LOW POWER SIGNAL PROCESSING STRATEGIES

Alfred O. Hero, III
Riten Gupta
Mahesh Godavarti

Nov. 16, 2000

ARO-MURI Research Progress 11/96 - 11/00

1. Reduced wordlength strategies for channel equalization and detection

   - Fixed point LMS: Transient and steady-state behavior
     (Hero&Gupta ISIT-98, SSAP-98, MILCOM-98, ICASSP-99, ICASSP-00, SP-10/2000).

   - Fixed point effects in OFDM
     (Hong, Gupta, Hero, Stark VTC-00, Gupta Rockwell-00).

   - Compression of multidimensional data for detection
     (Gupta&Hero URSI-00, ISIT-01, Gupta Thesis-01)

2. Partial update strategies for equalization and beamforming
   - Stability analysis of S-LMS for stationary signals (Godavarti&Hero, ICASSP-99)

   - Stability analysis of S-LMS and P-LMS for cyclostationary signals (Godavarti&Hero ICASSP-01, SPL-01).

   - Stochastic PU-LMS introduced with better convergence properties (Godavarti&Hero SAM-00).
3. Space-time processing and space-time coding
   - Optimal power allocation for multiple transmit antennas (Hero&Marzetta IT-00, Hero&Marzetta SAM-00, Hero&Marzetta:ISIT-00)

   - Channel Capacity for Rician-Rayleigh ST Channels (Godavarti, Hero, Marzetta ISIT-01)

Industry/Govt. Interactions

1. Invited seminars on low power MURI topics

   Hughes Network Systems, Lucent Technologies, MIT Lincoln Laboratory, France Telecom

2. Industrial Internships

   - Bell Labs, Lucent Technologies (Murray Hill) - Hero (1999)
   - Bell Labs, Lucent Technologies (Murray Hill) - Godavarti (2000)
   - Rockwell Inc. (LA) - Gupta (2000)

3. Government Panels and Study Groups
   DARPA - Hero (2000)

Reduced Wordlength Strategies: Motivation

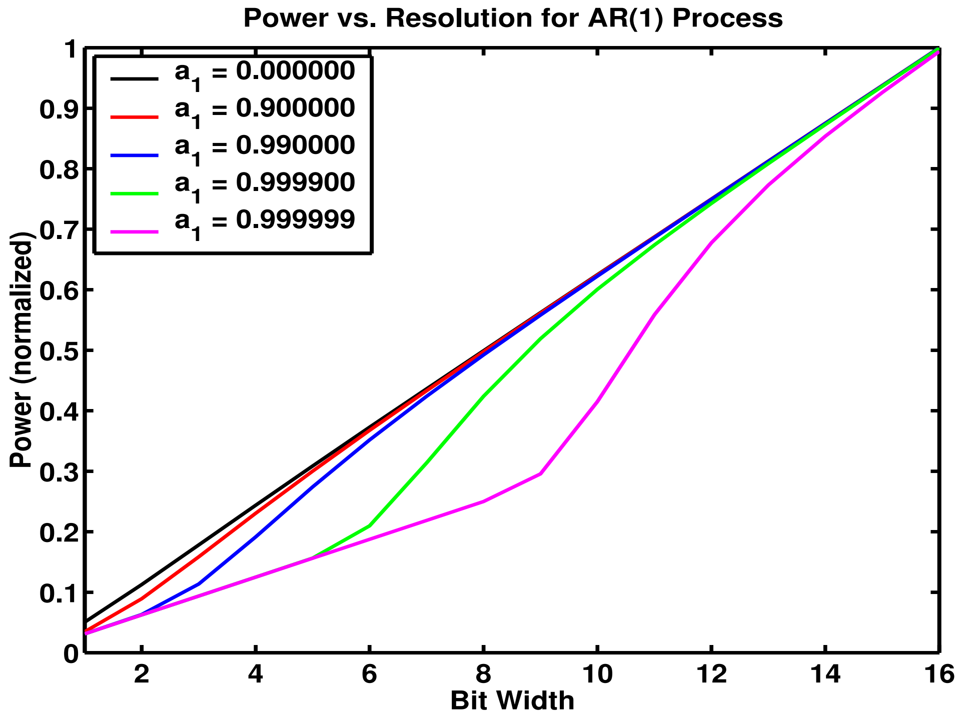
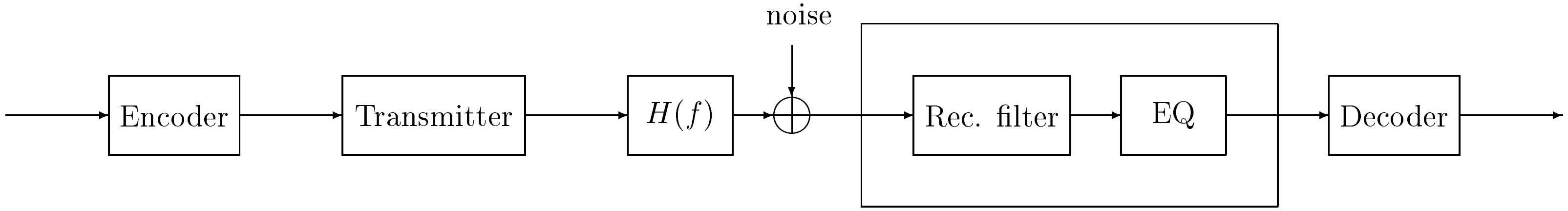


Figure 1: Normalized power vs. wordlength for Gaussian AR(1) process.

Fixed Point LMS: Motivation



Objective: Reduce power consumption of channel equalizer with minimal loss in adaptation performance.

Other applications:

- System ID (channel identification)
- Adaptive beamforming
- Adaptive interference cancellation

Innovation on previous work: Power-optimal bit allocation. (Caraiscos&Liu 84, Alexander 87, Douglas&Meng 91, Bershad&Bermudez 96)

Fixed Point LMS Adaptation

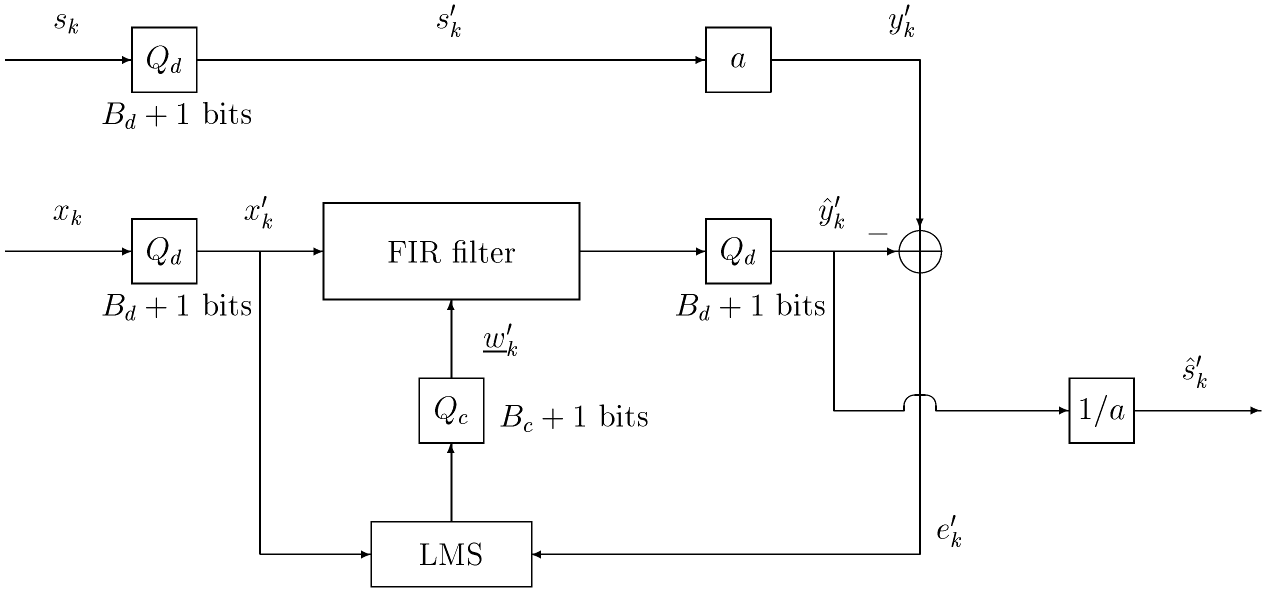
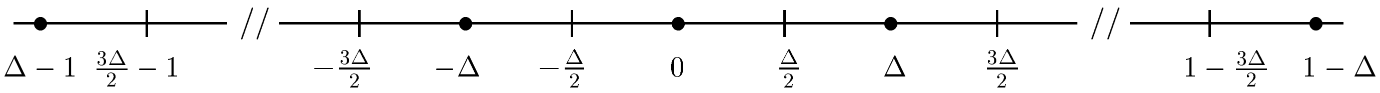


Figure 2: Finite precision LMS algorithm.



Fixed Point LMS Adaptation

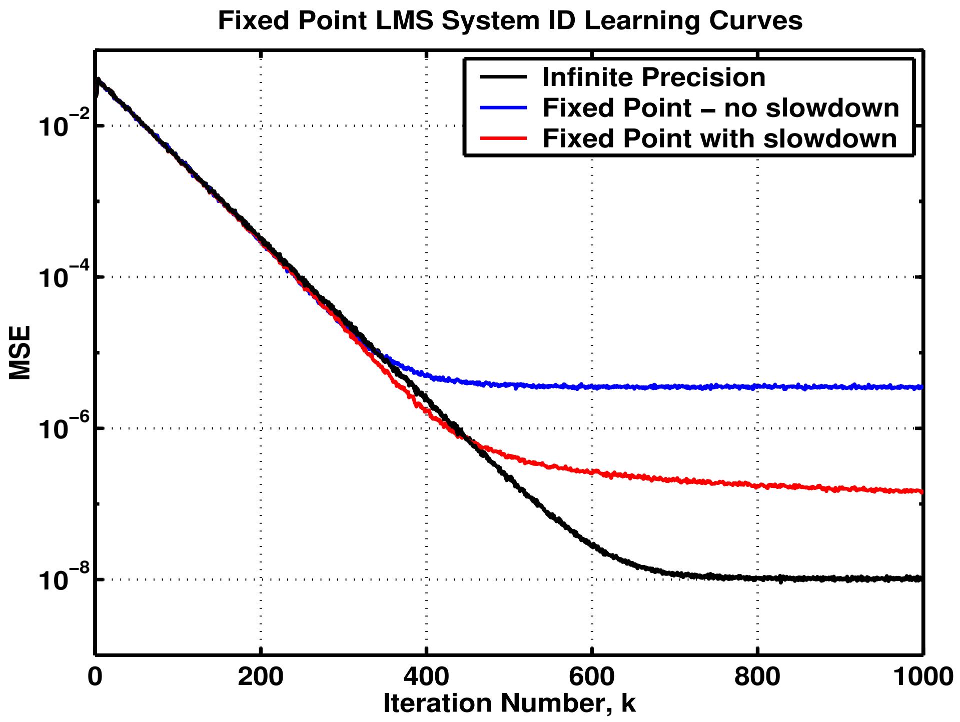


Figure 3: Fixed point LMS learning curves.

Note **"slowdown"** in red curve and **residual MSE** in blue curve.

Prediction of Slowdown: Theory

MSE floor assuming no slowdown:

$$\xi'_k > \frac{2^{-2B_d}}{6}(\|\underline{w}^o\|^2 + p + 1) + \xi_{min} \triangleq \xi'_{floor}$$

Slowdown MSE:

$$\xi'_{slow} = \frac{2^{-2(B_c+1-q)}}{\sigma_x^2[\mathrm{erf}^{-1}(1-\epsilon)]^2}, \quad q = -\log_2 \mu, \quad \epsilon \ll 1$$

No-Slowdown condition:

$$\xi'_{slow} \; < \; \xi'_{floor}$$

$$\Rightarrow B_c \; > \; B_d + \nu(\mu, \epsilon, \|\underline{w}^o\|^2, \sigma_x^2, p)$$

Prediction of Slowdown: Validation

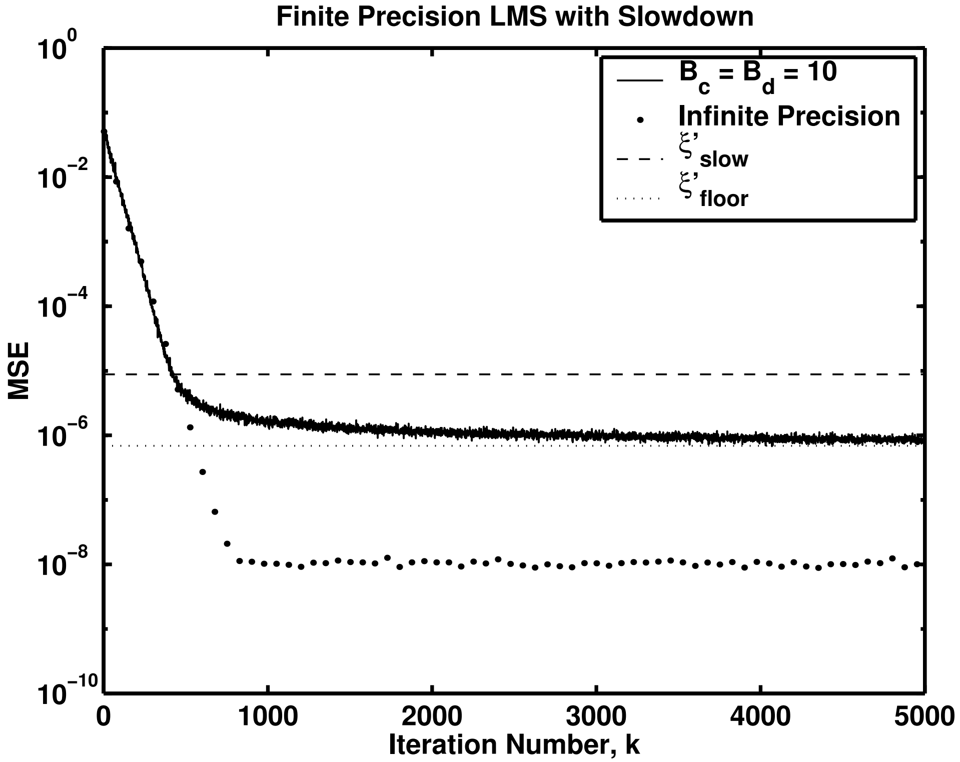


Figure 4: Fixed point LMS learning curves.

Prediction of Finite Resolution Steady-State MSE

Assuming no slowdown, increase in steady state MSE due to finite precision of data and coefficients is:

$$\xi_q' \approx \left[\frac{p}{12\mu}\right] 2^{-2B_c} + \left[\frac{\|\underline{w}^o\|^2 + p + 1}{6}\right] 2^{-2B_d}.$$

Power Optimization of Finite Resolution MSE

Finite resolution LMS power consumption:

$$\begin{aligned} P_T &= 4p[(2B_d + B_c + 2)\eta_a + (3B_d + B_c)\eta_t] \\ \eta_a, \eta_t &= \text{power per add, table lookup} \end{aligned}$$

Bit allocation factor ρ:

$$\rho = \frac{B_d}{B_d + B_c}$$

Can **optimize** ρ with respect to steady-state MSE with **power constraint**, assuming no slowdown.

$$\rho^P = \min\{\rho_{slow}, \rho^{**}\}$$

ρ_{slow} = maximum ρ such that slowdown does not occur.
ρ^{**} = optimal ρ with respect to MSE.

Power Optimization: Example

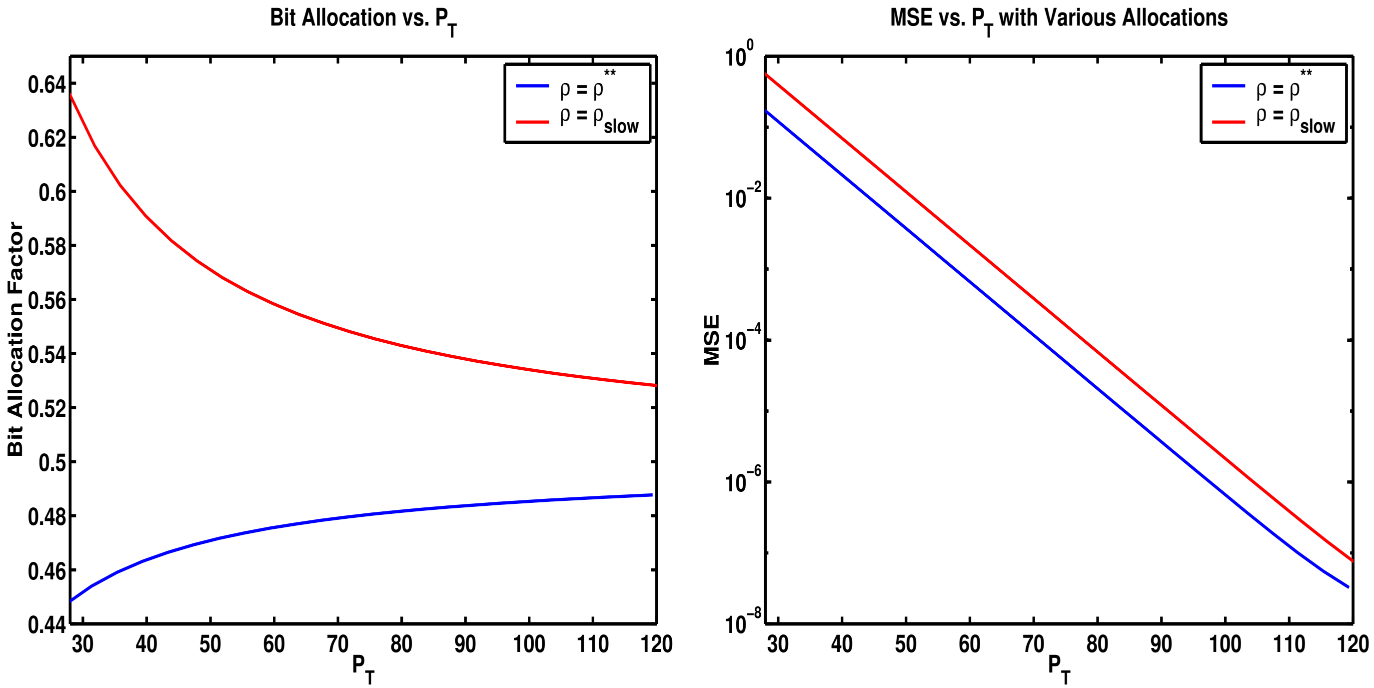
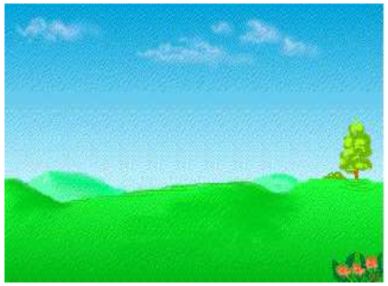


Figure 5: Optimal bit allocation and resultant MSE for LMS channel equalization example.

Future Work on Reduced Resolution Strategies

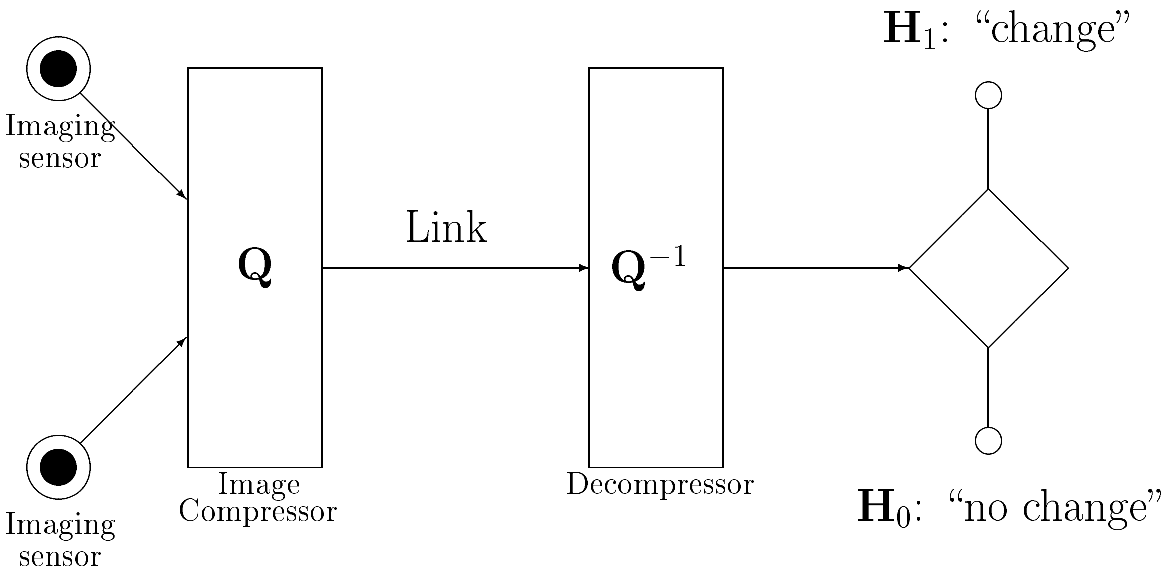
- Analyze effects of register overflow on LMS performance; incorporate power estimation / AGC into overall analysis.
- Investigate "dynamic allocation" strategies, i.e. ρ different for transient and steady-state regions.
- Study slowdown "probability tolerance", ϵ.
- Extend LMS analysis to low-power blind equalization / CMA.
- Use similar analysis to study fixed point effects for low-power OFDM.

Compression for Detection: Motivation



Old scene: x_{t-1}

New scene: x_t



Question: Is there a change in scene ($\mathbf{H}_1$) or not ($\mathbf{H}_0$)?

$$\mathbf{H}_0 : \text{joint pdf} = q_0(x_{t,1}, x_{t,2}), \qquad \mathbf{H}_1 : \text{joint pdf} = q_1(x_{t,1}, x_{t,2})$$

Objective: What is best way to compress images sequence to minimize loss in detection performance?

Other applications:

- Target recognition
- Classification
- Indexing databases
- Medical imaging

Innovation on previous work: High-rate, detection-optimal vector quantization.

- VQ: Oehler, Gray 96; Perlmutter et. al. 96.
- Scalar Quantization: Poor, Thomas 77, 83; Benitz, Bucklew 89; Picinbono, Duvaut 88.
- MSE-opt VQ: Gersho 79; Zador 82; Na, Neuhoff 95.

Vector Quantization

2-D estimation-optimal VQ for Gaussian source (known mean and covariance):

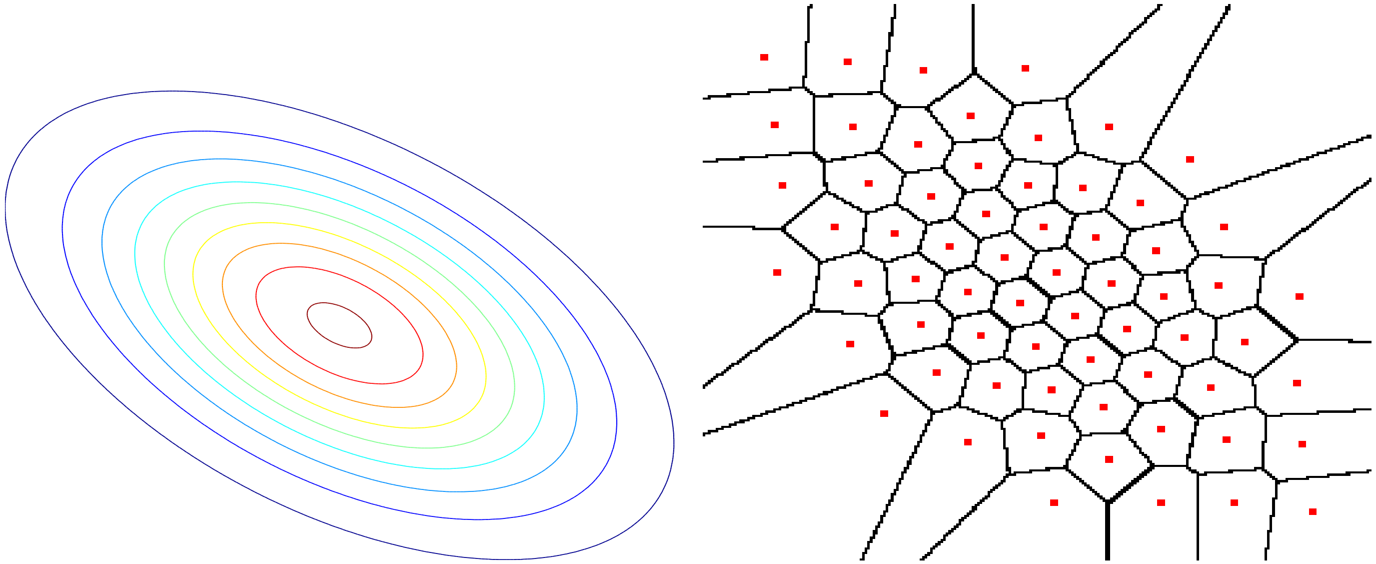


Figure 6: 2-D Gaussian source (left) and estimation-optimal VQ (right)

Compression for Detection

The optimal (Neyman-Pearson) decision rule is

$$\Lambda(x) \quad \underset{H_1}{\overset{H_0}{\gtrless}} \quad T(\lambda)$$

Error probabilities (false alarm and miss) are <u>exponentially decreasing</u> in $L(q_\lambda, q_0)$ and $L(q_\lambda, q_1)$ where

$$q_\lambda(x) = \frac{q_0(x)^{1-\lambda} q_1(x)^\lambda}{\int q_0(y)^{1-\lambda} q_1(y)^\lambda dy} \qquad \lambda \in [0, 1].$$

$L(\cdot, \cdot)$ = Kullback-Leibler divergence (discrimination)

<u>Goal</u>: Minimize loss in $L(q_\lambda, q_0)$ and $L(q_\lambda, q_1)$ due to compression.

Asymptotic Loss in Discrimination with Vector Quantization

For dense encoder cells:

$$\Delta L_0 \approx \frac{\lambda^2}{2} \int \frac{q_\lambda(x)\mathcal{F}(x)}{\zeta(x)^{2/k}} \, dx,$$

$$\Delta L_1 \approx \frac{(1-\lambda)^2}{2} \int \frac{q_\lambda(x)\mathcal{F}(x)}{\zeta(x)^{2/k}} \, dx.$$

- $\zeta(x)$ = VQ point density. Describes distribution of cell centroids.
- $\mathcal{F}(x)$ = "Fisher" covariation profile. Describes cell shapes.

Gaussian Example: Detection and Estimation VQ's

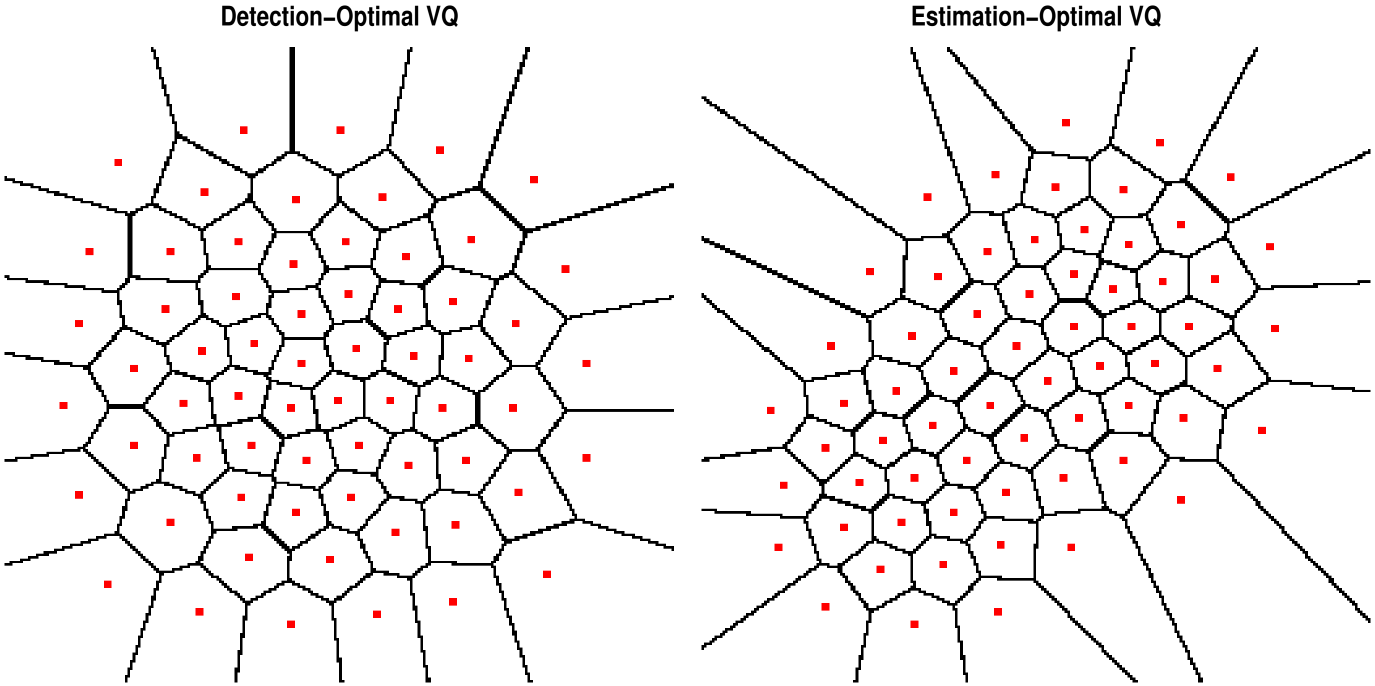
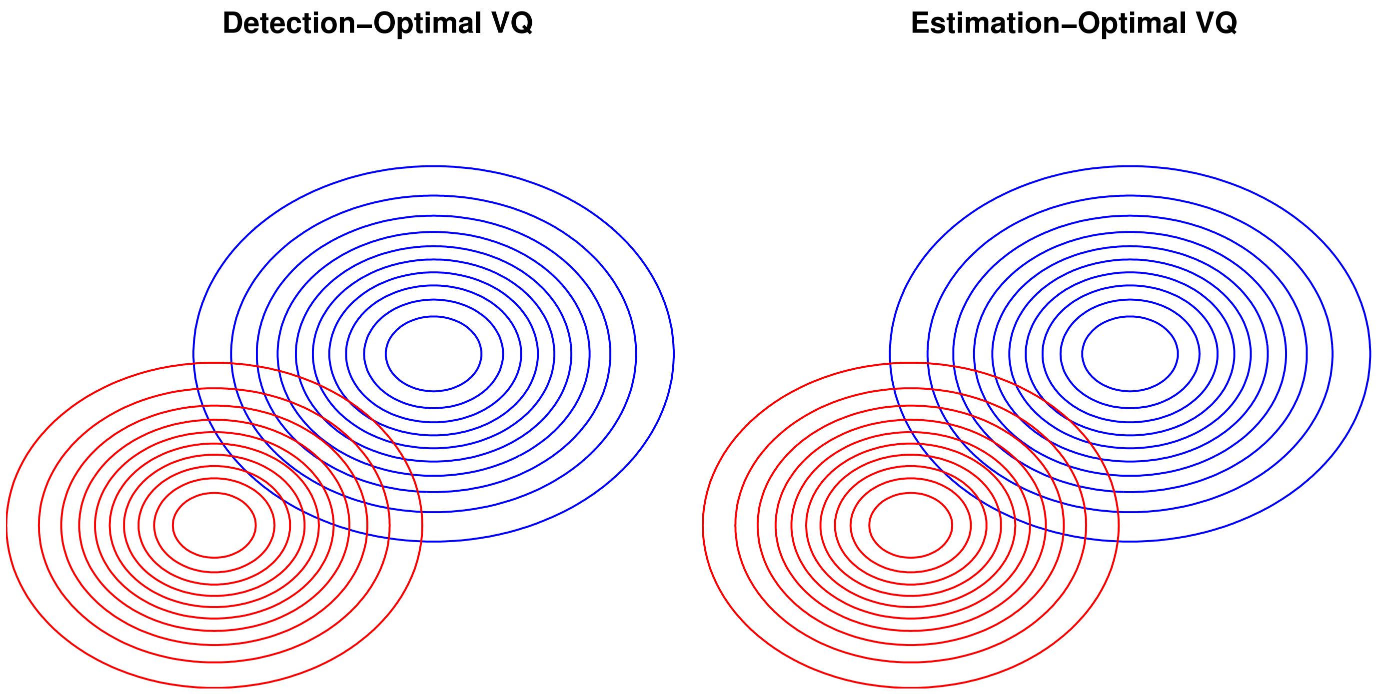


Figure 7: Detection VQ (left) and Estimation VQ (right) for min-max priors.

Gaussian Example: Detection and Estimation VQ's



Gaussian Example

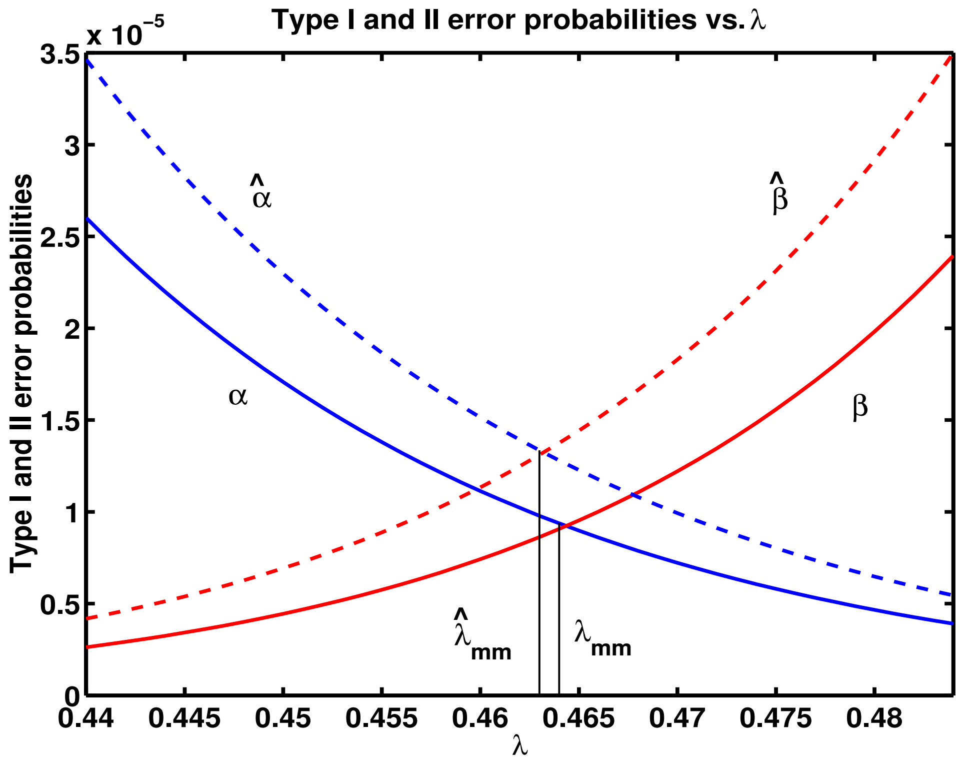


Figure 8: Type I and II error probabilities before and after quantization.

Gaussian Example: Detection and Estimation Performance

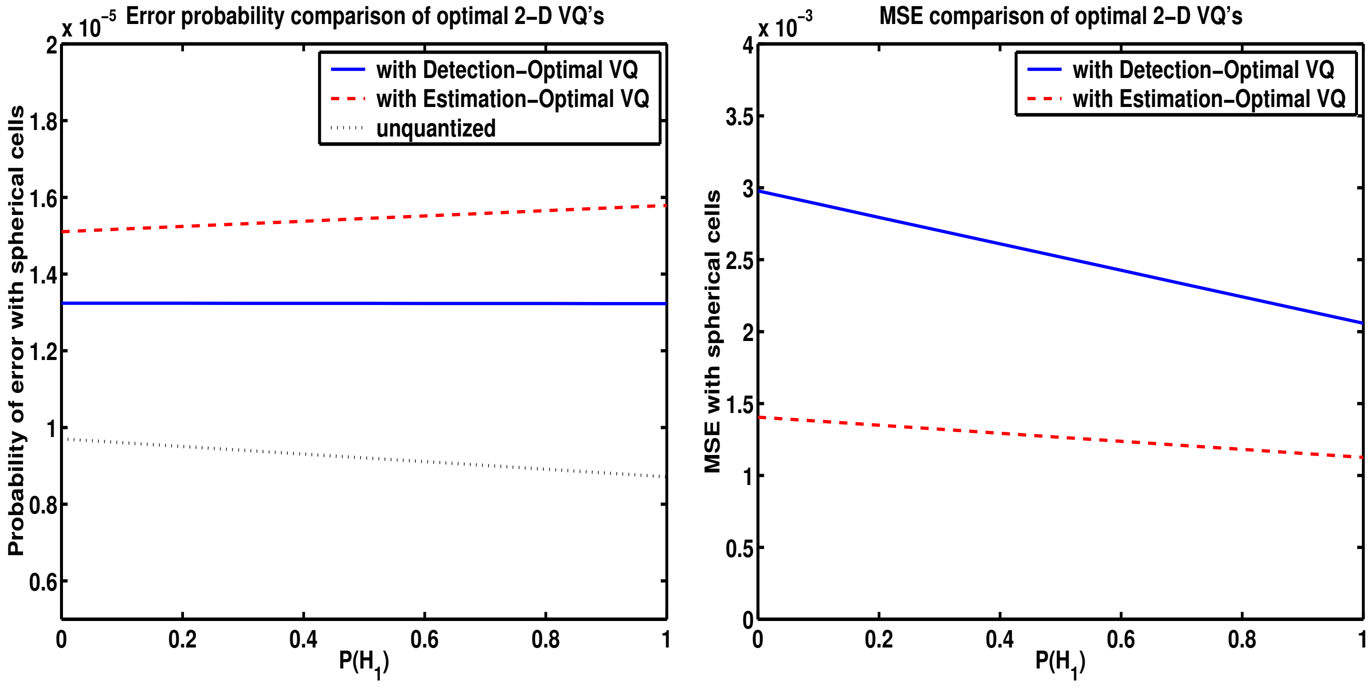


Figure 9: Detection performance (error probability) and estimation performance (MSE).

Future Work on Compression for Detection

- Compare to suboptimal structured compression strategies
- Investigate optimal compression when source densities q_0, q_1 not fully known.
- Implement compression algorithms for combined detection / reconstruction.
- Application: content-based retrieval.
- Application: succesive refinement for feature detection.

Partial Update LMS Algorithm

- Partial Update LMS Algorithm: Only a subset of LMS coefficients updated per iteration (Douglas, Godavarti, Messerschmitt)
- Advantages:
  - Computational and memory savings
  - Power savings
- Existing Partial Update LMS (PU-LMS) Algorithms
  - Periodic (**P-LMS**) and Sequential (**S-LMS**) Partial Update LMS
- Contributions
  - New Stochastic PU-LMS Algorithm (**SPU-LMS**) with better convergence properties (Godavarti&Hero SAM00)
  - Convergence criteria for **P-LMS** and **S-LMS** established (Godavarti&Hero ICASSP99, ICASSP01)

Sequential Partial Update LMS Algorithm

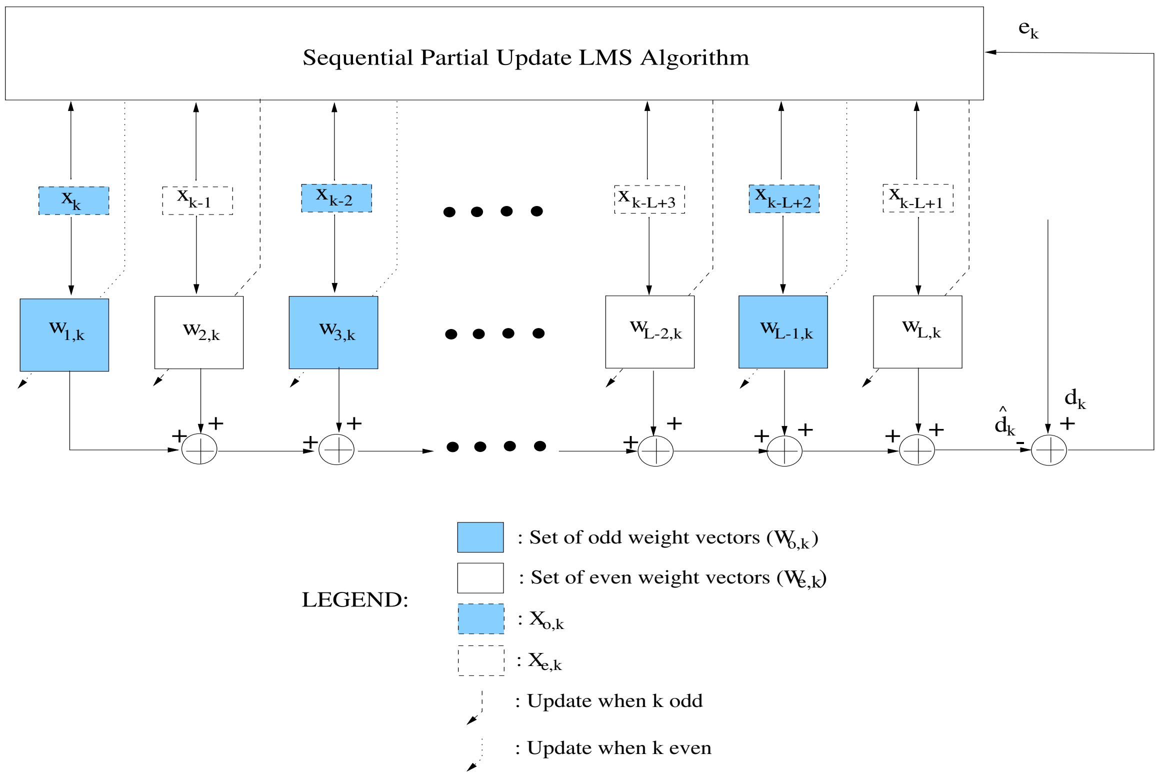


Figure 10: Block diagram of the sequential LMS algorithm.

Description of Algorithms

- Full Update LMS algorithm (**LMS**):

$\boxed{W_{k+1} = W_k + \mu e_k^* \ X_k}$

- Periodic Partial Update LMS algorithm (**P-LMS**):

$\boxed{W_{k+1} = W_k + \mu e_k^* \mathbf{1}_k \ X_k}$, $\mathbf{1}_k = 1, \text{ if } (\ k\%P = 0) \text{ and } = 0, \text{ otherwise}$

- Sequential Partial Update LMS algorithm (**S-LMS**):

$\boxed{W_{k+1} = W_k + \mu e_k^* \ \mathbf{I_k} X_k}$, $\mathbf{I_k} = \sum_{i \in S^k} \epsilon_i \epsilon_i^T, \ \epsilon_1 = [1 \ 0 \ \ldots \ 0]^T$

- Stochastic Partial Update LMS algorithm (**SPU-LMS**):

$\boxed{W_k = W_k + \mu e_k^* \ \mathbf{L_k} X_k}$, $\mathbf{L_k} \in \{I_i\}$, randomly selected

Example: Non-Stationary Signal Scenario

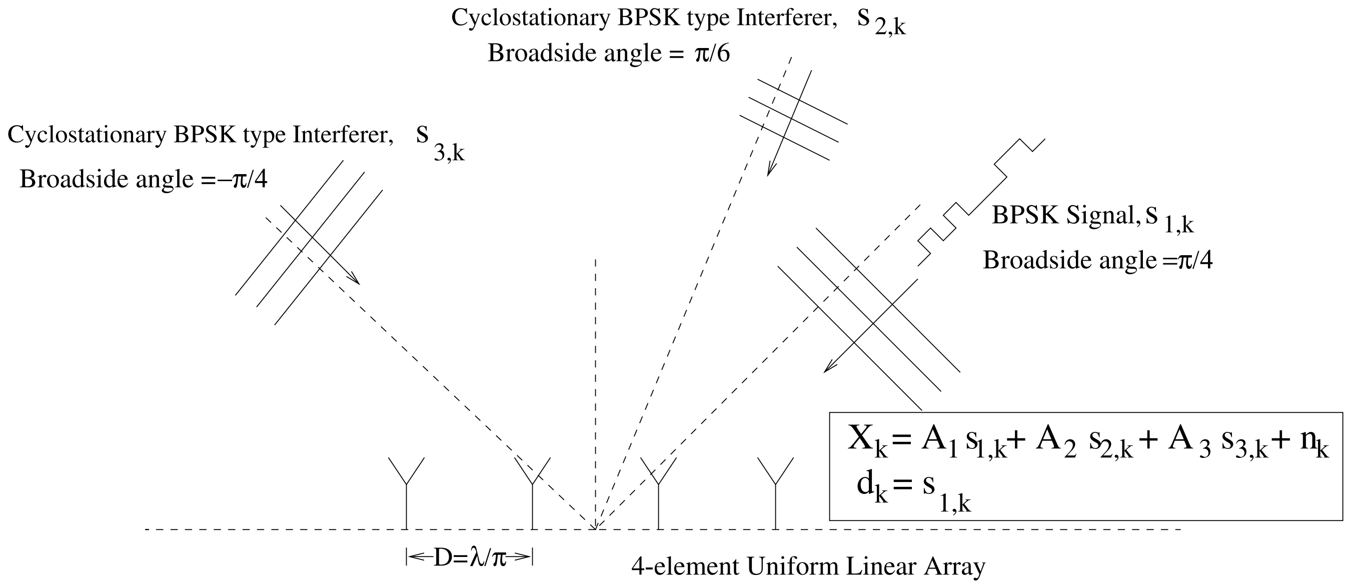


Figure 11: Signal scenario for the example

- **SPU-LMS** randomly updates one coefficient per iteration
- **P-LMS** updates every fourth instant
- **S-LMS** updates one coefficient per iteration in the order $\{1, 2, 3, 4\}$.

Example cont'd

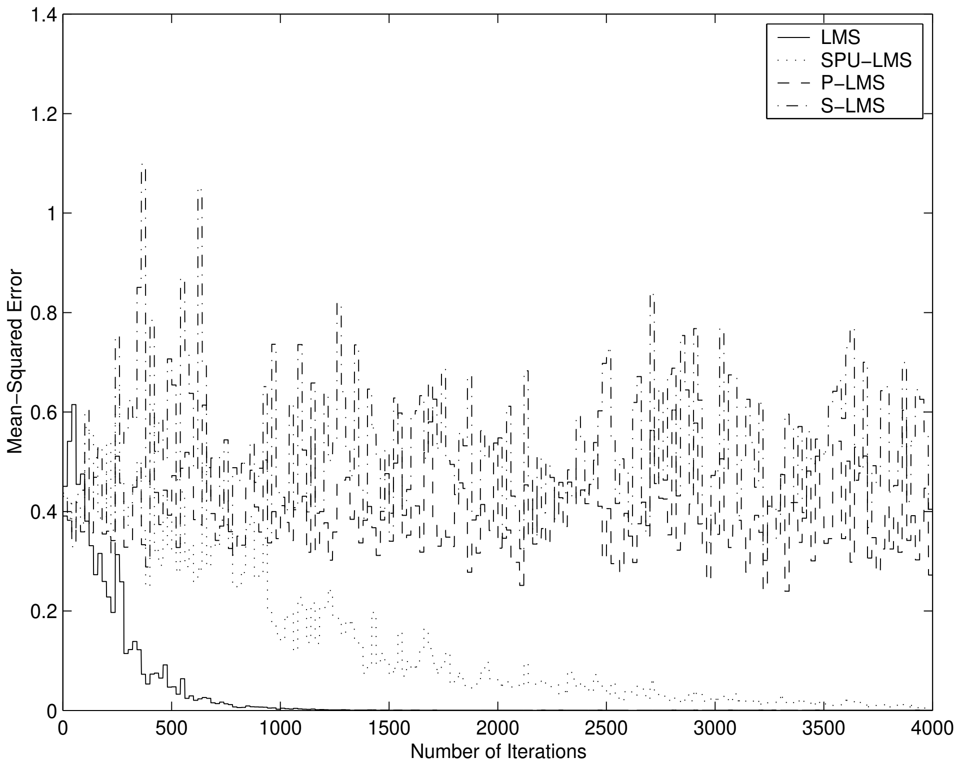


Figure 12: Envelopes of trajectories of all algorithms for Example 2. S-LMS and P-LMS failed to converge for all observed values of $\mu > 0$

Example cont'd

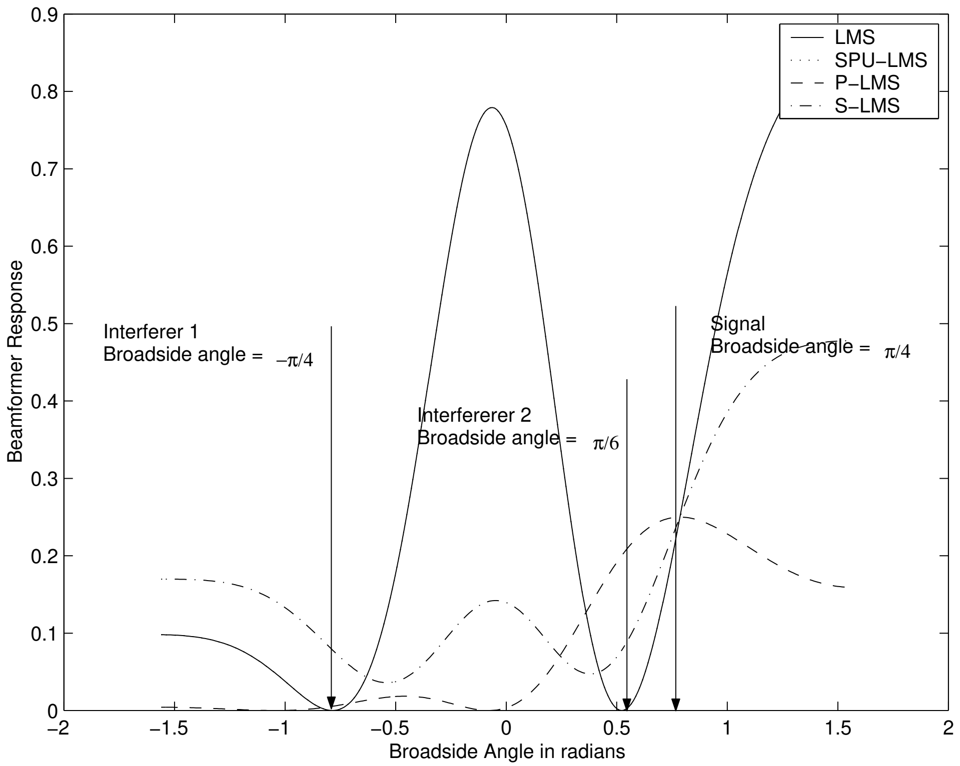


Figure 13: Best possible beampatterns generated by all algorithms. LMS and SPU-LMS have indentical beamformer response characteristics.

Space-Time Coding (STC)

Advantages:
- Combating co-channel interference and Fading
- Higher Data rates (Lower BER)
- Reduced power usage (Extension in range coverage)

Previous Work:
- BLAST (Foschini BLTJ96, WPC98)
- Channel capacity Rayleigh fading (Telatar ET 98; Marzetta&Hochwald IT99)
- Sub-optimal Space-Time codes (Seshadri IT98, SP00; Hochwald&Marzetta IT00, Hughes IT00)

Contributions:
- capacity and cutoff rate bounds
- optimal sequentially decodable STC
- Rician fading extensions

$$\boxed{X = \eta S H + N, \qquad H = (\sqrt{1-r})G + (\sqrt{rMN})\alpha\beta^{\dagger}}$$

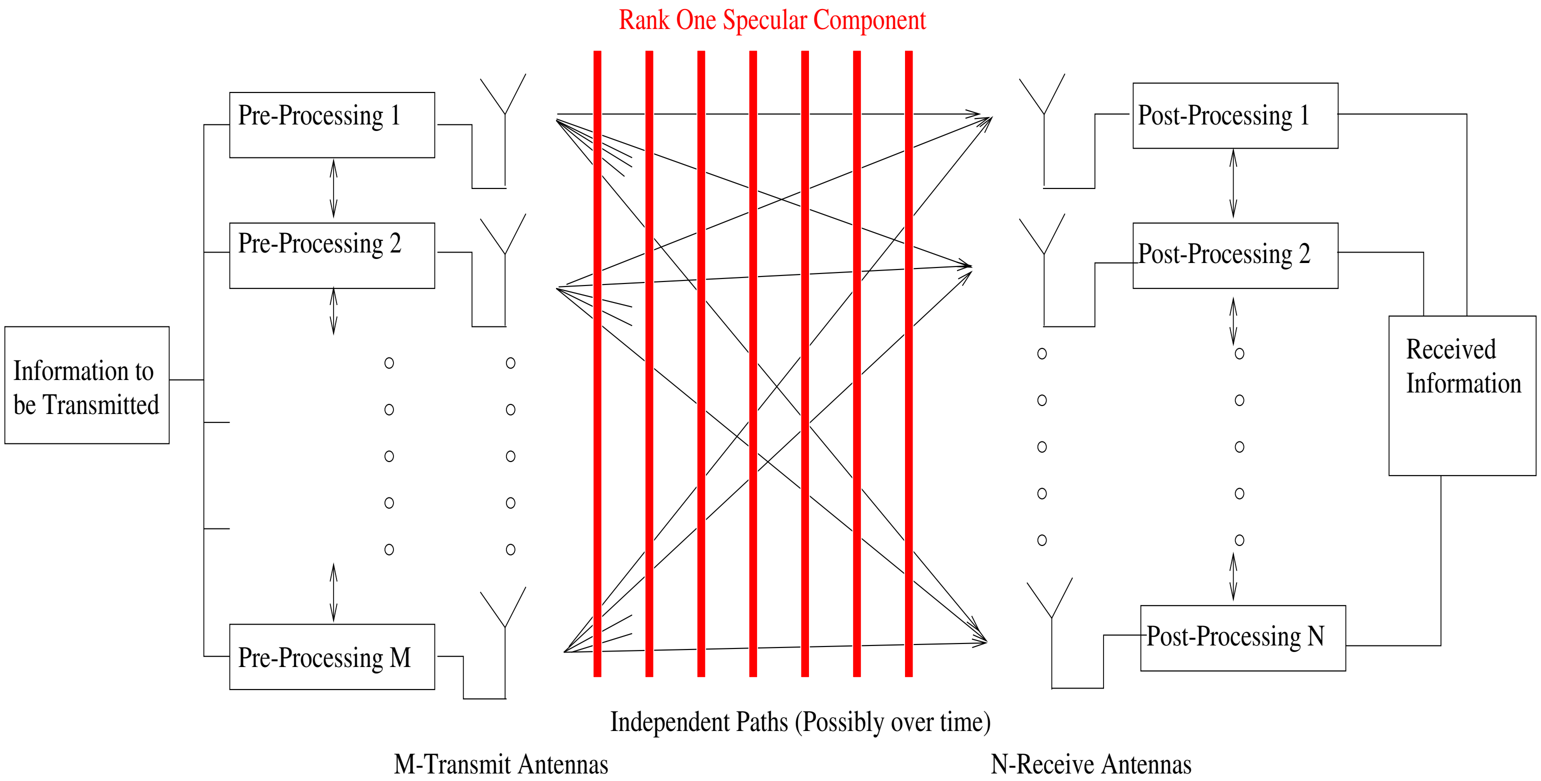


Figure 14: Multiple antenna transmitter and receiver channel. M transmitters, N receivers, T time samples.

Capacity Comparisons

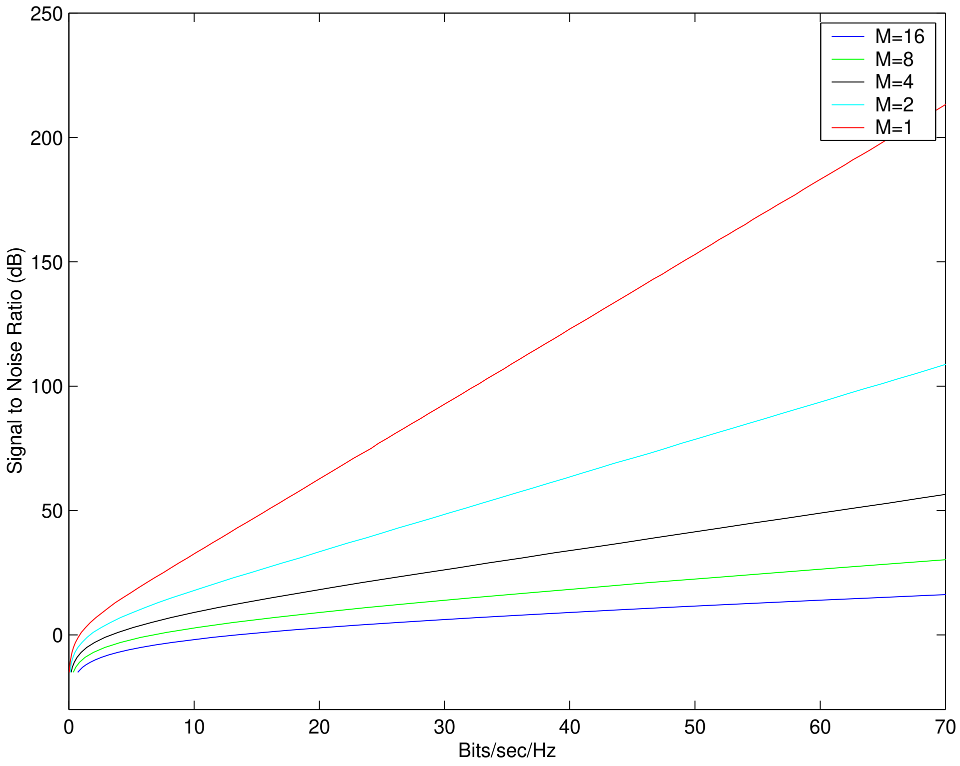


Figure 15: Capacity Upper Bounds as a function of SNR for a Rayleigh Fading Channel ($r = 0$).

Power Savings

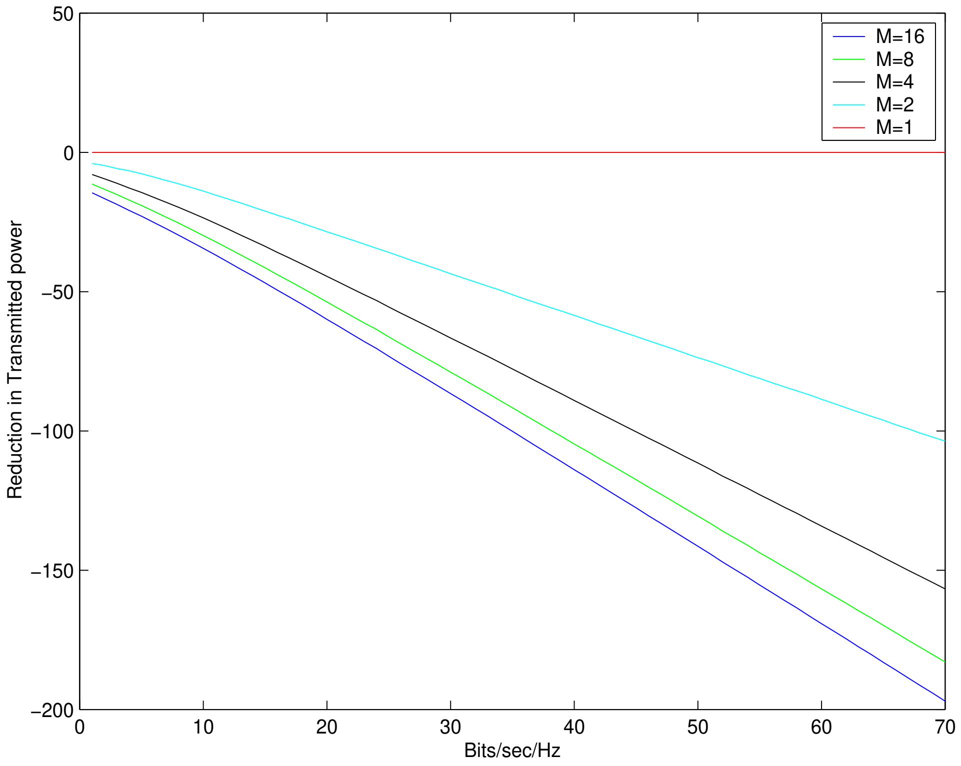


Figure 16: Reduction in Transmitted power for different number of antennas as a function of capacity (bps/hz). Channel is assumed known by receiver but not by transmitter.

Rate Optimal Constellations (Background)

$\Rightarrow$ rate-optimal constellation $\{S_k\}$ satisfies

1. H known to Txt/Rcv, $r = 0$ (Telatar:BLTM96)
$$C(H) = \log(\left|I + \eta H^H R_s^o H\right|), \;\Rightarrow\; [S]_{i*} \sim \mathcal{N}(0, R_s^o)$$

2. H unknown to Txt, $r = 0$: (Telatar:BLTM96)
$$C = E\left[\log\left|I + H^H R_s^o H\right|\right], \;\Rightarrow\; [S]_{i*} \sim \mathcal{N}(0, \rho I)$$

3. H unknown to Txt/Rcv, $r = 0$, $T \leq M$: (Marzetta&Hochwald:IT00)
$$C = E\left[\log P(X|S)/P(X)\right] \;\Rightarrow\; S = \underbrace{\Phi}_{T\times T}\underbrace{D}_{T\times M}$$
where Φ is isotropically random, D is diagonal, $\lim_{\eta\to\infty} D = \rho I$.

Rate Optimal Constellations (New Results)

1. Extension of Hochwald&Marzetta's results to Rician channels (Godavarti, Hero, Marzetta ISIT01)

2. H unknown to Txt/Rcv, $r = 0$: (Hero&Marzetta:IT00)

   $$R(K) = \ln\left(\frac{K}{1 + (K-1)e^{-ND_{\max}}}\right) \;\Rightarrow\; S_i^H S_j = \rho I\,\Delta_{ij}, \;\; \mathrm{rank}(S_i) = M_o$$

   where

   2.1. $R(K)$ is computational cutoff rate for constellation of size K

   2.2. $D_{\max}$ is "distance" between pairs in optimal constellation

   2.3. $M_o \leq M$ is number of transmit antennas used by optimal constellation

Rate-optimal Antenna Power Allocation

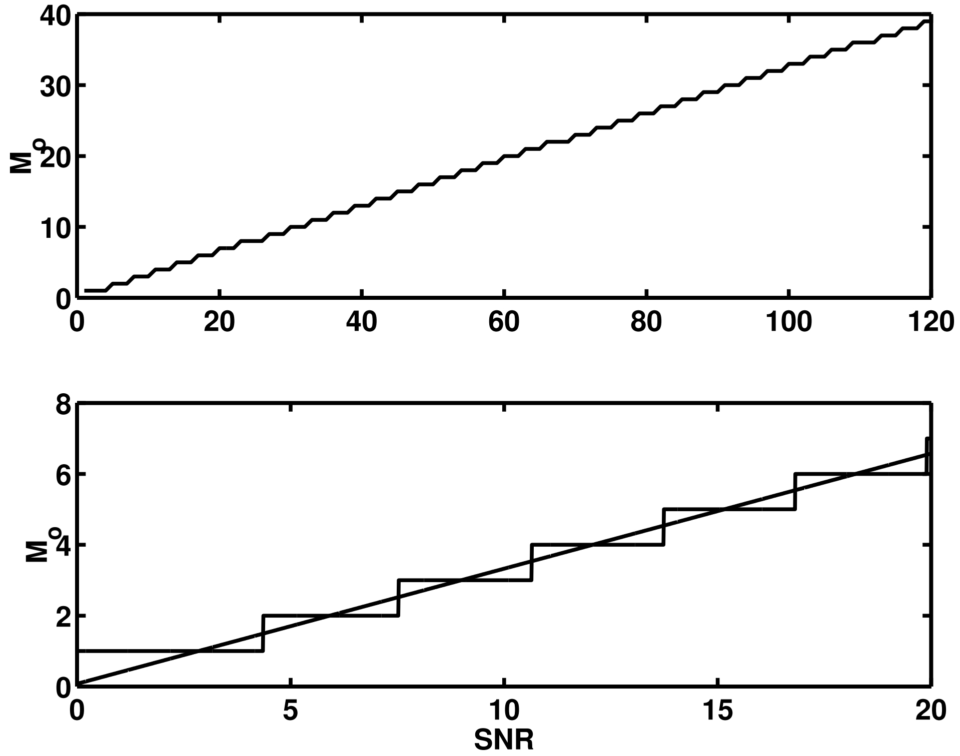


Figure 17: Number of antennas utilised by the low rate optimal constellation for various SNR's $= \eta TM$.

Future Work on Space Time Coding

- Power advantage: training vs. no training
- Power allocation: training vs. data transmission
- Sensitivity: calibration, lost nodes, training error
- Evaluation of specific space-time codes for Rician model
- Design of space-time codes for persistent specular components
- Reduced complexity optimal STC decoding strategies