Markov Chain Monte Carlo: Applications to MIMO detection and channel equalization

Rong-Rong Chen, Ronghui Peng, Behrouz Farhang-Boroujeny

Dept. of ECE., University of Utah

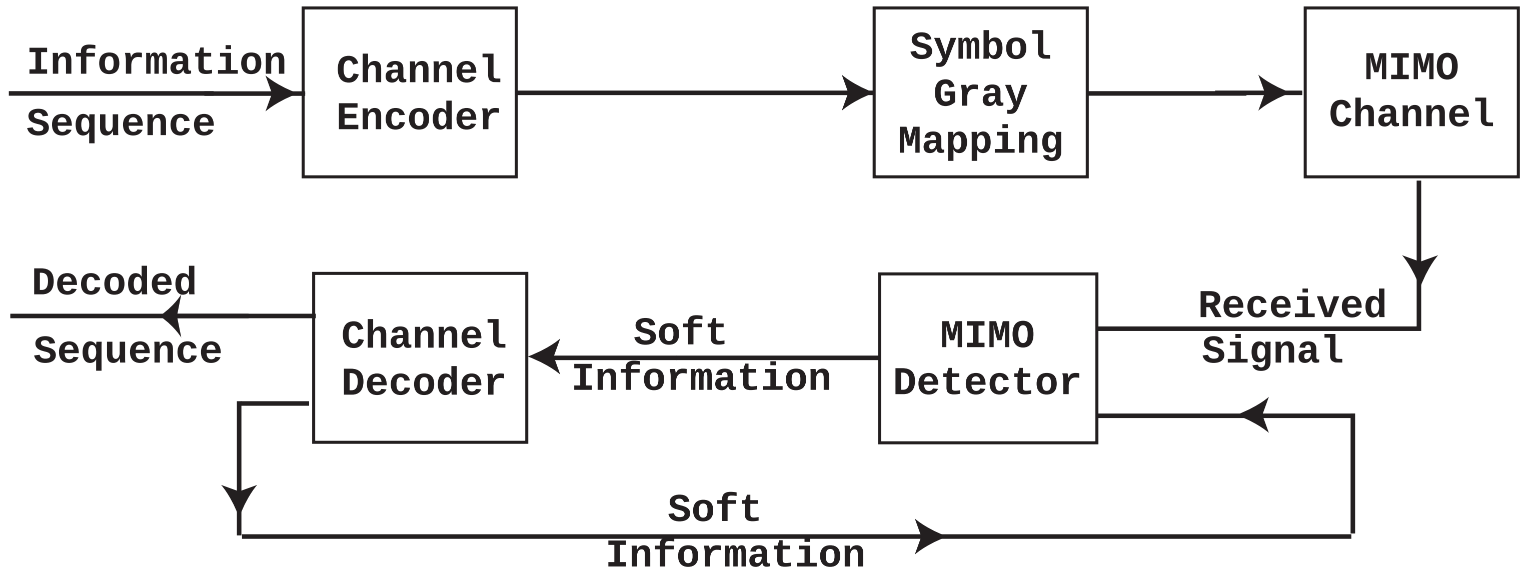
Background – MCMC methods

- Markov Chain Monte Carlo (MCMC) is a statistical method to generate random samples from arbitrary distributions.
- Previous work on MCMC methods for signal processing and communication (Doucet-Wang'05):
  - ♦ Bit-counting: use MCMC simulations to determine the frequency over which a bit occurs.
  - ♦ Many samples are needed.
  - ♦ Requires a burning period to allow Markov chain to converge.
- Our proposed MCMC methods:
  - ♦ Do not use bit-counting
  - ♦ No burning period is needed
  - ♦ Require very few samples even for large systems

Background– comparisons of suboptimal detectors

- Linear detectors: zero-forcing (ZF), minimum mean square error (MMSE): low complexity, limited performance
- List Sphere Decoding (LSD) detectors (Hochwald-Brink'03), Soft-in Soft-out (SISO)-LSD (Vikalo-Hassibi-Kailath'04):
  - Use tree search to find a sample set $\mathcal{B}$ containing likely samples.
  - Excellent performance at high SNR
  - Variable complexity
  - Exponential complexity at low SNR
- MCMC detectors:
  - Use Gibbs sampler to find $\mathcal{B}$
  - Constant complexity
  - Excellent performance at low SNR
  - High SNR problem

System diagram– Joint MIMO detection and channel decoding



MIMO channel model

- Channel Model:

$$\mathbf{y} = \sqrt{\frac{\rho}{t}}\mathbf{H}\mathbf{d} + \mathbf{n}, \tag{1}$$

  - : t transmit and r receive antenna
  - ρ: SNR per receive antenna.
  - $\mathbf{H}$: r by t channel fading matrix, i.i.d. complex Gaussian $\mathcal{CN}(0,1)$;
  - $\mathbf{d} = (d_1, \cdots, d_t)^T$: transmitted signal vector
  - $\mathbf{y} = (y_1, \cdots, y_r)^T$: received signal vector
  - $\mathbf{n}$: i.i.d. $\mathcal{CN}(0,1)$ distributed entries;
- Assumes coherent detection: receiver knows $\mathbf{H}$ perfectly.
- Noncoherent MCMC detector (Chen-Peng'09)

Optimal detector

- $\mathbf{d} = (d_1, \cdots, d_t)^T \leftrightarrow \mathbf{b} = (b_0, b_1, \cdots, b_{K-1})^T$, where $K = tM_c$.
- $\boldsymbol{\lambda} = (\lambda_0, \cdots, \lambda_{K-1})^T$, λ_i is LLR of the i-th bit provided by channel decoder.
- Given $\mathbf{y}$ and $\boldsymbol{\lambda}$, the LLR of bit b_k is

$$\gamma_k = \ln \frac{P(b_k = 1|\mathbf{y}, \boldsymbol{\lambda})}{P(b_k = -1|\mathbf{y}, \boldsymbol{\lambda})} = \ln \frac{\sum_{\overline{\mathbf{b}}_k} P(b_k = 1, \overline{\mathbf{b}}_k|\mathbf{y}, \boldsymbol{\lambda})}{\sum_{\overline{\mathbf{b}}_k} P(b_k = -1, \overline{\mathbf{b}}_\mathbf{k}|\mathbf{y}, \boldsymbol{\lambda})} \qquad (2)$$

where $\overline{\mathbf{b}}_k = (b_0, \cdots b_{k-1}, b_{k+1}, \cdots, b_{K-1})$; $b_j \in \{1, -1\}$.

- Optimal detector has exponential complexity 2^K.

Bitwise MCMC (b-MCMC) MIMO detector

- Gibbs sampler – an example: $t = 2$, QPSK

		TX1		TX2		
		b_0	b_1	b_2	b_3	
	$\mathbf{b}^{(0)} \rightarrow$	1	−1	1	1	random initialization
1st iteration		−1	−1	1	1	update b_0
		−1	1	1	1	update b_1
		−1	1	1	1	update b_2
	$\mathbf{b}^{(1)} \rightarrow$	−1	1	1	−1	update b_3
2nd iteration		−1	1	1	−1	update b_0
		−1	−1	1	−1	update b_1
		−1	−1	1	−1	update b_2
	$\mathbf{b}^{(2)} \rightarrow$	−1	−1	1	1	update b_3

- Run Q Gibbs sampler in parallel, with I iterations each.
- Sample set $\mathcal{B} = \{\mathbf{b}^{(1)}, \mathbf{b}^{(2)}, \cdots, \mathbf{b}^{(QI)}\}$. Remove redundant samples.

b-MCMC MIMO detector – Compute LLR

- Max-Log:

$$\mathrm{LLR}_k = \ln \frac{P(b_k = 1|\mathbf{y}, \boldsymbol{\lambda})}{P(b_k = -1|\mathbf{y}, \boldsymbol{\lambda})}$$

$$\approx \max_{\{\mathbf{b}:\ \mathbf{b}\in \mathcal{B}^k_{+1}\}} \left\{ -\left\|\mathbf{y} - \sqrt{\frac{\rho}{t}}\mathbf{Hd}(\mathbf{b})\right\|^2 + \frac{1}{2}\boldsymbol{\lambda}^T\mathbf{b}\right\} - \max_{\{\mathbf{b}:\ \mathbf{b}\in \mathcal{B}^k_{-1}\}} \left\{ -\left\|\mathbf{y} - \sqrt{\frac{\rho}{t}}\mathbf{Hd}(\mathbf{b})\right\|^2 + \frac{1}{2}\boldsymbol{\lambda}^T\mathbf{b}\right\}.$$

- Expanded set:

$\mathcal{B} =$

-1	1	1	-1
-1	-1	1	1

then the expanded set for bit b_0 is

$\mathcal{B}^0 =$

−1	1	1	-1
1	1	1	-1
−1	-1	1	1
1	-1	1	1

$= \mathcal{B}^0_{-1} \cup \mathcal{B}^0_{+1}$

b-MCMC MIMO detector – Compute LLR

- Log-MAP-table-tb (Chen-Peng-Ashikhmin-Farhang'08):

$$\begin{aligned} \mathrm{LLR}_k \approx & \ln \sum_{\{\mathbf{b}:\, \mathbf{b} \in \mathcal{B}_{+1}^k\}} \exp\left\{ -\left\| \mathbf{y} - \sqrt{\frac{\rho}{t}} \mathbf{H} \mathbf{d}(\mathbf{b}) \right\|^2 + \frac{1}{2} \lambda^T \mathbf{b} \right\} \\ & - \ln \sum_{\{\mathbf{b}:\, \mathbf{b} \in \mathcal{B}_{-1}^k\}} \exp\left\{ -\left\| \mathbf{y} - \sqrt{\frac{\rho}{t}} \mathbf{H} \mathbf{d}(\mathbf{b}) \right\|^2 + \frac{1}{2} \lambda^T \mathbf{b} \right\}. \end{aligned} \quad (3)$$

To further reduce complexity,

$$\begin{aligned} \ln(e^{\delta_1} + e^{\delta_2}) &= \max(\delta_1, \delta_2) + \ln(1 + e^{-|\delta_2 - \delta_1|}) \\ &= \max(\delta_1, \delta_2) + f_c(|\delta_1 - \delta_2|), \end{aligned} \quad (4)$$

Compared to Max-Log

- better performance
- less samples

b-MCMC MIMO detector – Simulation results

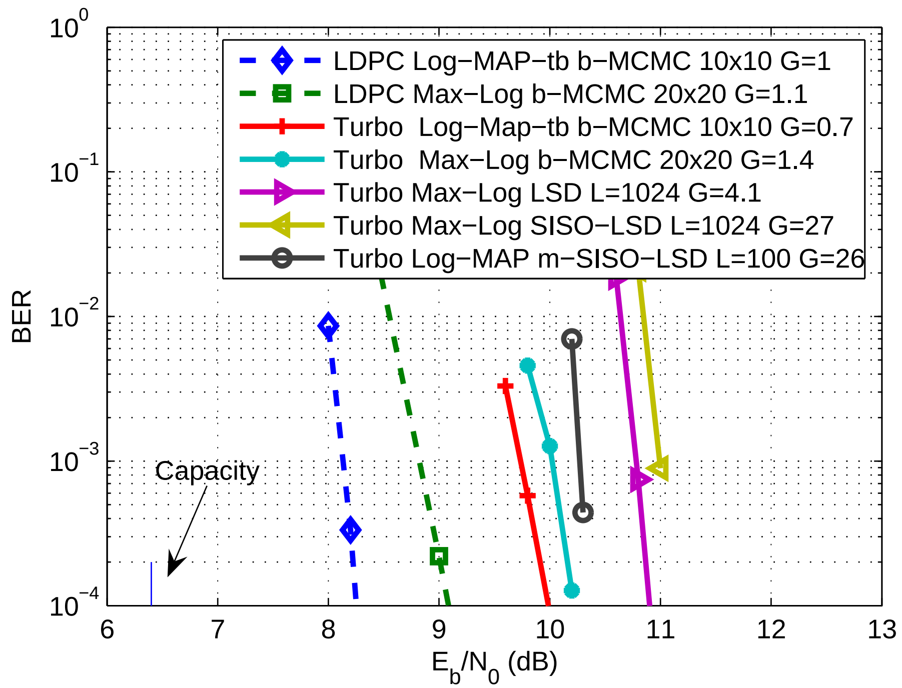


Figure 1: Performance of turbo and LDPC coded TX8 64QAM systems [Chen-Peng-Ashikhmin-Farhang'08].

- Compare Max-Log, Log-MAP-tb b-MCMC, Max-Log LSD (Hochwald-Brink'03), Max-Log-SISO LSD (Vikalo-Hassibi-Kailath'04) and their Log-MAP versions
- Log-MAP-tb 10x10 MCMC performs the best in performance and complexity
- Within 1.8 dB of capacity at 24 bits/channel use.

MCMC MIMO detectors for high SNR

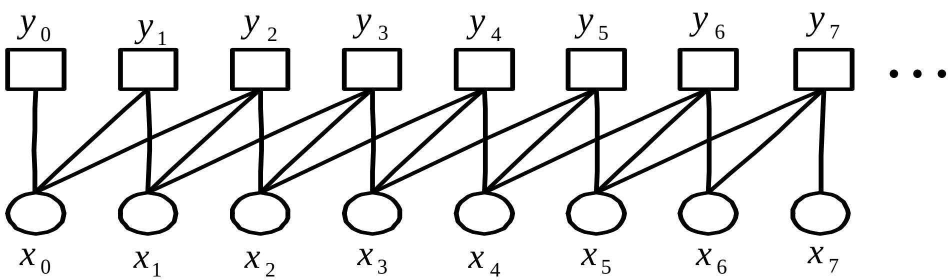
To alleviate high SNR problems:

- Assume a larger noise variance than actual noise variance (Farhang-Zhu-Shi'06).
- Initialize one of the Gibbs sampler using ZF or MMSE solutions (Mao-Amini-Farhang'07).
- Constrained MCMC (Akoum-Peng-Chen-Farhang'09):

$$\text{LLR}_k = \max_{\{\mathbf{b}:\, b_k=1\}} \left\{ -\left\| \mathbf{y} - \sqrt{\frac{\rho}{t}} \mathbf{Hd}(\mathbf{b}) \right\|^2 \right\} - \max_{\{\mathbf{b}:\, b_k=-1\}} \left\{ -\left\| \mathbf{y} - \sqrt{\frac{\rho}{t}} \mathbf{Hd}(\mathbf{b}) \right\|^2 \right\}, \quad (5)$$

  - First find ML solution $b_{\text{ML}}^{(0)}$.
  - If $b_{\text{ML},k}^{(0)} = 1$, $b_{\text{ML}}^{(0)}$ achieves the first maximum.
  - N-ML solution: the vector that attains the second maximum.
  - run constrained MCMC for each bit k to obtain good approximations of N-ML.

MCMC equalizers for frequency selective channels



Channel Model:

$$y_n = \sum_{l=0}^{L} h_l x_{n-l} + z_n, \qquad n = 0, 1, \cdots, N + L - 1, \tag{6}$$

- L: channel memory
- h_l: channel gain of l-th tap
- $\{x_0, \cdots, x_{N-1}\}$: transmitted symbols
- $\{y_0, \cdots, y_{N-1}\}$: received signals
- $\{z_n\}$: i.i.d. channel noise $CN(0, N_0/2)$.

MCMC equalizers – Gibbs sampler

Group MCMC (g-MCMC) (Peng-Chen-Farhang'09):

- Inside Gibbs sampler, update G_{max} symbols at a time.
- g-MCMC performs better than b-MCMC for channels with strong ISI.

An example: Assume $L = 2$, $G_{\text{max}} = 2$, QPSK modulation.

		x_0	x_1	x_2	x_3	x_4	x_5	x_6	x_7	
1st iteration	$\mathbf{x}^{(0)} \rightarrow$	1	3	2	2	0	3	1	0	initial vector
		3	0	2	2	0	3	1	0	update (x_0, x_1)
		3	0	1	2	0	3	1	0	update (x_2, x_3)
		3	0	1	2	2	3	1	0	update (x_4, x_5)
	$\mathbf{x}^{(1)} \rightarrow$	3	0	1	2	2	3	1	3	update (x_6, x_7)
2nd iteration		0	0	1	2	2	3	1	3	update (x_0)
		0	1	1	2	2	3	1	3	update (x_1, x_2)
		0	1	1	0	3	3	1	3	update (x_3, x_4)
		0	1	1	0	3	2	0	3	update (x_5, x_6)
	$\mathbf{x}^{(2)} \rightarrow$	0	1	1	0	3	2	0	0	update (x_7)

MCMC equalizers –compute LLR

Compute LLR for bit b_k:

$$\gamma_k = \ln \frac{\sum_{\mathbf{b}_{i_1:i_2} \in \mathcal{B}^{k,+1}_{i_1:i_2}} p(\mathbf{y}_{i:i+L} | \mathbf{b}_{i_1:i_2}) \prod_{l=i_1}^{i_2} P(b_l)}{\sum_{\mathbf{b}_{i_1:i_2} \in \mathcal{B}^{k,-1}_{i_1:i_2}} p(\mathbf{y}_{i:i+L} | \mathbf{b}_{i_1:i_2}) \prod_{l=i_1}^{i_2} P(b_l)} \tag{7}$$

- Bit k is mapped to symbol x_i
- $\mathbf{y}_{i:i+L}$: received signals that are affected by b_k.
- $\mathbf{y}_{i:i+L}$ depends only on bits $\{b_l,\ i_1 = M_b(i-L) \leq l \leq M_b(i+L) = i_2\}$.

MCMC equalizers –simulation results

A channel with strong ISI

$$h_1[n] = 0.227\delta[n] + 0.46\delta[n-1] + 0.688\delta[n-2] + 0.46\delta[n-3] + 0.227\delta[n-4]$$

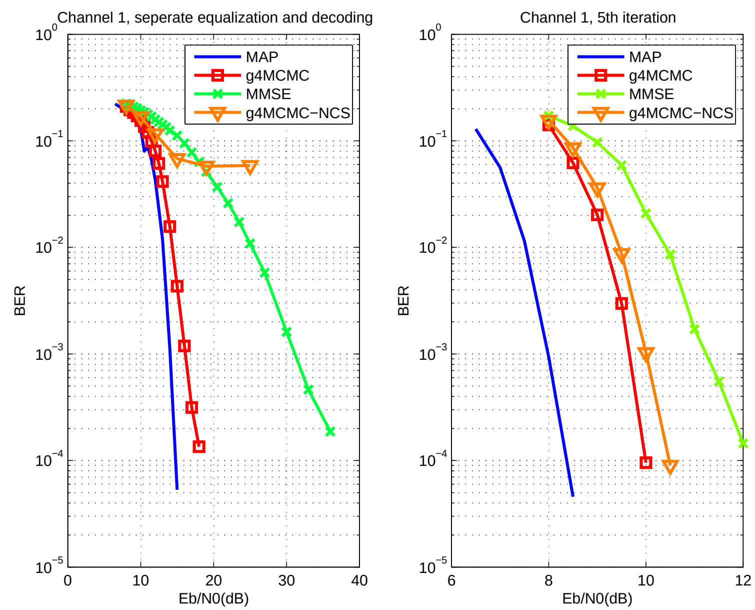


Figure 2: Performance comparisons of MAP, MMSE, g-MCMC equalizers for a strong ISI channel [Chen-Peng-Farhang'09].

- g-MCMC significantly outperforms MMSE equalizer (Tuchler-Singer-Koetter'02)
- b-MCMC does not work for such channel with strong ISI.

Conclusions and future work

- MCMC detectors are capable of achieving excellent performance at low complexity for both MIMO channels and frequency selective channels.
- Amicable for hardware implementation (Laraway-Farhang'09)
- MCMC equalizers allow for parallel implementation (Peng-Chen-Farhang'09)
- Ongoing research:
  - ♦ MCMC detectors for continuously time-varying channels and channels with imperfect channel state information.
  - ♦ Applications of MCMC equalizers to underwater acoustic channels.
  - ♦ Theoretical analysis of MCMC techniques.

Thank You !