PERFORMANCE OF THE A PRIORI AND A POSTERIORI QR-LSL ALGORITHMS IN A LIMITED PRECISION ENVIRONMENT

Maria D. Miranda *Leonardo Aguayo* *Max Gerken*

LCS / DEE / EPUSP - PO Box: 61548 - São Paulo 01065-970 - Brazil
maria@lcs.poli.usp.br aguayo@lcs.poli.usp.br mgk@lcs.poli.usp.br

ABSTRACT

The performance of two minimal QR-LSL algorithms in a low precision environment is investigated. For both algorithms backward consistency and backward stability become guaranteed under simple numerical conventions. They present stable behavior even when excited with ill conditioned signals such as predictable signals. Since the problem of ensuring numerical stability is solved for these algorithms, an investigation about their accuracy is in place. By simulating a channel equalizer configuration it is shown that, for small mantissa wordlengths and forgetting factors λ not too close to 1, the *a priori* algorithm performs better due to its dispensing with passive rotations. For forgetting factors very close to one and small wordlengths, both algorithms are sensitive to the accuracy of some well-identified computations. They are compared to an LSL algorithm, based on *a priori* prediction errors, whose good performance in limited precision environments is known.

1. INTRODUCTION

Recursive Least Squares (RLS) algorithms are of interest for a variety of signal processing and control applications due to their robust convergence properties and consistent parameter estimates. Classical RLS algorithms have a numerical complexity of $\mathcal{O}(M^2)$, where M is the order of the filtering problem, while in fast versions complexity is reduced to $\mathcal{O}(M)$ by using the sequential nature of input data. Unfortunately, many of the fast RLS algorithms are plagued by numerical instability problems. Only recently the suspicion that fast RLS algorithm are numerically unstable has been dispelled [1]. This was achieved with the introduction of concepts such as minimality, backward consistency and backward stability, which have been used to establish sufficient conditions for stable propagation of numerical errors. To the best knowledge of the authors, the only known algorithms which are rigorously proven to be backward stable are two hybrid QR-LSL (Least Squares Lattice) algorithms, one based on *a posteriori* and the other based on *a priori* prediction errors. A compact description of the *a priori* QR-LSL algorithm [7, 8] is given in Tables I and II, while the *a posteriori* QR-LSL is described in [1] and [2].

These algorithms differ by the kind of prediction errors they compute, and have some characteristics that are worth mentioning:

- They are based on Givens rotations and are minimal in a system theory sense;
- Backward consistency and backward stability become guaranteed under simple numerical conventions;
- Contrasting with the *a posteriori* QR-LSL algorithm, in the *a priori* version backward consistency is guaranteed without the constraint of passive rotations in the recursive lattice section;
- They present stable behavior even when excited with ill conditioned signals such as predictable signals;
- The *a priori* version presents some degree of parallelism that can be exploited for fast implementations.

For these two algorithms the problem of ensuring numerical stability is completely solved. Thus, due to their numerical robustness, results about their accuracy are of great interest. Simulations in [3] show that the *a posteriori* QR-LSL algorithm gives the best accuracy among numerous other algorithms. The *a priori* QR-LSL algorithm and the LSL algorithm of [4], based on *a priori* prediction errors, have not been considered in [3]. The latter algorithm is known to diverge in adverse situations and conditions that guarantee its backward stability are not known. Despite this, the algorithm produces accurate results even in low precision environments. Therefore, to further investigate this question we made some comparisons between the *a posteriori* and *a priori* QR-LSL algorithms and the *a priori* error-feedback LSL algorithm of [4]. We performed several simulations to investigate the influence of the mantissa wordlength, the order and the forgetting factor on the performance of the mentioned algorithms.

2. IMPLEMENTATION ISSUES

In this section, based on the methodology introduced in [1], we address some properties and implementation issues that determine the numerical behavior of the algorithm given in Table II.

The prediction section of the RLS algorithm can be described by a non-linear system

$$\boldsymbol{\zeta}(n+1) = T\{\boldsymbol{\zeta}(n), u(n+1)\}\,, \tag{1}$$

where $\boldsymbol{\zeta}(n)$ is the state vector that stores all the variables necessary to propagate the LS solution, and $u(n+1)$ is the input data sequence. Considering the *a priori* QR-LSL algorithm in Table II, it follows that the prediction section can be described by (1) through

$$\boldsymbol{\zeta}(n) = \left[\ \mathbf{q}_M^{fT}(n),\ \xi_M^{f/2}(n),\ \overline{\boldsymbol{\psi}}_M^T(n)\ \right]^T\,, \tag{2}$$

where $\mathbf{q}_M^f(n) = [q_{i-1}^f]$, $\xi_M^{f/2}(n)$ and $\overline{\boldsymbol{\psi}}_M(n) = [\overline{\psi}_{i-1}(n)]$, for $1 \le i \le M$, are respectively the column vector of the transformed input data (forward prediction), the square root of the forward prediction error energy and the column vector of the normalized *a priori* backward prediction error. Since this state vector has the dimension $2M+1$, M being

This work was supported by FAPESP, CAPES and CNPq.

the order of the prediction problem, the QR-LSL *a priori* algorithm is minimal [1].

The sets $\mathcal{S}_e(n)$ and $\mathcal{S}_f(n)$ are the sets of all states $\boldsymbol{\zeta}(n)$ reachable at time n using respectively exact and finite precision arithmetic. We can show, by analogy to [1], that the set $\mathcal{S}_e(n)$ consists of all states satisfying

a) $\xi_M^{f/2}(n) > 0 \;\Leftrightarrow\; |\phi_i(n)| < \pi/2,$

b) $\|\overline{\boldsymbol{\psi}}_M(n)\| < \infty \;\Leftrightarrow\; \left|\theta_i^b(n)\right| < \pi/2 \;\Leftrightarrow\; 1 \leq \gamma_M^{-1} < \infty,$

for $1 \leq i \leq M$. To assure that $\mathcal{S}_f(n) \subset \mathcal{S}_e(n)$, it is only necessary to guarantee that, regardless of numerical errors, conditions (a) and (b) are satisfied. Condition (a) can be easily satisfied since $\xi_M^{f/2}(n)$ is computed as a norm of a two element vector which is nonzero in case of persistent excitation. The inequality $\|\overline{\boldsymbol{\psi}}_M(n)\| < \infty$ turns out to be automatically satisfied when using any usual finite precision arithmetic. No constraints are imposed to $\mathbf{q}_M^f(n)$ and $\overline{\boldsymbol{\psi}}_M(n)$, thus the rotations $\theta_i^b(n)$ and $\phi_i(n)$ which propagate these variables do not need to be passive in finite precision. Therefore, it is possible to use rounding in all operations of the proposed algorithm and guarantee backward stability, which is not the case of the algorithm described in [2]. By the use of rounding, the accumulation of bias, which otherwise results with passive rotations, is reduced.

The essential difference between the *a priori* and *a posteriori* QR-LSL algorithms is the type of prediction error used. Here the column vector of the normalized *a posteriori* backward prediction error is represented as $\overline{\mathbf{b}}_M(n) = [\overline{b}_{i-1}(n)]$, for $1 \leq i \leq M$. As already shown, one consequence of using *a priori* normalized prediction errors is that the stability domain of the proposed algorithm is open in what concerns these errors. More precisely, the reachable ranges for the normalized *a posteriori* and *a priori* errors are respectively $0 \leq \|\overline{\mathbf{b}}_M(n)\| < 1$ and $0 \leq \|\overline{\boldsymbol{\psi}}_M(n)\| < \infty$. Therefore, one might expect the normalized *a priori* errors $\overline{\psi}_{i-1}(n)$ to be greater than the *a posteriori* errors $\overline{b}_{i-1}(n)$, implying smaller relative errors in the numerical representation of $\overline{\psi}_{i-1}(n)$ and, consequently, resulting in better accuracy of the *a priori* algorithm. For stationary persistent excitations, λ close to one and for large n, the following approximations are valid for $1 \leq i \leq M$ (see the appendix):

$$E[|\overline{b}_{i-1}(n)|^2] \approx \lambda^{i-1}(1-\lambda), \tag{3}$$

$$E[|\overline{\psi}_{i-1}(n)|^2] \approx \lambda^{-i}(1-\lambda), \tag{4}$$

$$E[\gamma_i(n)] \approx \lambda^i, \tag{5}$$

where $E[.]$ represents the expectation operator. Since $E[|\overline{b}_{i-1}(n)|] \approx 0$ and $E[|\overline{\psi}_{i-1}(n)|] \approx 0$ for stationary signals, both the *a posteriori* and the *a priori* backward prediction errors tend to assume small values for λ slightly less than one. In fact, they make little use of their reachable ranges $0 \leq \|\overline{\mathbf{b}}_M(n)\| < 1$ and $0 \leq \|\overline{\boldsymbol{\psi}}_M(n)\| < \infty$. Anyway, as the order M increases and λ decreases, the variance of the *a priori* backward prediction error increases more than the variance of the corresponding *a posteriori* error.

It is also worth noting that (3), (4) and (5) indicate that the precision in step 1, the angle θ^b solving part, may be specially important when $\lambda^{1/2}$ is close to one. To compute the norm $\gamma_i^{-1/2}(n) = (|\gamma_{i-1}^{-1/2}(n)|^2 + |\bar{\psi}_{i-1}(n)|^2)^{1/2}$ using a general purpose computer or a signal processor, a number close to one, $|\gamma_{i-1}^{-1/2}(n)|^2$, is to be added to a number close to zero, $|\bar{\psi}_{i-1}(n)|^2$. This suggests that the accuracy of this computation is sensitive to the mantissa wordlength. Some simulation results in the next section confirm this conjecture. Finally, it must be observed that the *a posteriori* QR-LSL algorithm presents almost the same problem in the corresponding angle θ^b solving part. There, the norm $\gamma_{i-1}^{1/2}(n) = (|\gamma_i^{1/2}(n)|^2 + |\bar{b}_{i-1}(n)|^2)^{1/2}$ is computed.

To close this section, it should be stressed that backward consistency is not per se wordlength dependent. However, the accuracy of the computed signals and variables is obviously affected by the wordlength. Simulation results [3] show that, in a low-precision environment, the *a posteriori* QR-LSL algorithm gives the best results among numerous other known RLS algorithms. In the next section we present some simulation results that show the performance of the *a priori* and *a posteriori* QR-LSL algorithms.

3. SIMULATION RESULTS

We considered an adaptive channel equalizer in the same configuration and with the same data channel used in [6, Chapters 9 and 13]. The data channel has a raised cosine unit pulse response with only three non zero samples: $h(n) = [1 + \cos(2\pi(n-1)/W)]/2$ for $n = 0, 1, 2$. The parameter W controls the amount of intersymbol interference and the eigenvalue spread of the correlation matrix at the equalizer input. We considered $W = 3.5$ which corresponds to an eigenvalue spread of the equalizer input correlation matrix of approximately 47. The channel was fed with a polar (± 1) pseudorandom sequence, and a delayed version of the input sequence was used as reference signal for the filtering algorithm. Zero mean gaussian noise was added at the channel output resulting in an SNR of 30dB at the equalizer input.

Using this configuration we measured the mean square *a priori* estimation error for several mantissa wordlengths, several values of $\lambda^{1/2}$ and several equalization orders M. The results for two values of $\lambda^{1/2}$ are shown in Figure 1. The exponent wordlength was fixed in 8 bits and it was also supposed that a multiply-accumulator machine similar to a commercial signal processor was available. To prevent the results from being distorted by initialization effects, we discarded the output of the first 5000 iterations. The mean squared *a priori* estimation error was then averaged in time during 5000 iterations. The two QR-LSL algorithms use $\lambda^{1/2}$ in their computations. Therefore, $\lambda^{1/2}$ was always chosen to be exactly represented by the lower mantissa wordlength used in a set of measurements. The corresponding value of λ used in the *a priori* error-feedback LSL algorithm was obtained rounding the square of the quantized $\lambda^{1/2}$ to the lower mantissa precision used in a set of measurements.

Observing the plots in Figure 1 one immediately identifies two ranges of mantissa wordlengths: one where the MSE is insensitive to mantissa wordlength variations and the other where the performance of the algorithms suffers a degradation. For fixed λ and fixed order M, in the first wordlength range the measured MSE is very close to the corresponding expected value (MSE_o). This behavior suggests the definition of a minimal mantissa wordlength w_{min}, above which appropriate performance occurs. It may be defined as the minimal number of bits that ensure a deviation from MSE_o lower than a threshold level of 0.5 dB. The actual threshold level is not very important as long as it is the same for all comparisons. Here, w_{min} is useful only in describing qualitatively the behavior of the QR-LSL algorithms. Generally, w_{min} increases with λ

and with M. This behavior, illustrated by Figure 1, was observed for several filter orders and several values of $\lambda^{1/2}$.

Figure 1 clearly shows that, for short mantissa wordlengths, the *a priori* error-feedback LSL algorithm performs better than the QR-LSL algorithms. However, it should be remembered that the better accuracy is achieved at the expense of stability, not guaranteed for the former algorithm.

Now, considering only the QR-LSL algorithms, some comments about their behavior are in place. Figures 1.a-b illustrate the typical performance observed for moderate values of $\lambda^{1/2}$, i. e. not so close to one. Specifically for the simulation results in Figure 1, moderate values of $\lambda^{1/2}$ are understood as those lower than $1 - 5/256 = 0.98046875$. The *a priori* QR-LSL algorithm shows better performance for mantissa wordlengths lower than w_{min}. For a fixed wordlength the performance difference increases with increasing order M and decreasing λ, as long as $\lambda^{1/2}$ is not very close to 1.

Figures 1.c-d illustrate the typical performance observed for values of $\lambda^{1/2}$ close to one. For the displayed simulation results, this implies values of $\lambda^{1/2}$ greater than 0.98046875. In this case the minimal wordlength w_{min} is better defined and both algorithms show a pronounced performance degradation for mantissa wordlengths lower than w_{min}. The degradation can be so strong as to preclude the use of either algorithm. Despite this observation, the *a posteriori* algorithm performs better. Our attempts to explain this degradation revealed that computations done in the θ^b angle solving part (first line of equation in step 1 of Table II or its equivalent in the *a posteriori* algorithm) were the primary source of error. To stress this, Figures 1.c-d include simulation results where only the square root arguments in the θ^b angle solving part were represented using double precision mantissa wordlengths. For all other variables single precision was used. As shown in Figures 1.c-d, for short wordlengths and $\lambda^{1/2}$ close to one, there are significant performance improvements for both algorithms which show similar behavior in this case. For moderate values of $\lambda^{1/2}$, increasing the mantissa wordlength as described does not cause any significant improvement. The displayed results also indicate that the *a priori* algorithm is more sensitive to the computational accuracy of the θ^b angle solving part.

In order to explain how the passivity constraint may affect the performance of the QR-LSL algorithms, we simulated the *a priori* algorithm with passive rotations in the recursive lattice section. In this case, Figure 1.a-b shows that the two algorithms have almost the same performance for moderate values of $\lambda^{1/2}$. Thus, the *a priori* QR-LSL algorithm shows better performance because passive rotations are not needed. In fact, the passivity constraint degrades performance only for mantissa wordlengths smaller than w_{min} and moderate values of $\lambda^{1/2}$. As displayed in Figures 1.c-d, for $\lambda^{1/2}$ very close to one the effect of passive rotations is not significant, even for mantissa wordlengths lower than w_{min}.

4. CONCLUDING REMARKS

The performance of the *a priori* and *a posteriori* QR-LSL algorithms was compared and it results that, for small mantissa wordlengths and forgetting factors not too close to 1, the *a priori* QR-LSL algorithm performs better, the reason being that passive rotations are not necessary. Moreover, for small wordlengths and forgetting factors very close to 1, the overall accuracy of both QR-LSL algorithms is specially sensitive to the accuracy of the θ^b angle solving part. The comparison of the two QR-LSL algorithms with the *a priori* error feedback LSL algorithm shows that, for low mantissa wordlengths, there is some margin for accuracy improvements in both QR-LSL algorithms. Finally, it should be emphasized that both QR-LSL algorithms are backward stable and present stable behavior even when excited with ill conditioned signals such as predictable signals.

APPENDIX

To show that the approximations (3), (4) and (5) are valid for stationary signals, $n \to \infty$ and $\lambda \approx 1$, we use the fact that the conversion factor $\gamma_i(n)$ can be expressed as $\gamma_i(n) = \lambda^i \xi_i^b(n) \xi_i^{-f}(n)$. Therefore

$$E\left[\gamma_i(n)\right] = \lambda^i E\left[\xi_i^b(n)\xi_i^{-f}(n)\right]. \quad (6)$$

Since $\xi_i^b(n) = \sum_{\ell=0}^{n} \lambda^{n-\ell} |e_i^b(\ell)|^2$, with $e_i^b(\ell)$ being the backward prediction error of order i, the above becomes

$$E\left[\gamma_i(n)\right] = \lambda^i \sum_{\ell=0}^{n} \lambda^{n-\ell} E\left[|e_i^b(\ell)|^2 \xi_i^{-f}(n)\right]. \quad (7)$$

Employing the averaging principle [5], which is valid for $n \to \infty$ and λ close to one, we obtain

$$E\left[|e_i^b(\ell)|^2 \xi_i^{-f}(n)\right] \approx E\left[|e_i^b(\ell)|^2\right] E\left[\xi_i^{-f}(n)\right]. \quad (8)$$

Considering now $\xi_i^f(n) = \sum_{\kappa=0}^{n} \lambda^{n-\kappa} |e_i^f(\kappa)|^2$, with $e_i^f(\kappa)$ being the forward prediction error of order i, and at time κ, $0 \le \kappa \le n$, and $E\left[|e_i^f(\kappa)|^2\right] \approx E\left[|e_i^b(\ell)|^2\right]$, which is valid for stationary signals, it follows readily from (7) and (8) that

$$E\left[\xi_i^b(n)\xi_i^{-f}(n)\right] \approx \sum_{\ell=0}^{n} \lambda^{n-\ell} \left(\sum_{\kappa=0}^{n} \lambda^{n-\kappa}\right)^{-1} = 1. \quad (9)$$

The approximations (5) and (3) result from substituting (9) respectively into (6) and from $|\bar{b}_{i-1}(n)|^2 = \gamma_{i-1}(n) - \gamma_i(n)$. Applying the expectation operator on $|\overline{\psi}_{i-1}(n)|^2 = \gamma_i^{-1}(n) - \gamma_{i-1}^{-1}(n)$ and with the same arguments used in the derivation of (3), the approximation (4) follows.

REFERENCES

[1] P.A. Regalia: *Numerical stability properties of a QR-based fast least squares algorithm*, IEEE Trans. on SP, vol. SP-41, pp. 2096-2109, June 1993.

[2] P.A. Regalia, M.G. Bellanger: *On the duality between fast QR methods and lattice methods in least squares adaptive filtering*, IEEE Trans. on SP, vol. SP-39, pp. 879-891, April 1991.

[3] M.D. Levin, C.F.N. Cowan: *The performance of 8 recursive least squares adaptive filtering algorithms in a limited precision environment*, Signal Processing VII. Theories and Applications, Edinburg, Scotland, pp. 1261-1264, September 1994.

[4] F. Ling, D. Manolakis, J.G. Proakis: *Numerically robust least-squares lattice-ladder algorithms with direct updating of the reflection coefficients*, IEEE Trans. on ASSP, vol. ASSP-34, pp. 837-845, August 1986.

[5] C. Samson, V.U. Reddy: *Fixed point error analysis of the normalized ladder algorithms*, IEEE Trans. on ASSP, vol. ASSP-31, pp. 1177-1191, October 1983.

[6] S. Haykin: *Adaptive Filter Theory*, Prentice Hall Int., 2nd. Ed., 1991.

[7] M.D. Miranda, M. Gerken: *A hybrid QR-lattice least squares algorithm using a priori errors.* Proc. 38*th* MWSCAS, Rio de Janeiro, pp. 983-986, August 1995.

[8] M.D. Miranda, M. Gerken: *A hybrid QR-lattice least squares algorithm using a priori errors.* Submitted to IEEE Trans. on SP.

Table I - Variables

λ - forgetting factor;

$u(n)$ and $d(n)$ - input and reference signal;

Θ_i^b, Θ_i^f and Φ_i - Givens rotation of angle θ_i^b, θ_M^f and ϕ_i. For example:

$$\Theta_i^b = \begin{bmatrix} \cos\theta_i^b(n) & \sin\theta_i^b(n) \\ -\sin\theta_i^b(n) & \cos\theta_i^b(n) \end{bmatrix};$$

$\varepsilon_i^f(n)$ - rotated forward prediction errors;

$\varepsilon_i(n)$ - rotated estimation errors;

$q_i^f(n)$ and $q_i(n)$ - rotated input and reference signals;

$\xi_i^{f/2}(n)$ - square root of forward prediction error energy;

$\gamma_M^{1/2}(n)$ and $\gamma_M^{-1/2}(n)$ - square root of conversion factor and its inverse;

$\bar{\psi}_i(n)$ - normalized *a priori* backward prediction errors;

$\bar{\eta}_i(n)$ - normalized *a priori* forward prediction errors.

Table II - *A priori* QR-LSL algorithm

Initialization: $n = 0$, $\xi_M^{f/2}(0)$ = small positive const.
For $i = 0,\ 1,\ ...,\ M-1$ do
$\{\ q_i^f(0) = q_i(-1) = 0;\quad \bar{\psi}_i(0) = 0\ \}$.

Step 1:
$\gamma_0^{-1/2}(n) = 1$; $\varepsilon_0^f(n+1) = u(n+1)$; $\varepsilon_0(n) = d(n)$
For $i = 0,\ 1,\ ...,\ M-1$ do

$$\begin{bmatrix} 0 & \gamma_{i+1}^{-1/2}(n) \\ \varepsilon_{i+1}^f(n+1) & q_i^f(n+1) \\ \varepsilon_{i+1}(n) & q_i(n) \end{bmatrix} = \begin{bmatrix} \bar{\psi}_i(n) & \gamma_i^{-1/2}(n) \\ \varepsilon_i^f(n+1) & \lambda^{1/2} q_i^f(n) \\ \varepsilon_i(n) & \lambda^{1/2} q_i(n-1) \end{bmatrix} \Theta_{i+1}^b$$

Step 2a:

$$\begin{bmatrix} 0 & \xi_M^{f/2}(n+1) \end{bmatrix} = \begin{bmatrix} \varepsilon_M^f(n+1) & \lambda^{1/2}\xi_M^{f/2}(n) \end{bmatrix} \Theta_M^f$$

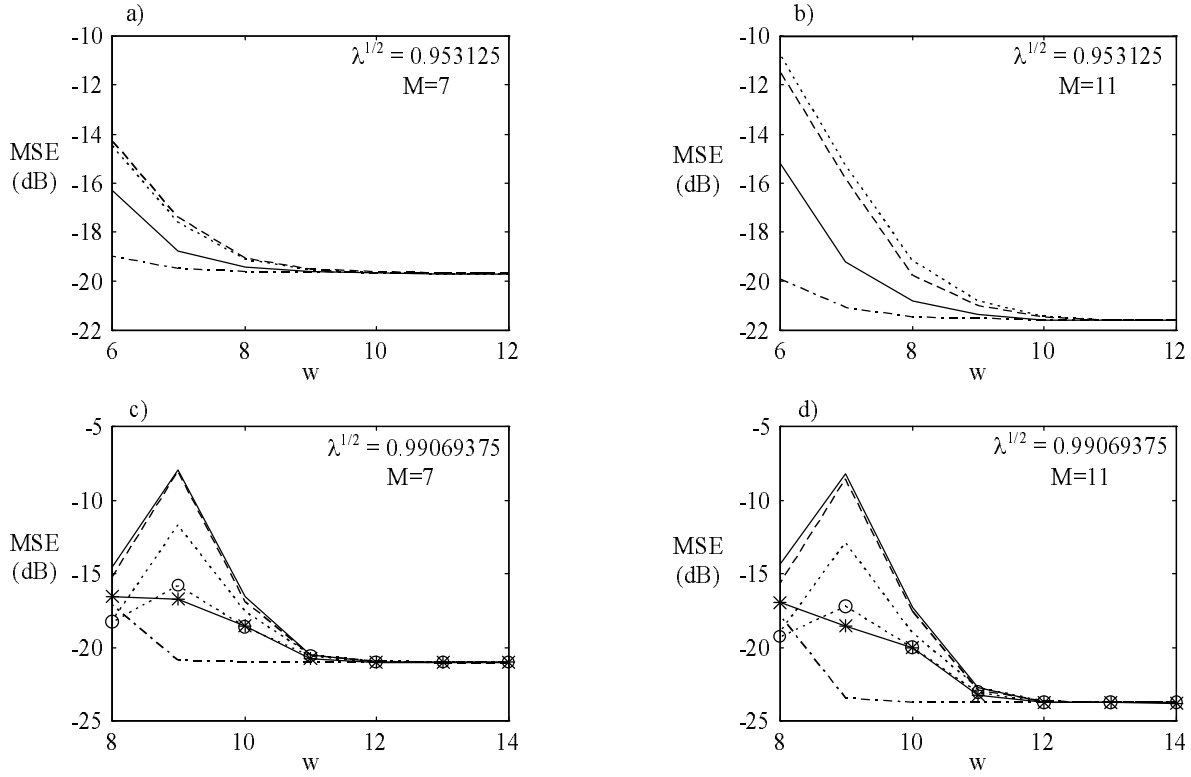
$\bar{\eta}_M(n+1) = \gamma_M^{-1/2}(n)\sin\theta_M^f / \cos\theta_M^f$

Step 2b: For $i = M-1,\ M-2,\ ...,\ 0$ do

$$\begin{bmatrix} 0 & \xi_i^{f/2}(n) \\ \bar{\psi}_{i+1}(n+1) & \bar{\eta}_i(n+1) \end{bmatrix} = \begin{bmatrix} q_i^f(n) & \xi_{i+1}^{f/2}(n) \\ \bar{\psi}_i(n) & \bar{\eta}_{i+1}(n+1) \end{bmatrix} \Phi_{i+1}$$

with $\bar{\psi}_0(n+1) = \bar{\eta}_0(n+1)$.

A posteriori estimation error: $e_M(n) = \varepsilon_M(n)/\gamma_M^{-1/2}(n)$



· · · · · · *a posteriori* QR-LSL satisfying passivity constraints.
——— *a priori* QR-LSL with rounding in all operations.
– – – *a priori* QR-LSL with passive rotations at the recursive lattice section.
– · – · *a priori* error-feedback LSL algorithm.
(∗) or (o) indicate use of double mantissa wordlength to represent arguments of square-roots in the θ^b angle solving part of the QR-LSL algorithms.

Figura 1. Averaged *a priori* MSE (Mean Square Estimation) error for channel equalizer experiment, w is the mantissa wordlength excluding the sign bit and M is the order of the equalizer.