PROGRESS ON LOW POWER STRATEGIES FOR ADAPTIVE FILTERING

A. O. Hero EECS, University of Michigan, Ann Arbor

Nov. 18, 1998

Adaptive equalization with training (LMS)

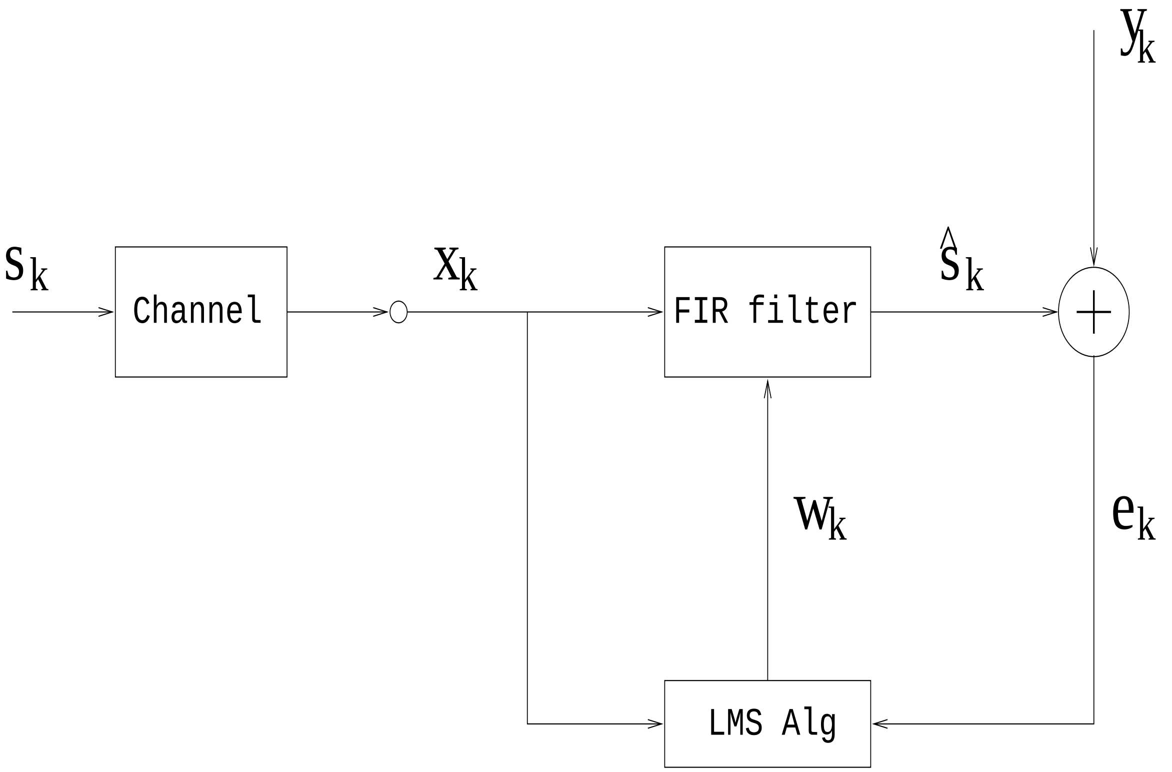

Figure 1: *LMS adaptive channel equalizer with training sequence* y_k.

Blind adaptive equalization (CMA)

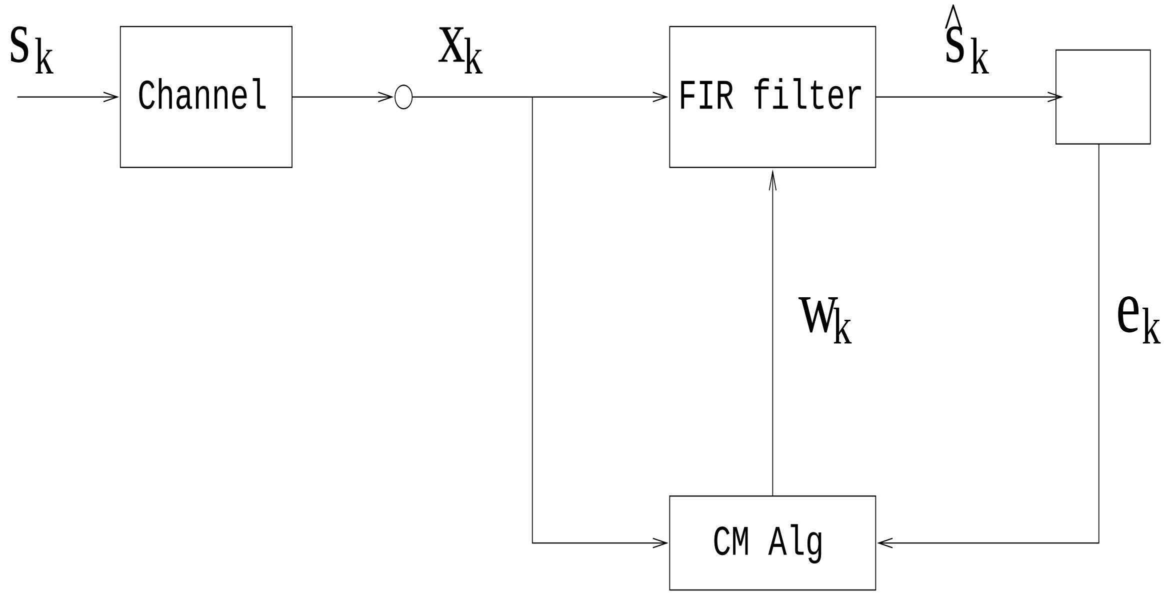

Figure 2: *Blind CMA adaptive channel equalizer.*

Relevant Applications

- adaptive channel equalization
- adaptive anti-jam and noise cancelation
- adaptive multipath combining
- adaptive nulling and beamforming arrays
- adaptive source separation
- adaptive channel identification

Low Power Adaptive SP Strategies

1. **Register length (bit width) reduction**

   1.1. Reduced resolution LMS: data and coefficient quantization

   1.2. Reduced resolution CMA: data and coefficient quantization

2. **Acceleration of filter convergence rates**

   2.1. Gauss-Seidel iterations: partial updating of LMS

   2.2. *Proximal point iterations: bundle implementation of CMA*

Students and post-docs:

Robby Gupta (G), Mahesh Godavarti (G), Hafez Hadinejad-Mehram (UG), Dr. Stephane Chretien (P)

Blind l_2 CMA Equalization

<u>CM algorithm</u>

$$\underline{W}_{k+1} = \underline{W}_k - \mu \hat{S}_k(\underline{W}_k)\ \underline{X}_k e_k$$

where

$$e_k = \left|\hat{S}_k(\underline{W}_k)\right|^2 - 1$$

Unquantized CMA Trajectory

Quantized CMA Trajectory

CMA Learning curves quantized and unquantized

Register Length Reduction

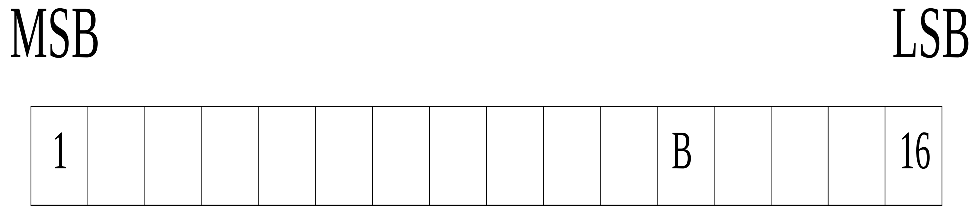

Figure 3: *Fixed point B-bit register*

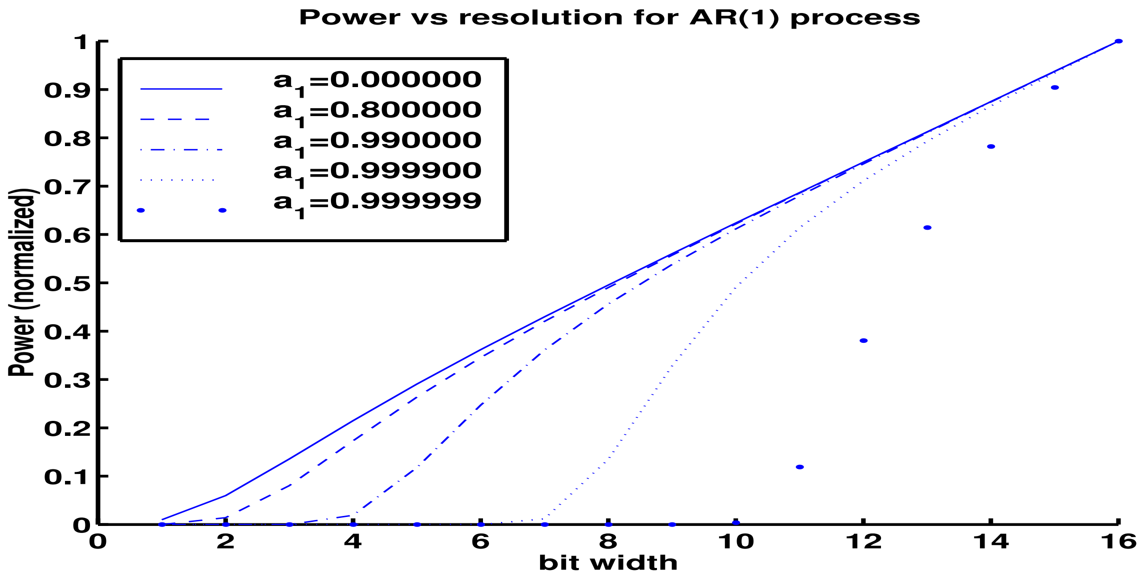

Figure 4: *Normalized power versus bit width b for various AR parameters a_1.*

Reduced Resolution LMS

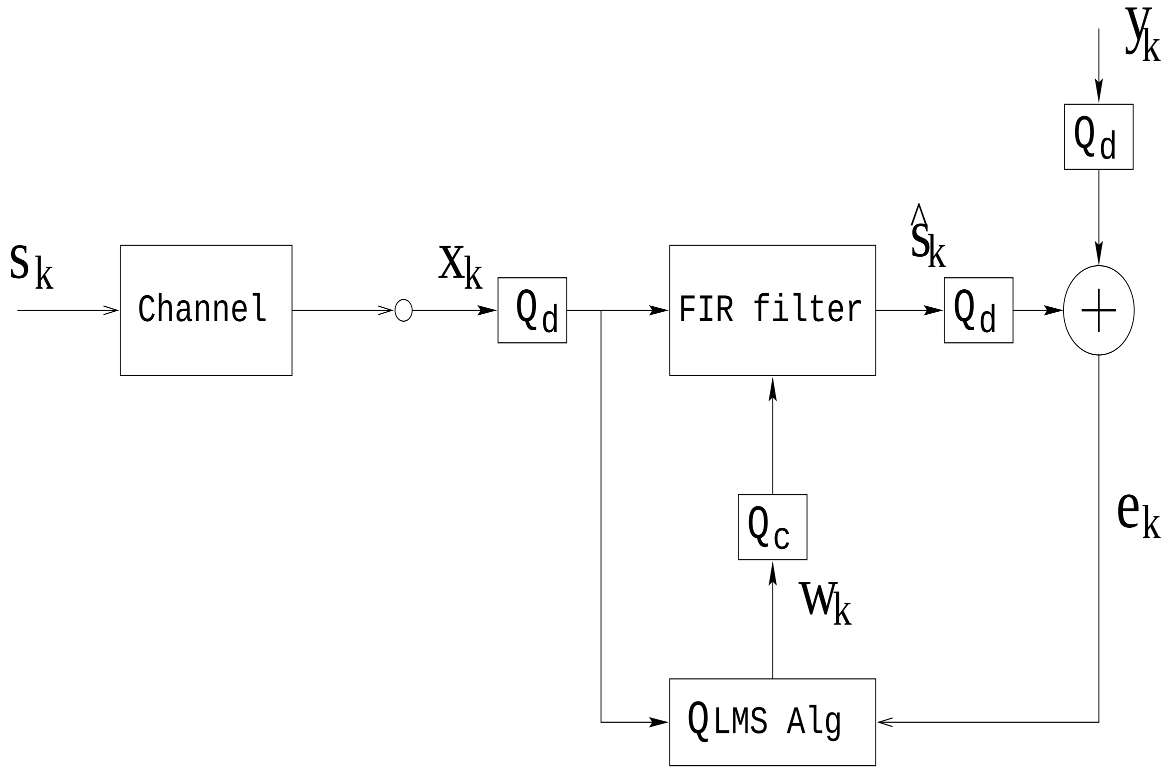

Figure 5: *LMS adaptive channel equalizer using LMS with training sequence y_k.*

Quantized LMS weight update algorithm

• $Q_d()$ quantizes to B_d data bits

• $Q_c()$ quantizes to B_c coefficient bits

Quantized (fixed point) LMS

$$\underline{W}_{k+1} = \underline{W}_k + Q_c\left(\mu \, Q_d(\underline{X}_k) e_k\right)$$

$$e_k = Q_d(y_k) - Q_d\left(\underline{X}_k^H \underline{W}_k\right)$$

Total power per iteration (lookup table mult, $\mu = 2^{-q}$)

$$P_T = \left[24p(3B_d + B_c - 2) + 32p\right] \eta_g + 24pB_d\eta_t$$

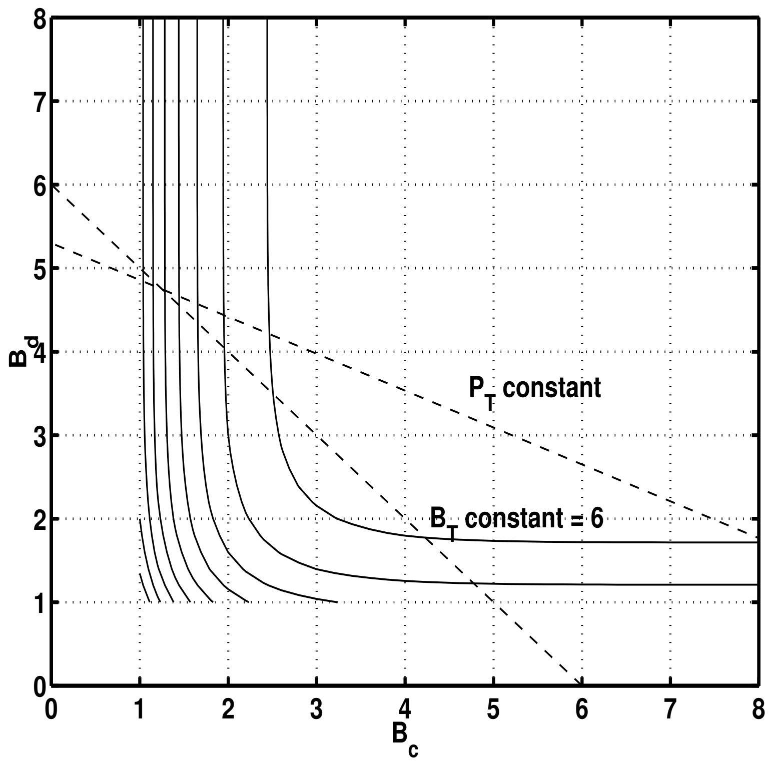

Figure 6: *Excess MSE as a function of B_d and B_c for single pole IIR channel.*

Expression for MSE Increase

$$\Delta\epsilon = \alpha_c \, 2^{-2B_c} + \alpha_d \, 2^{-2B_d}$$

where,

$$\alpha_c = \frac{A_c^2}{\mu} \, \frac{\Sigma_{i=1}^p \, q\lambda_i + r}{1 - \mu \, \Sigma_{i=1}^p \, \frac{\lambda_i}{2-\mu\lambda_i}}, \qquad \alpha_d = \frac{(\|\underline{W}^o\|^2 + p(1+\mu/2)\xi_{min})A_d^2}{6}$$

Steady state $p \times p$ weight covariance matrix

$$P = P_o + \mu(2I - \mu R_x)^{-1} \left(\alpha I + 2qR_x + r R_x^{-1}\right)$$

Optimal Data Bit Allocation Factor $\rho = B_d/B_T$

Optimal ρ for fixed bit budget $B_T = B_d + B_c$

$$\rho^* = \frac{1}{4B_T}\log_2\left(2\mu\frac{\|\underline{W}^o\|^2 + p}{p}\right) + \tfrac{1}{2}$$

Optimal ρ for fixed power budget P_T

$$\rho^{**} = \frac{a/(P_T + b) + 2}{(a + c)/(P_T + b) + 4}$$

- a, b, c functions of $\eta_g, \eta_t, p, \alpha_d/\alpha_c$

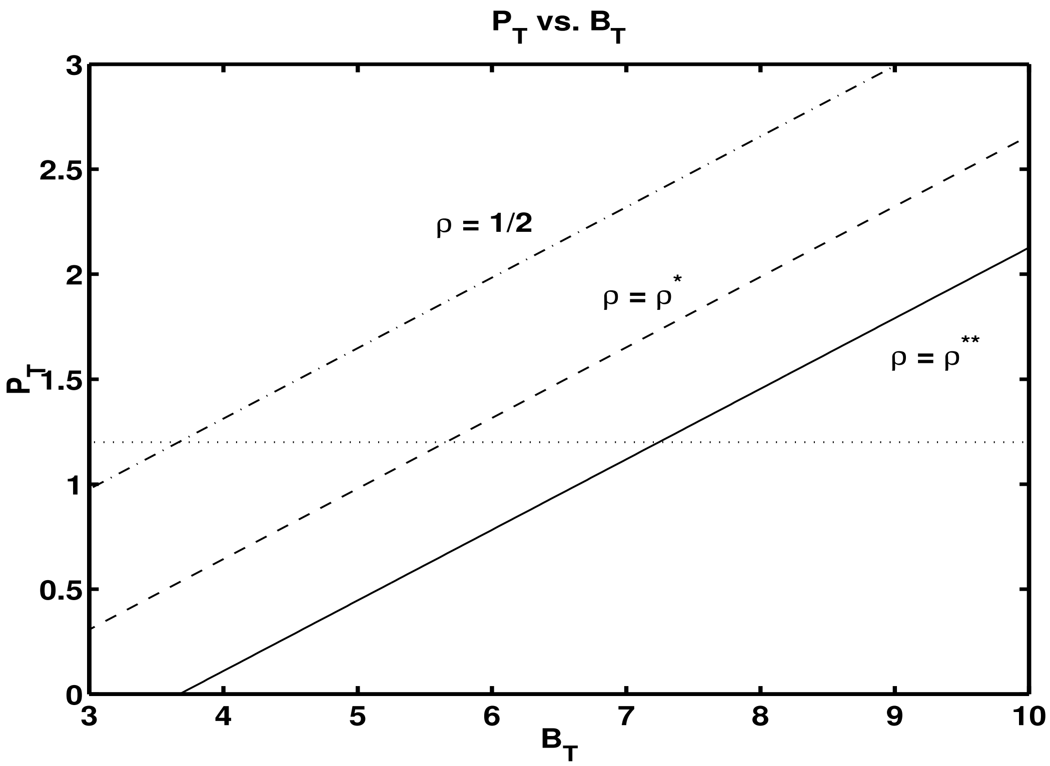

Figure 7: B_T *vs* P_T *for* $\rho = \frac{1}{2}$, ρ^*, ρ^{**}.

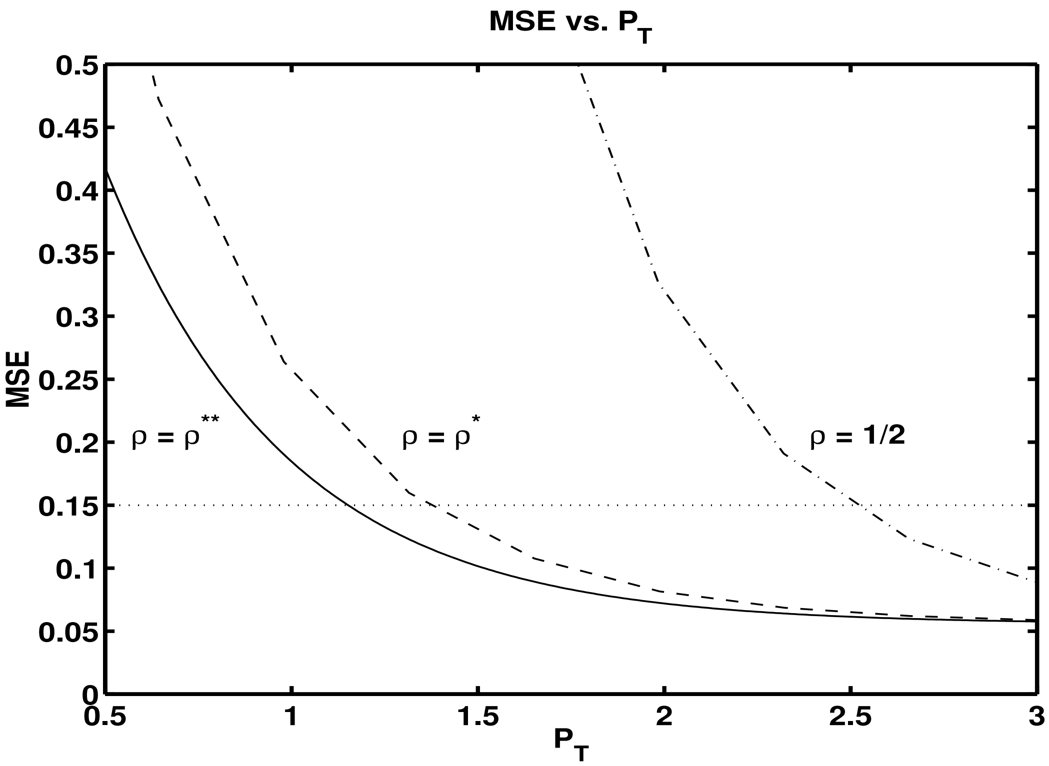

Figure 8: *IIR Channel with 4-PSK input. MSE as function of* P_T *for* $\rho = \frac{1}{2}$, ρ^*, ρ^{**}.

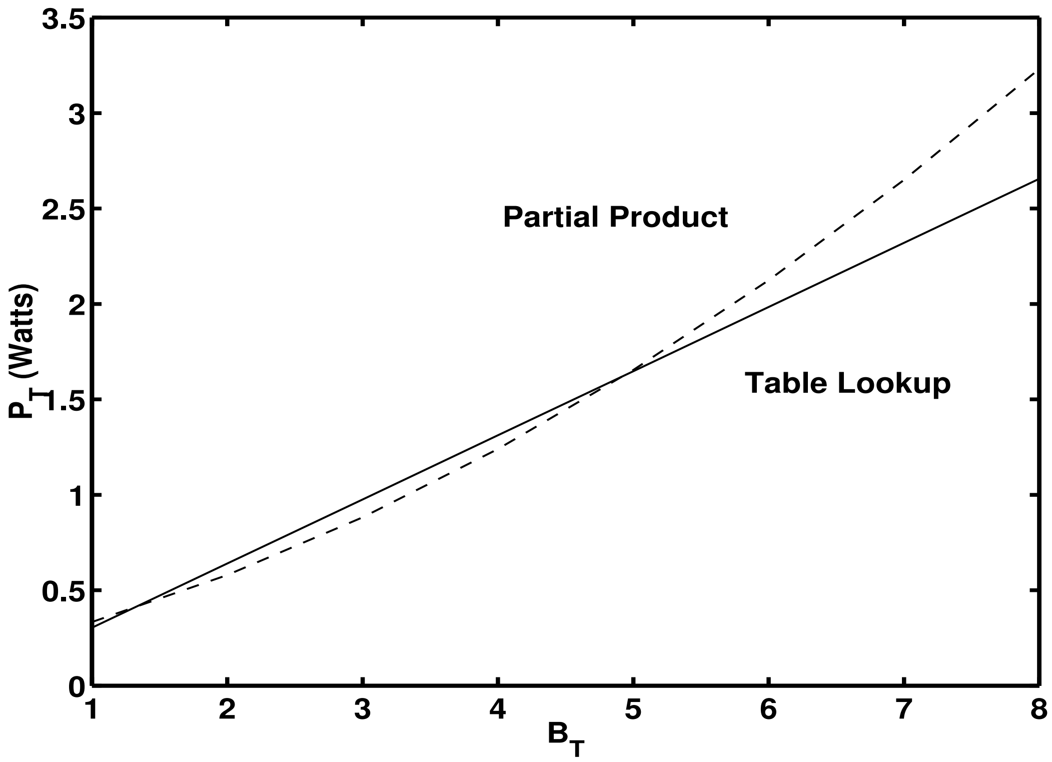

Figure 9: P_T *vs.* B_T *for multiplication using table lookup vs. shift add of partial products.*

LMS learning Curves: optimal ρ

LMS learning Curves: suboptimal ρ

Summary of Findings for Reduced Resolution LMS

- transient convergence behavior is primarily determined by B_d
- steady state behavior is determined by both B_d and B_c
- Optimal B_T and P_T bit-allocations better than "texbook" $\rho^o = 0.5$
  - Require only $B_c = 7$ and $B_d = 3$ $(B_T = 10)$ for B_T-optimal QLMS at $MSE = 0.15$
  - Require only $B_c = 7$ and $B_d = 2$ $(B_T = 9)$ for P_T-optimal QLMS at $MSE = 0.15$
  - As compared to standard 16bit architecture: $P(\rho^o)/P(\rho^{**}) \approx 4$
- $B_c = B_d$ is optimal for unlimited P_T (B_T)
- $B_c \gg B_d$ is optimal for limited P_T (B_T) and small μ

Acceleration of Filter Convergence Rates

Steepest Descent (SD) algorithm:

$$\underline{W}_{k+1} = \underline{W}_k - \mu \, \nabla J_k(\underline{W}_k)$$

LMS Algorithm is SD applied to $\lambda = 0$ version of

$$J_t^{\mathrm{LMS}}(\underline{W}) = \sum_{i=1}^{t} \lambda^{t-i} \, |e_t(\underline{W})|^2$$

l_d **CMA** is SD applied to $\lambda = 0$ version of

$$J_t^{\mathrm{CMA}}(\underline{W}) = \sum_{i=1}^{t} \lambda^{t-i} \, \left[1 - |\hat{S}_t(\underline{W})|^d\right|^d$$

Gauss-Seidel iterations

General offline coordinate descent algorithm (fixed t):

$$\underline{W}_{k+1}^{S_k} = \underline{W}_k^{S_k} - \mu F^{S_k} \nabla J_t^{\mathrm{LMS}}(\underline{W}_k^{S_k})$$

Advantages:

- Only part of weight vector is updated at each iteration
- Can obtain online partial update LMS algorithm (*Treichler, Douglass*) by setting $t = k$.

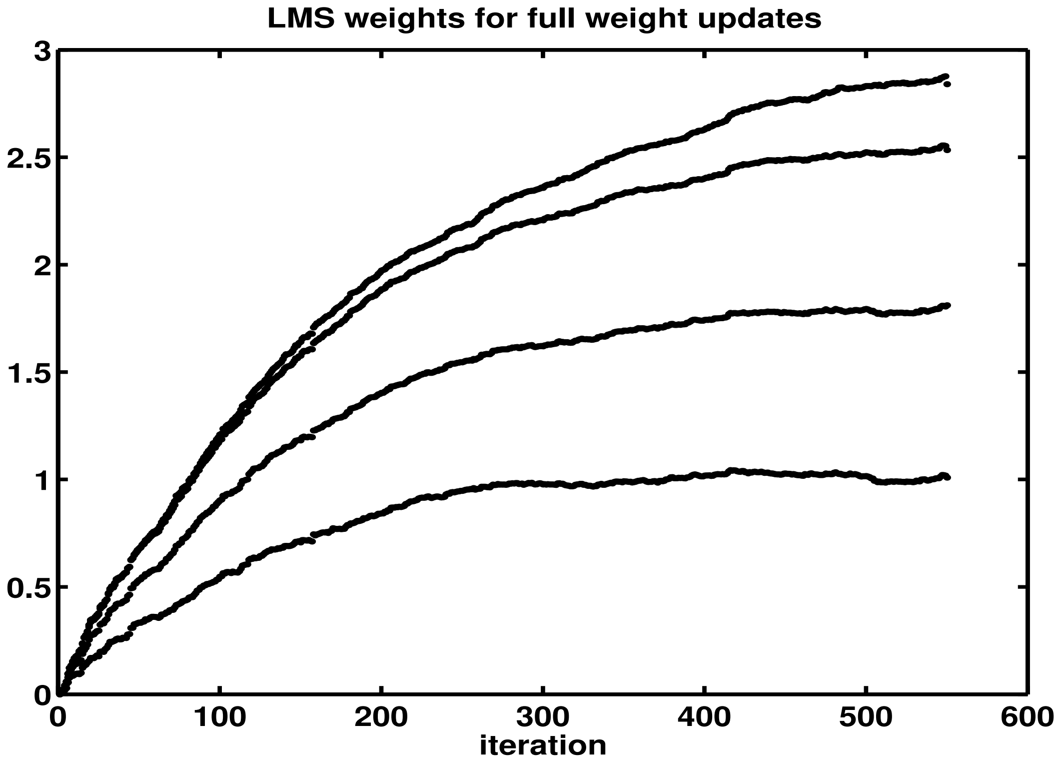

Figure 10: *LMS Weight sequence with full updating*

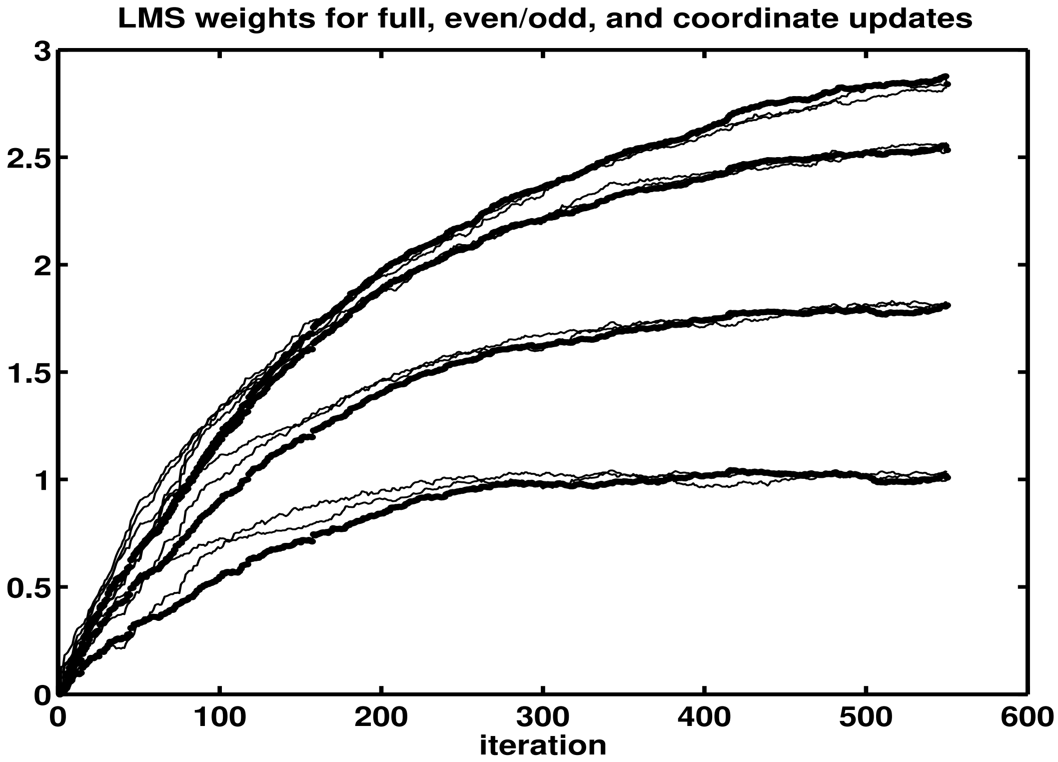

Figure 11: *LMS weight sequences with full, e/o, and coordinate updating (lighter pitch*

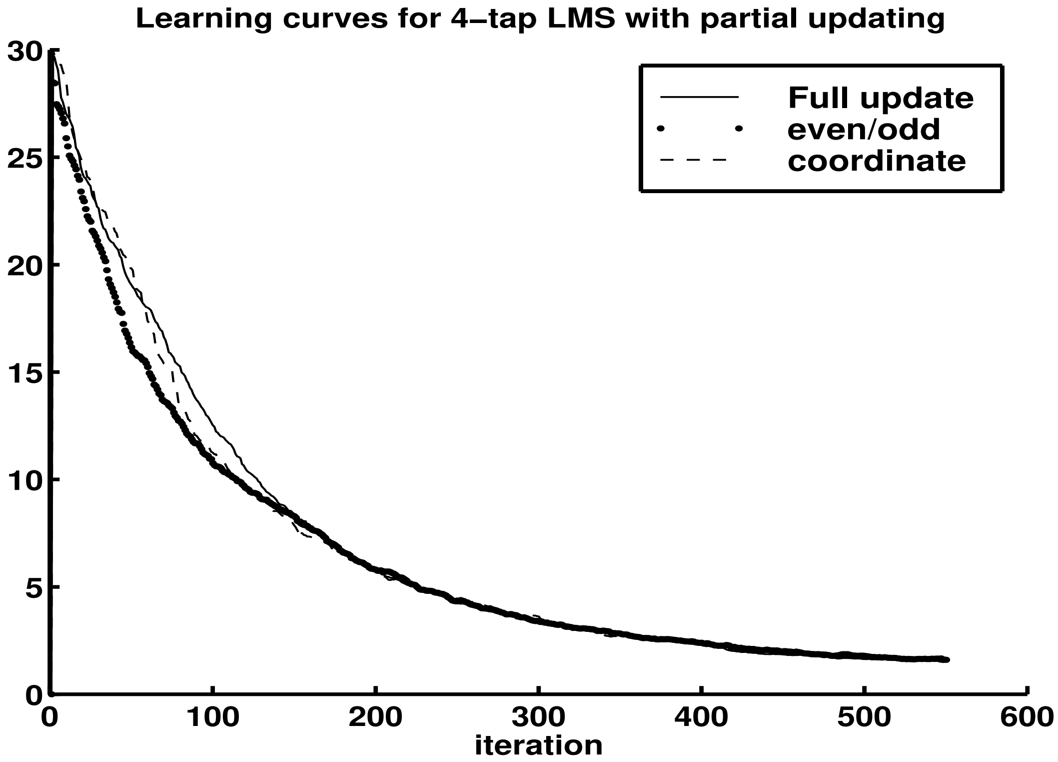

Figure 12: *MSE comparisons of LMS weight sequence with full, e/o, and coordinate updating*

Special Case: Online Even/Odd update LMS

Update for odd k:

$$\begin{bmatrix} W_{e,k+1} \\ W_{o,k+1} \end{bmatrix} = \begin{bmatrix} W_{e,k} \\ W_{o,k} \end{bmatrix} + \begin{bmatrix} \mu e_k X_{e,k} \\ 0 \end{bmatrix} \tag{1}$$

Update for even k:

$$\begin{bmatrix} W_{e,k+1} \\ W_{o,k+1} \end{bmatrix} = \begin{bmatrix} W_{e,k} \\ W_{o,k} \end{bmatrix} + \begin{bmatrix} 0 \\ \mu e_k X_{o,k} \end{bmatrix} \tag{2}$$

Ref: *Godavarti and Hero, ICASSP-99.*

Theoretical Result

As noted by Douglass, standard LMS condition on μ fails to guarantee convergence of even/odd updated LMS

$$0 < \mu < \frac{2}{\lambda_{max}}$$

$\Rightarrow$ **New sufficient condition on** μ

$$0 < \mu < \frac{2}{\lambda_{max}} \frac{\lambda_{min}}{tr(R)}$$

Next Year's Milestones

- Optimal bit allocation theory for QCMA
- Full first and second order convergence analysis for PU-LMS and NPU-LMS
- Extension of QLMS/QCMA analysis to other performance functions, e.g. P_e
- Incorporation of QLMS/QCMA/PU-LMS into system simulation software

Publications

1. “Maximum Likelihood Digital Receiver Using Coordinate Ascent and the Discrete Wavelet Transform,” I. Sharfer and A. O. Hero, to appear in **IEEE Transactions on Signal Processing**, Mar. 1999.

2. “Stability bounds on the step size for the partial update LMS algorithm,” M. Godavarti and A. O. Hero, submitted to Proc. of 1998 Int. Conf. on Acoust., Speech, and Sig. Proc., Phoenix, March 1999.

3. “Theoretical analysis of power-performance tradeoffs in reduced resolution adaptive filtering,” R. Gupta and A. O. Hero, submitted to Proc. of 1998 Int. Conf. on Acoust., Speech, and Sig. Proc., Phoenix, March 1999.

4. S. Chretien and A. Hero, “Generalized proximal point algorithms and bundle implementations,” **SIAM Journ on Control and Optimization**, vol. Submitted, Sept., 1998.

5. “Turbo codes for fading and burst channels,” J. H. Kang, W.E. Stark and A.O. Hero, Proceedings GLOBECOM, Sydney, Australia Oct. 1998.

6. “Power vs. Performance Tradeoffs for Reduced Resolution Adaptive Equalizers,” A.O. Hero and R. Gupta, IEEE Conf. on Military Communications (MILCOM), Oct. 1998.

7. ”Optimal Bit Allocation for the Quantized LMS Adaptive Algorithm,” R. Gupta and A. O. Hero, Proc. of the 1998 IEEE Workshop on Statistical Signal and Array Processing , Sept. 1998.

8. ”Data-recursive algorithms for blind channel identification in oversampled communication systems”, D. Goeckel, A. O. Hero, and W. E. Stark, **IEEE Transactions on Signal Processing**, Vol. 46, No. 8, pp. 2217-2220, Aug. 1998.

9. ”Modulation discrimination in digital communications using higher order moments”,” A. O. Hero and H. Hadinejad-Marham, Proc. of 1998 Int. Conf. on Acoust., Speech, and Sig. Proc., Seattle, May 1998.