Power vs. Performance Tradeoffs for Reduced Resolution Adaptive Equalizers

Alfred O. Hero and Robby Gupta
Department of EECS
University of Michigan
Ann Arbor, MI 48109-2122

Abstract— **Adaptive channel equalization algorithms are commonly used in wireless communications receivers to counter intersymbol interference, multi-path dispersion, and other time varying channel degradations. In this paper we obtain approximate expressions for the increase in mean square error of the LMS adaptive algorithm when the total processing power is decreased by reducing the number of data and filter coefficient bits used by the algorithm. We also obtain expressions for the power-optimal bit-allocation factor which determines the proportion of the bits allocated to the data vs. allocated to the coefficients. Numerical studies are presented for an exponential memory ISI channel and 4-ary PSK signalling. These studies indicate that as few as 8 bits total are needed to equalize the channel and that most of these bits (6 out of 8) should be allocated to the filter coefficients.**

I. INTRODUCTION

In a battery powered receiver, the adaptive equalization function consumes a significant portion of total processing power. For example the SINCGARS combat radio used by the US Army consumes on the average 7 Watts in receive mode of which more than 1 Watt is consumed by the channel equalizer [1]. Therefore the channel equalization function is a prime target for power reduction in these handsets. There have been many digital hardware design strategies proposed for power reduction including: reduction of supply voltage, reduction of clock speed and data rate, parallelization and pipelining of operations, using sign-magnitude arithmetic, and differential encoding of data [2], [3]. Another technique, which is the springboard for this paper, is the reduction of the number of bits used to represent the data and control variables in the digital circuit. This bit width reduction strategy is very highly leveraged since it reduces the power dissipation everywhere in the data and control flow paths. This strategy is also very versatile since it can be applied to any hardware architecture and can be easily adjusted in real time by dynamically switching-off bus lines and register bits. However, bit width reduction generally entails a degradation in algorithm performance, as measured by adaptive algorithm convergence rate, steady state mean square error (MSE), and subsequent probability of bit error. This paper provides an analysis of MSE degradation versus power reduction for a widespread class of adaptive algorithms popularly known as the LMS algorithm.

The LMS algorithm was introduced by Widrow [4] and is one of the most common adaptation algorithms found in practical systems such as channel equalizers [5], adaptive antenna arrays [6], and interference cancellation systems [7]. The LMS algorithm adapts the filter coefficients of an FIR filter driven by an error signal formed by subtracting the training signal from the received data. We consider a quantized version of the LMS algorithm, called QLMS, which is an LMS algorithm implemented with separate uniform scalar quantizers in the data path and the filter coefficient path, where the quantizers can have different resolutions. To obtain significant power reduction, we propose applying different pairs of quantizers during the transient acquisition phase and the steady state tracking phase of the algorithm. In this paper we focus on the steady state tracking phase.

We first present a formula for the increase in steady state mean square error (MSE) due to quantization which generalizes the formulas of Caraiscos and Liu [8] to the case of complex data and coefficients. We then derive a pair of optimal bit-allocation factors which minimize the increase in MSE subject to: 1) a total bit-width constraint; and 2) a total power consumption constraint. Finally we show that QLMS with optimal bit-allocation consumes significantly less power than LMS at little expense in performance. While these results hold for the generic LMS algorithm in a variety of applications, we concentrate on the case of channel equalization with training in this paper. Numerical examples will be given for an IIR (exponential memory) channel which illustrate that the power can be reduced by more than a factor of 4 relative to the standard LMS implemented with 16 bit arithmetic and at negligible increase in MSE. This power reduction is achieved by a QLMS algorithm having a total of 8 bits and optimal bit allocation strategy consisting of assigning 2 bits to the data and 6 bits to the coefficients.

II. REGISTER LENGTH AND POWER

It is well known that the power consumed by the operation of loading successive time samples of a random sequence into a B-bit register is proportional to the average number of bit flips induced in the register [9]. While for a white sequence the average number of bit flips is B, in

[1]This research was supported in part by the Department of Defense Research & Engineering (DDR&E) Multidisciplinary University Research Initiative (MURI) on Low Energy Electronics Design for Mobile Platforms and managed by the Army Research Office (ARO) under grant ARO DAAH04-96-1-0337. Corresponding author: Alfred Hero, hero@eecs.umich.edu, (734)-763-0564 (voice), (734)-763-8041 (fax).

general this average can be much less than B for a correlated sequence. This is because for correlated sequences higher order bits have lower probability of transitioning than lower order bits. Several models for the power consumption of register loading have been proposed [9]. We propose to use the following simple upper bound on the power consumption for a B bit register, derived under the assumption of a zero mean wide sense stationary Gaussian random sequence:

$$P_B \approx B\eta \cdot \operatorname{erfc}\left(\frac{2^{-B+1}}{\sqrt{R(0)-R(1)}}\right), \tag{1}$$

where η is the power dissipation-per-bit, which depends on factors such as load capacitance and supply voltage, and $R(\tau)$ is the autocorrelation function of the random sequence. Note that it is important that the bound (1) is conservative: if we constrain the right hand side of (1) to some maximum tolerable power dissipation, P_{max} say, then a circuit design which uses register bit width B_{max} which solves the equation $P_{max} = B\eta \cdot \operatorname{erfc}\left(\frac{2^{-B+1}}{\sqrt{R(0)-R(1)}}\right)$ is guaranteed to consume less power than P_{max}.

A plot of P_B versus B is given in Fig. 1 for an AR(1) sequence with real pole located at a_1. Note that as the pole approaches the unit circle the P_B curve displays an abrupt threshold occurring at an increasingly high bit-width. The threshold bit-width can be specified by the formula

$$B_{thresh} = -\frac{1}{2}\log_2\left(R_d(0) - R_d(1)\right) + 1.$$

For $|a_1| < 0.8$ the threshold occurs at or below one bit so in this range of a_1 the power dissipation increases almost linearly as a function of B.

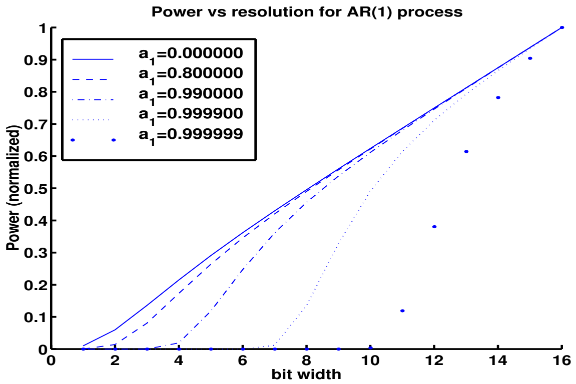

Fig. 1. *Power versus bit width B as function of AR parameter a_1 for loading a B-bit register with successive samples of an AR(1) process. Curves are normalized relative to power P_{16} consumed for a white sequence in a 16 bit register.*

III. Quantized Adaptive Channel Equalization

Figure 2 shows the block diagram of an adaptive equalization system with two different quantizers, denoted Q_d and Q_c, applied to the data and to the filter coefficients of an adaptive p-tap FIR filter. The quantizers Q_d and Q_c are allocated B_d bits plus sign and B_c bits plus sign, respectively. Here y_k is a (baseband) training signal, s_k is the transmitted signal, x_k is the received (baseband) data, and $\hat{s}_k$ is a linear estimate of the transmitted signal given the p samples $\underline{x}_k = [x_{k-1}, \ldots, x_{k-p}]^H$ and filter coefficients $\underline{w}_k = [w_{1k}, \ldots, w_{pk}]^T$.

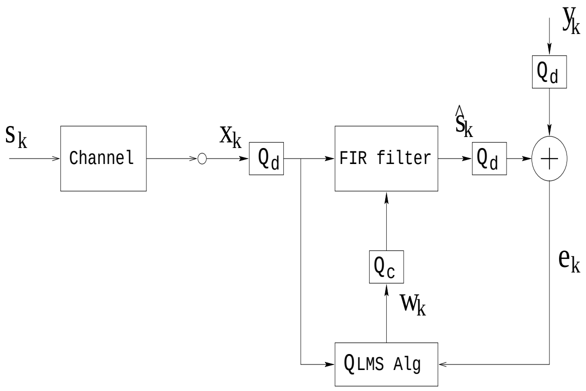

Fig. 2. *Quantized LMS adaptive channel equalizer with training sequence y_k.*

The standard LMS algorithm adapts the filter coefficients in an attempt to minimize the quadratic surface specified by the mean squared error: $E[|y_k - \hat{s}_k|^2]$ [4]. The quantized LMS algorithm, which we call QLMS, implements a recursive filter coefficient update of the form:

$$\underline{w}_{k+1} = \underline{w}_k + Q_c\left(\mu Q_d(\underline{x}_k) \cdot e_k\right)$$

where

$$e_k = Q_d(y_k) - Q_d\left(Q_d(\underline{x}_k^H) \cdot Q_c(\underline{w}_k)\right).$$

is the quantized error signal. Here μ is the gain parameter which controls the convergence properties of the algorithm.

The total power per iteration of the quantized LMS algorithm is determined by power dissipation of shift, add, multiply, memory load, and memory store operations. This depends on the specific design of the FIR filter and control circuitry. For illustration we will use the following expression for total power dissipation per iteration of LMS:

$$P_T = [24p(3B_d + B_c - 2) + 32p]\eta_g + 24p(B_d\eta_t). \tag{2}$$

This expression is linear in the number of bits B_d and B_c and assumes fixed point complex arithmetic, overwriting the data stack without using shift operations, multiplication using table lookup as opposed to adding partial products, and generic power coefficients η_g representing logic gate power consumption and η_t representing table lookup operation.

A. Performance of LMS Algorithm

The performance of the LMS adaptive algorithm is typically characterized by two quantities: the speed of convergence and the excess MSE. We assume that x_k, y_k and

s_k are all wide sense stationary random sequences. For a moderate number of filter coefficient quantization bits B_c, it can be shown that the convergence of QLMS in the mean is primarily determined by the quantized data $Q_d(x_k)$. On the other hand, the excess MSE is affected by both the number of data quantization bits B_d and the number of filter coefficient quantization bits B_c.

A.1 Mean Convergence

Define the $p \times p$ covariance matrix $R_{Q_d(x)} = \text{cov}(Q_d(\underline{x}_k))$ of the quantized data vector $Q_d(\underline{x}_k)$ and let this matrix have real non-negative eigenvalues $\{\lambda_i\}_{i=1}^p$. Further define the cross correlation vector $R_{Q_d(x),Q_d(y)} = \text{cov}(Q_d(\underline{x}_k), Q_d(y_k))$. Then the mean filter coefficients of the QLMS algorithm converge to the Wiener filter coefficients $\underline{w}^o$ of the quantized processes:

$$\lim_{k\to\infty} E[\underline{w}_k] = \underline{w}^o = R^{-1}_{Q_d(x)} R_{Q_d(x),Q_d(y)}$$

as long as the gain parameter μ satisfies the condition

$$0 < |1 - \mu\lambda_i| < 1, \qquad i = 1, \ldots, p.$$

When the QLMS algorithm converges, the MSE converges as a decaying exponential with the $1/e$ time constant of the slowest mode equal to $\tau_{3dB} = 1/(-\max_i \ln(|1 - \mu\lambda_i|)$, called the adaptation time constant. Note that the speed of convergence generally increases as μ increases.

A.2 Excess Mean Square Error

When the above condition for mean convergence of QLMS is met, an expression for the steady state mean square error can be derived.

$$\xi = E[|y_k - \hat{s}_k|^2] = \sigma_y^2 - R_{xy}^H R_x^{-1} R_{xy} + \mu \text{tr}(R_x) + \Delta\epsilon$$

where

$$\Delta\epsilon = \alpha_c \, 2^{-2B_c} + \alpha_d \, 2^{-2B_d}, \tag{3}$$

is the increase in MSE due to quantization,

$$\alpha_c = \frac{pA_c^2}{12\mu}, \qquad \alpha_d = \frac{(\|\underline{w}^o\|^2 + p)A_d^2}{6}$$

and A_d, A_c are the maximum amplitude ranges of the data and filter coefficient quantizers.

The expression above applies to complex sequences and is derived in a very similar manner to the derivation of Caraiscos and Liu [8] for real-valued sequences. As in [8] we make several standard assumptions including: the process x_k is circularly Gaussian, the quantization error is a zero mean white sequence, the quantization errors are independent of the data x_k and the filter coefficients $\underline{w}_k$. These assumptions are fairly restrictive but enable us to obtain useful closed form expressions.

With these relations the increase in MSE due to quantization, $\Delta\epsilon$, can be plotted as a function of B_d and B_c. A plot for the increase in MSE is given in Figs. 3 and 4

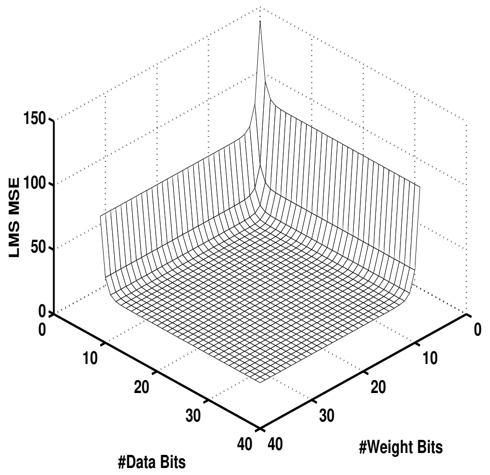

Fig. 3. *Excess MSE as a function of B_d and B_c for single pole IIR channel.*

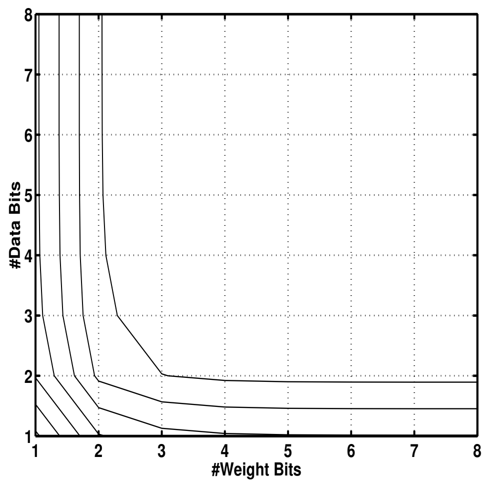

Fig. 4. *Excess MSE as a function of B_d and B_c for single pole IIR channel.*

for the case of white signals $s_k = y_k$ and for x_k generated by passing s_k through a single pole IIR filter ($p = 2$) with pole at $a_1 = 0.8$.

Observe from the surface plot that the increase in MSE is fairly flat over the range $B_c, B_d > 4$ but undergoes a rapid increase as either B_c or B_d decreases much below this range. The objective of the following section is to specify optimal choices of B_d and B_c which minimize this increase under various constraints on B_d and B_c.

IV. OPTIMAL BIT ALLOCATION STRATEGIES

We present expressions for the optimum allocation of bits to data versus filter coefficients under two constraints: total number of bits and total power consumption. While a combined study of bit allocation as a function of convergence rate and excess MSE is of importance, for concreteness we limit the focus in this paper to the excess MSE.

Assume that there are a total of $B_T + 2$ bits which are available to allocate between data and coefficients, i.e. $B_T = B_d + B_c$. Further define the data bit allocation factor

$\rho = B_d/B_T$. Then we have the obvious relations

$$B_d = \rho B_T, \qquad B_c = (1-\rho)B_T. \tag{4}$$

A. Total Bit-Width Constraint

Under a constraint on B_T the objective becomes to minimize the increase $\Delta\epsilon$ in MSE with respect to ρ. Graphically, this is the same as minimizing $\Delta\epsilon$ along the diagonal line $B_T = B_d + B_c$ of slope -1 in the B_c, B_d plane shown on Fig. 4. It is straightforward to show that $\Delta\epsilon$ is convex as a function of ρ with a single minimum occurring at the point $\rho = \rho^*$:

$$\begin{aligned}\rho^* &= \frac{1}{4B_T}\log_2\left(\frac{\alpha_d}{\alpha_c}\right) + \frac{1}{2} \\ &= \frac{1}{4B_T}\log_2\left(2\mu\frac{\|\underline{w}^o\|^2 + p}{p}\right) + \frac{1}{2}\end{aligned}$$

and the minimum value is

$$\begin{aligned}\min_\rho \Delta\epsilon &= 2^{-B_T+1}\sqrt{\alpha_d\alpha_c} \\ &= 2^{-B_T-1}\,\frac{1}{3}\sqrt{\frac{2p}{\mu}(\|\underline{w}^o\|^2 + p)}\end{aligned}$$

To obtain the concise relations involving $\underline{w}^o$ we have assumed that the ranges of the quantizers are identical $A_d = A_c = 1$.

Observe that the optimal bit allocation factor ρ^* converges to $1/2$ as the combined register length B_T goes to infinity. This is the regime where the standard allocation is optimal: allocate an equal number of bits to data as to filter coefficients. As register length decreases or convergence speed increases the standard allocation becomes suboptimal. In typical implementations, e.g. where AGC is implemented to prescale the data to unit variance, the gain parameter is chosen such that $\mu \ll 1$ to ensure convergence. In particular, if $\mu < 1/4$ then ρ^* is less than $1/2$ and more bits should be allocated to the filter coefficients than to the data. Also worth noting is that $\Delta\epsilon$ increases in p at a linear rate, decreases in μ at an inverse square root rate, and decreases in B_T at an exponential rate. Therefore, the total number of bits allocated gives more leverage over excess MSE than any other of the design parameters.

B. Total Power Constraint

Under the constraint on P_T, we can use (2) to re-express the total combined number of bits B_T as a function of ρ and P_T

$$B_T = \frac{P_T + 16p\eta_g}{p[\rho(48\eta_g + 24\eta_t) + 24\eta_g]} \tag{5}$$

Now using (4) in the expression for $\Delta\epsilon$ (3) we again obtain a convex function of ρ with unique minimum at $\rho = \rho^{**}$

$$\rho^{**} = \frac{\log_2\left[\frac{24\eta_g\alpha_d}{(72\eta_g+24\eta_t)\alpha_c}\right]\frac{24\eta_g p}{P_T+16\eta_g p} + 2}{-\log_2\left[\frac{24\eta_g\alpha_d}{(72\eta_g+24\eta_t)\alpha_c}\right]\frac{(48\eta_g+24\eta_t)p}{P_T+16\eta_g p} + 4},$$

which gives the corresponding minimum MSE

$$\min_\rho \Delta\epsilon = \alpha_c\, 2^{-2(1-\rho^{**})B_T} + \alpha_d\, 2^{-2\rho^{**}B_T}$$

where B_T is given in (5).

Observe that the optimal bit allocation factor ρ^{**} converges to the standard allocation $1/2$ as the total power constraint P_T is relaxed to infinity. As P_T decreases the standard allocation becomes suboptimal.

V. NUMERICAL EXAMPLE: 4-PSK

Here we briefly consider the case of a Gaussian noise IIR channel with a single pole at $a_1 = 0.8$, a 4-PSK signal s_k of unit variance, noiseless training sequence $y_k = s_k$, and a 2-tap LMS filter with gain coefficient $\mu = 0.01$. This corresponds to a rather severe exponential memory channel with intersymbol interference (ISI) extending over 5 to 10 data samples.

Figure 5 shows the B_T-constrained optimal data bit allocation factor ρ^* as a function of B_T superimposed on a plot of the resultant optimal MSE, ξ. Note that the formula (3) for $\Delta\epsilon$ is independent of any channel effects (i.e. independent of R_x and R_{xy}). Therefore, the channel affects only that portion of the MSE that is not attributed to quantization error. This quantity is visible at the large B_T region. Hence, the only effect of the channel on the MSE plot is to "shift" it up by the infinite precision MSE, $\sigma_y^2 - R_{xy}^H R_x^{-1} R_{xy} + \mu\mathrm{tr}(R_x)$. The channel has no effect on the shape of the plot, or, as will be shown, the optimal bit allocation factors. Figure 5 shows that MSE does not begin to degrade significantly until B_T falls below approximately 6 bits. For $B_T = 6$ bits the optimal data bit allocation factor is approximately $\rho^* = 0.4$ which means that 2 bits plus sign should be allocated to the data while 4 bits plus sign should be allocated to each of the filter coefficients.

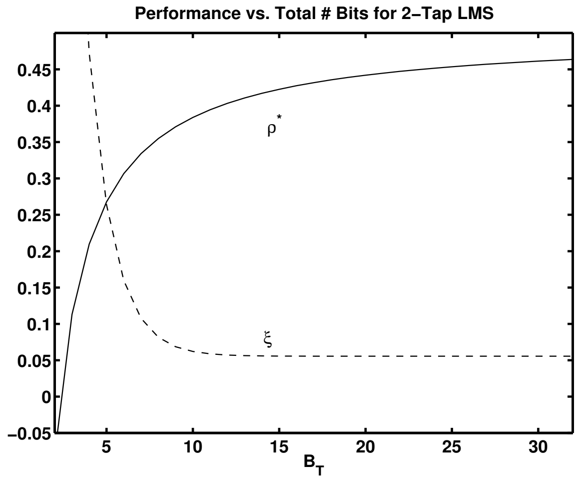

Fig. 5. *IIR channel with 4-PSK input. Optimal data bit allocation factor under B_T constraint and total MSE as function of B_T.*

Figure 6 shows the P_T-constrained optimal data bit allocation factor ρ^{**} as a function of P_T superimposed on a plot of MSE. The power coefficients used are $\eta_g = 1$ and

$\eta_t = 10$. Note that MSE does not degrade significantly until P_T falls below approximately 1200 Watts (normalized). At this breakdown point the optimal data bit allocation factor is approximately $\rho^* = 0.25$. We can use relation (5) with $\rho = \rho^{**}$, which is plotted in Fig. 7, to find the corresponding B_T as a function of P_T. From the plot we see that $P_T = 1200$ corresponds again to $B_T = 6$, but the optimal ρ^{**} tells us to allocate only 1 bit plus sign to the data and 5 bits plus sign to the filter coefficients. This reduction is because the data operations dominate the power relation (2).

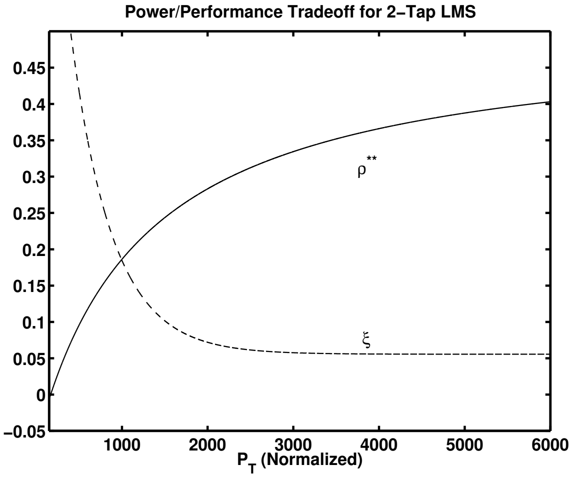

Fig. 6. *IIR channel with 4-PSK input. Optimal data bit allocation factor under P_T constraint and total MSE as function of P_T.*

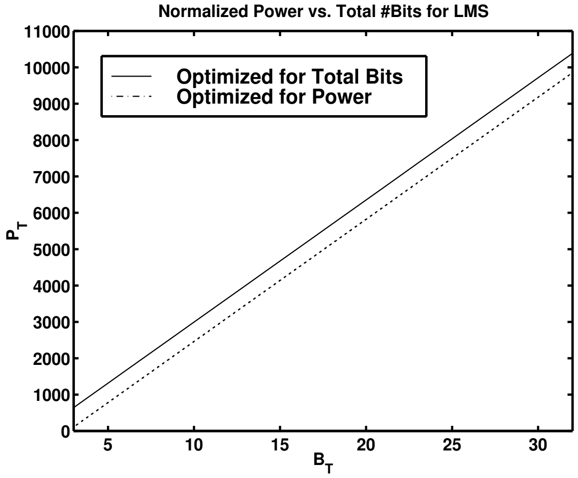

Fig. 7. *Power P_T as function of B_T for optimal bit allocation rules.*

Finally, in Fig. 8 we plot the MSE as a function of P_T for both the P_T-optimal and the B_T-optimal allocation factors ρ^{**} and ρ^*. Note that, as expected, for a given power constraint the P_T-optimal bit allocation yields lower MSE than the B_T-optimal bit allocation. While the difference between them is not great for this example, for other hardware designs, signal and channel statistics, the difference can be significant.

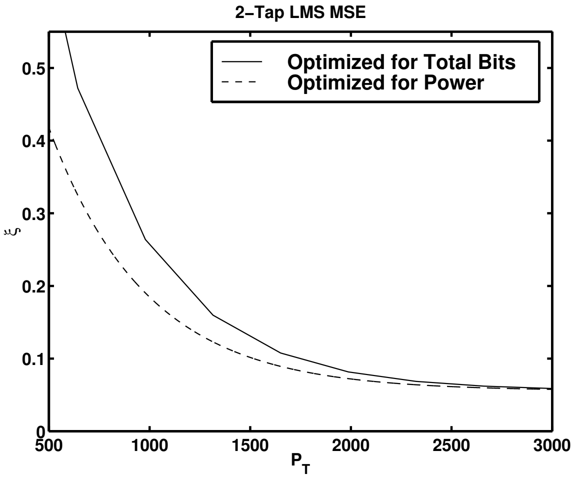

Fig. 8. *IIR channel with 4-PSK input. MSE as function of P_T for B_T-optimal and P_T-optimal allocation factors.*

VI. Conclusion

We have derived expressions for optimal bit allocation for adaptive LMS algorithms under combined register length constraints and total power constraints. These expressions can easily be specialized to a specific hardware implementation for computation of the number of bits to allocate to data and filter coefficients. A general conclusion is that the standard design strategy of allocating an equal number of bits to the data and filter coefficients is optimal only as the power or register length constraints get very large. For typical LMS implementations, it is optimal to allocate more bits to the coefficients than to the data. In particular, we have found that it is possible to reduce the number of bits in an LMS-based channel equalizer to 2 data bits and 6 coefficient bits without significant increase in steady state MSE.

References

[1] J. McChesney. *Notes for MURI Seminar on Low Energy Electronics Design for Mobile Platforms, The University of Michigan*, Oct. 1997.

[2] K. K. Parhi, "Algorithm transformation techniques for concurrent processors," *IEEE Proceedings*, vol. 77, pp. 1879–1895, Dec. 1989.

[3] A. P. Chandrasekan and R. W. Brodersen, "Minimizing power consumption in digital CMOS circuits," *IEEE Proceedings*, vol. 83, pp. 498–523, April 1995.

[4] B. Widrow and S. D. Stearns, *Adaptive signal processing*, Prentice Hall, Englewood Cliffs NJ, 1985.

[5] J. R. Treichler, C. R. Johnson, and M. G. Larimore, *Theory and Design of Adaptive Filters*, Wiley, New York, 1987.

[6] B. Widrow and /it et al, "Adaptive antenna systems," *IEEE Proceedings*, vol. 55, pp. 2143–2159, 1967.

[7] W. F. Gabriel, "Adaptive arrays – an introduction," *IEEE Proceedings*, vol. 64, pp. 239–272, Feb. 1976.

[8] C. Caraiscos and B. Liu, "A roundoff error analysis of the lms adaptive algorithm," *IEEE Trans. Acoust., Speech, and Sig. Proc.*, vol. ASSP-32, no. 1, pp. 34–41, Feb. 1984.

[9] D. Singh, J. M. Rabaey, M. Pedram, F. Catthoor, S. Rajgopal, N. Sehgal, and T. J. Mozden, "Power conscious CAD tools and methodologies: a perspective," *IEEE Proceedings*, vol. 83, pp. 5701–594, April 1995.