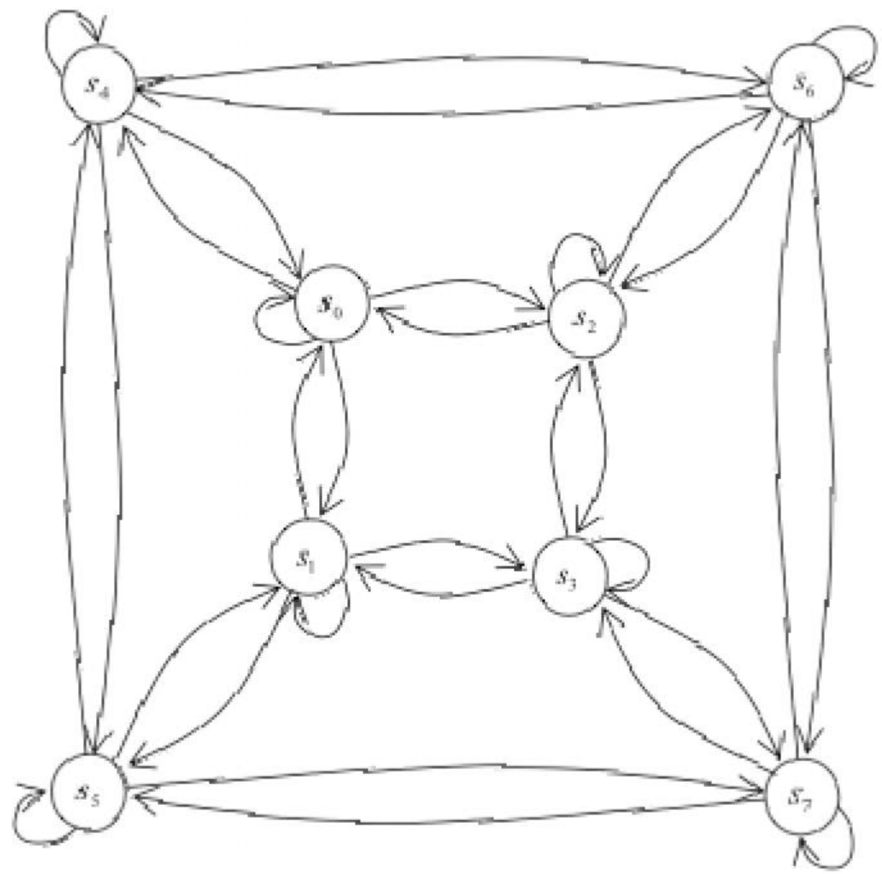
[3] B. Farhang-Boroujeny, H. Zhu, and Z. Shi, "Markov chain Monte Carlo algorithms for CDMA and MIMO communication systems," IEEE Trans. Signal Processing

Gibbs sampler

- Run full Markov chain is impossible because of huge number of states

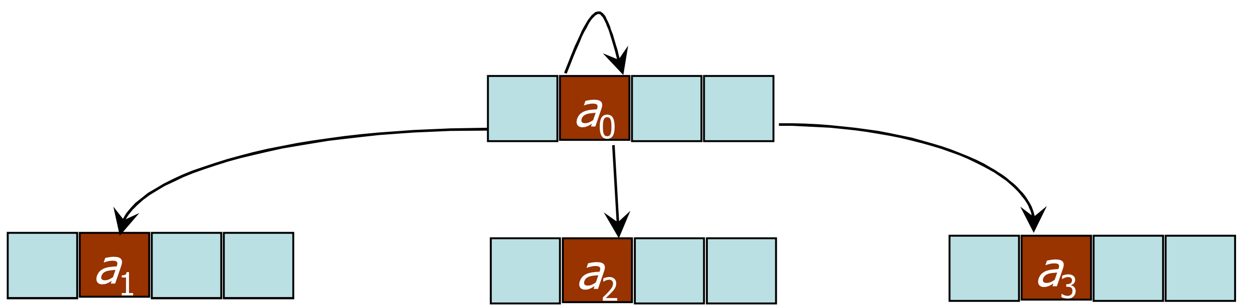
$$\mathbf{d} = \begin{bmatrix} d_1 \\ d_2 \\ d_3 \end{bmatrix}$$

State	d_3	d_2	d_1
S_0	-1	-1	-1
S_1	-1	-1	+1
S_2	-1	+1	-1
S_3	-1	+1	+1
S_4	+1	-1	-1
S_5	+1	-1	+1
S_6	+1	+1	-1
S_7	+1	+1	+1



Gibbs sampler

- Gibbs sampler limit the states jumping with only one variable change for each state jumping



Gibbs sampler

Generate initial $\mathbf{d}^{(0)}$ randomly

for n = 1 to I

generate $d_0^{(n)}$ from distribution $p(d_0 = b \mid d_1^{(n-1)}, d_2^{(n-1)}, \Lambda, d_{NM_c-1}^{(n-1)}, \mathbf{y})$

generate $d_1^{(n)}$ from distribution $p(d_1 = b \mid d_0^{(n)}, d_2^{(n-1)}, \Lambda, d_{NM_c-1}^{(n-1)}, \mathbf{y})$

M

generate $d_{NM_c-1}^{(n)}$ from distribution $p(d_{NM_c-1} = b \mid d_0^{(n)}, d_1^{(n)}, \Lambda, d_{NM_c-2}^{(n)}, \mathbf{y})$

end for

where $p(d_i = b \mid d_0^{(n)}, \Lambda\, d_{i-1}^{(n)}, d_{i+1}^{(n-1)}, \Lambda, d_{NM_c-1}^{(n-1)}, \mathbf{y})$

$$\propto p(\mathbf{y} \mid d_0^{(n)}, \Lambda\, d_{i-1}^{(n)}, d_i = b, d_{i+1}^{(n-1)}, \Lambda, d_{NM_c-1}^{(n-1)},) p(d_i = b)$$

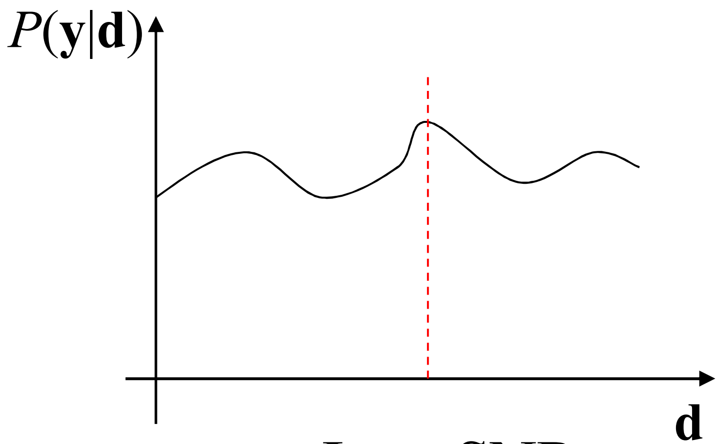
M_c is the # of bits per constellation symbol

MCMC

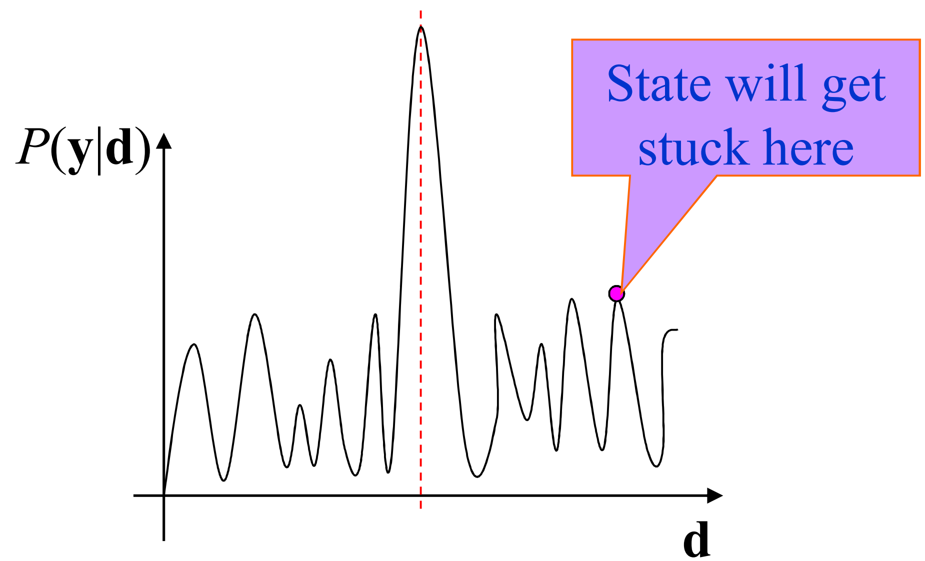
- The samples generated by Gibbs sampler are used to compute the soft message for soft decoder
- To accelerate Markov chain converge, *L* independent parallel Gibbs samplers are runned and each Gibbs sampler run *I* iterations

High SNR problem

- At high SNR, MCMC takes long time to converge, leads performance degradation, this is because of the multimodal property of channel PDF at high SNR



Low SNR



High SNR

Solutions

- Multimodal problem exists in many MCMC algorithm.
- No general method to overcome it
- For MIMO detection, one solution is to generate initial candidates using other low complexity detector (warm start)

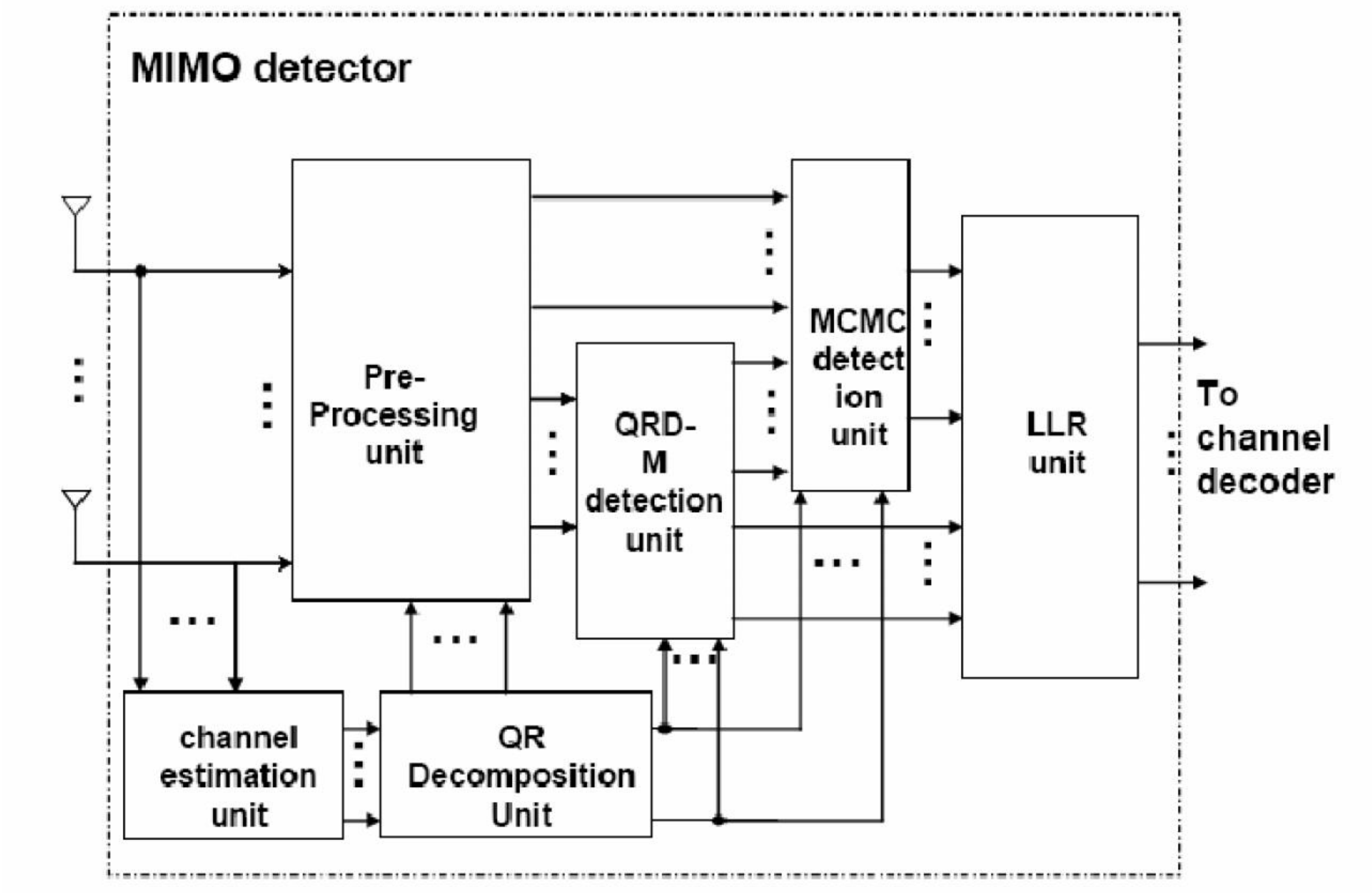
Tree search VS. MCMC

- Tree search based detection pruning paths on the tree
  - Exponential complexity at low SNR
  - Complexity is increased quickly with the dimension of problem
- MCMC finds important vectors using the P(**d**|**Y**)
  - Works very well at low SNR
  - Complexity is independent of SNR
  - Complexity is increased not too much with the dimension of problem
  - At high SNR, need the help of ZF or MMSE

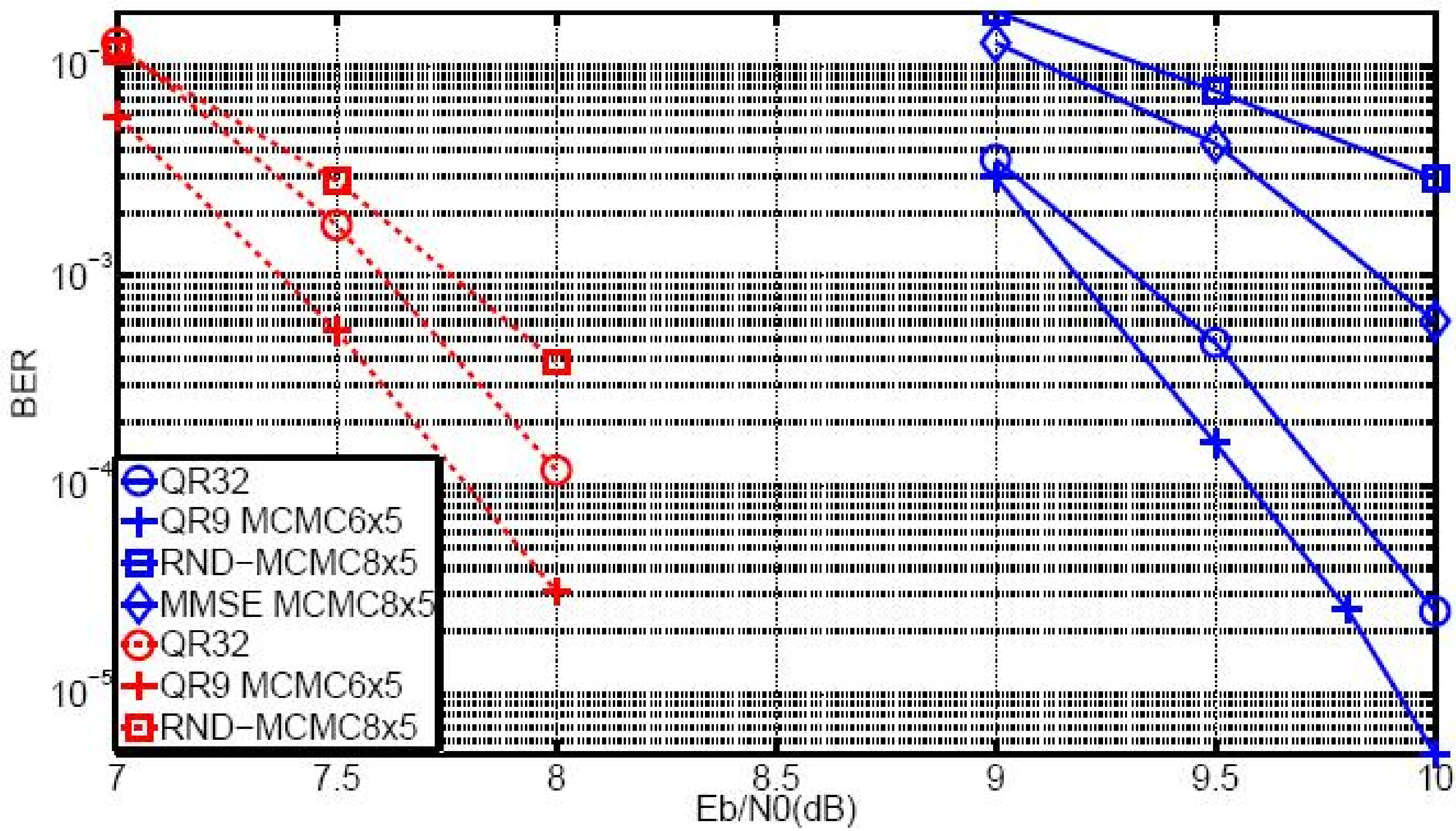
Hybrid QRD-MCMC

- Combine QRD-M and MCMC
  - A QRD-M with a small M is running first to generate initial important sets
  - The bit sequence with minimal path metric will be used to initialize one of L parallel MCMC
  - The important set produced by the QRD-M detector is incorporated by the MCMC detector
  - MCMC is running to generate refined important set
  - The soft message is computed using refined import set

Hybrid QRD-MCMC

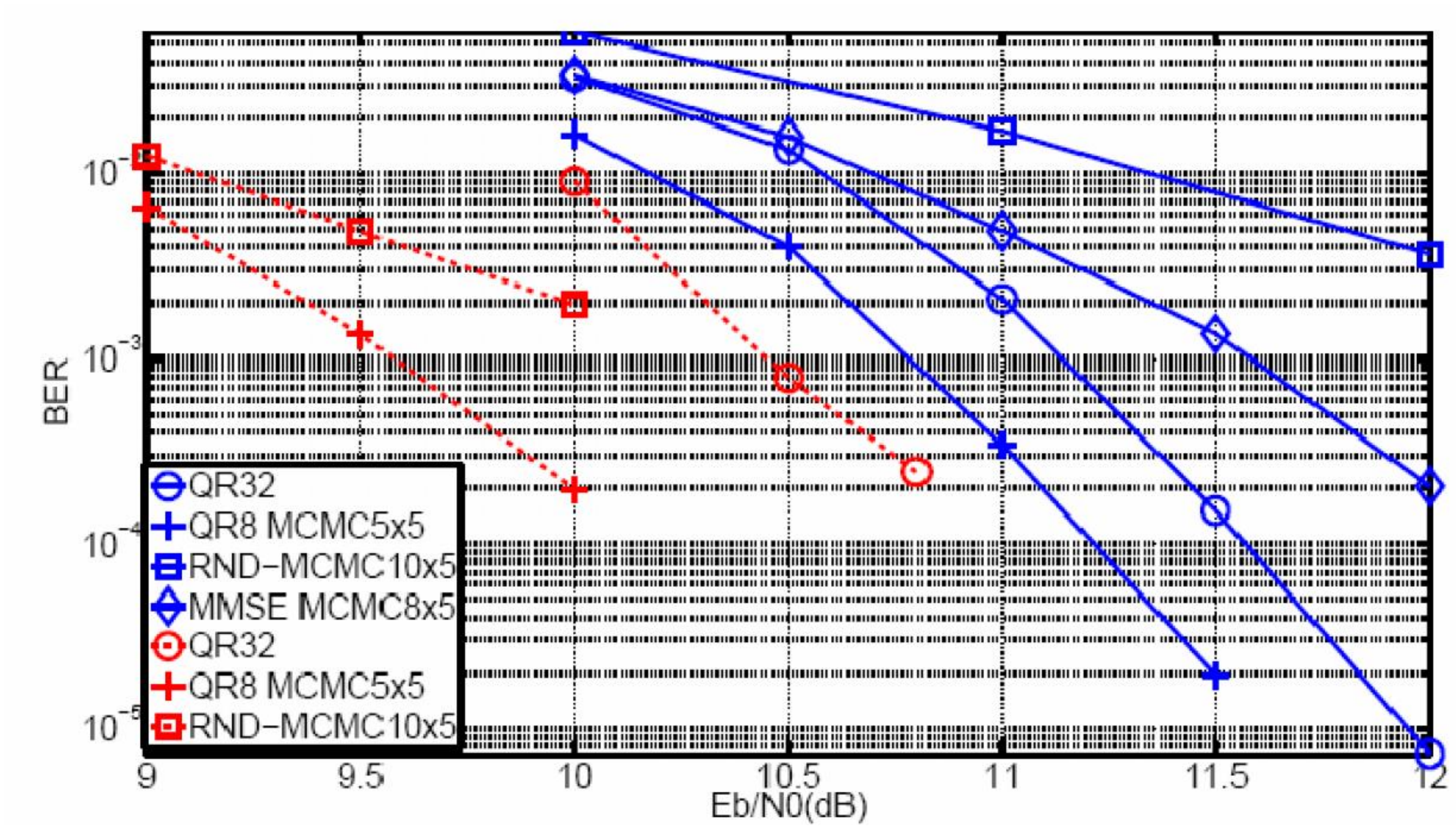


Results



Performance comparison of 4x4 16QAM SDD and JDD systems

Results



Performance comparison of 8x8 16QAM SDD and JDD systems

Complexity

- With the aid of QRD-M initialization, the MCMC detector starts from good initial vectors, which reduces the number of required parallel Gibbs samplers D and the number of iterations L per Gibbs sampler.
- Due to the use of MCMC, a small M is sufficient for the QRD-M detection, leading to reduced complexity and delay.
- Due to the QR decomposition of the channel matrix, the operations needed to compute path metric in MCMC detection can be reduced at least by 1/2

Complexity

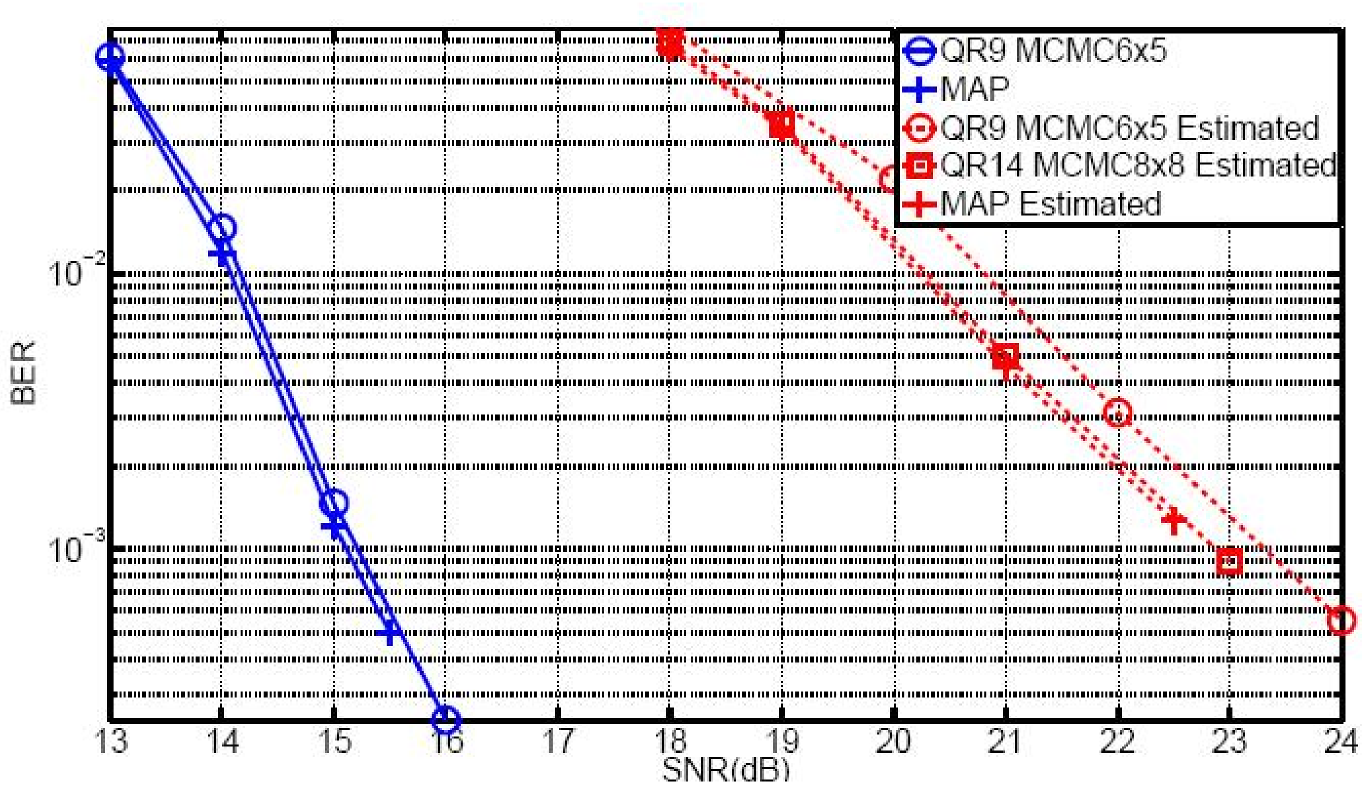
QRD-M	4x4 M=32	50256
	8x8 M=32	127344
RND-MCMC	4x4 $D = 8$ $L = 5$	73760
	8x8 $D = 10$ $L = 5$	331920
QRD-MCMC	4x4 $M = 9$ $D = 6$ $L = 5$	41498
	8x8 $M = 8$ $D = 5$ $L = 5$	114284

Application in 802.16e

SYSTEM PARAMETERS

Parameter	Value
Channel bandwidth	10 MHz
Number of subcarriers	1024
Subcarrier permutation	PUSC
Cyclic prefix	1/8
Channel coding	Convolutional turbo codes
Carrier frequency	2500 MHz
Sampling frequency	11.2 MHz
Multipath channel	ITU VehA
MS speed	120 km/hr

Results



Performance comparison of 4x4 16QAM MIMO-OFDMA system using R = 1/2 IEEE 802.16e convolutional turbo codes with perfect and 2D MMSE channel estimation.

Conclusion

- A hybrid MIMO detector is proposed: concatenation of QRD-M and MCMC
- The proposed detector reaps the advantages of QRD-M and MCMC
  - Work at wide range of SNR
  - Better performance and lower complexity
- Application to a practical IEEE802.16e system shows near optimal performance and is a good competing MIMO detector

Thank you !