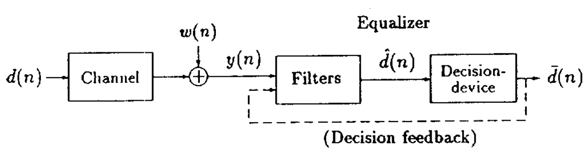
ROBUST DECISION FEEDBACK EQUALIZERS

Mikael Sternad, Anders Ahlén and Erik Lindskog
Systems and Control Group, Uppsala University
P.O. Box 27, S–751 03 Uppsala, Sweden

Abstract. Design equations are presented for robust and realizable decision feedback equalizers, for IIR channels with coloured noise. Given a probabilistic measure of model uncertainty, the mean MSE, averaged over a whole class of possible models, is minimized. A second type of robustification, which reduces the error propagation due to the feedback, is also introduced. The resulting design equations define a large class of equalizers, with DFE's and linear equalizers based on nominal models being special cases.

1. INTRODUCTION

If data sequences $\{d(n)\}$ are transmitted in the presence of intersymbol interference, they have to be reconstructed from the received sequences $\{y(n)\}$. Equalizers compute estimates $\bar{d}(n)$ on a symbol by symbol basis. Their main advantage, as compared to the MLSE Viterbi detector, is a low computational complexity.



If channels are slowly time–varying, filter coefficients can be adjusted during known training sequences, and held fixed until the next training. For fast time–variations, adaptive structures have to be used. The conventional approach to adaptive equalization is to make a decision–directed adjustment of the *filter* coefficients directly. An interesting alternative is ***indirect adaptation***: a model of the *channel* (and possibly also of the noise) is adjusted, often with decisioned data $\bar{d}(n)$ being used as channel model input [1].

One advantage of the indirect approach is that in e.g. Rayleigh–fading environments, channel parameters change much more "smoothly" than do the optimal values of equalizer parameters. Thus, it is much easier for an adaptive algorithm to track them. Furthermore, the number of channel parameters is mostly smaller than the required number of equalizer coefficients. Another advantage is that effects of modelling errors can be minimized analytically in an indirect approach, as described below.

In an indirect approach, equations are needed which optimize realizable equalizers, for given channel and noise models. For IIR channels with coloured noise, calculation of *linear recursive equalizers* was described in [1]. Equations for the more high–performance *Decision Feedback Equalizer* (DFE) were presented in [8]. While simple to use, these methods have two main limitations:

- Modelling errors are not taken into account. The resulting equalizer performance can be sensitive to such errors, in particular for channels with resonance peaks or deep nulls.
- The DFE in [8] was optimized under the assumption that past decisions were correct. Sensitivity to erroneous past decisions was not considered explicitly. The resulting DFE's sometimes generate long error propagation events.

Design equations which take these two problems into account are presented in Section 3 below.

2. MODEL AND FILTER STRUCTURE

The robustification is based on stochastic representation of the mismodelling and of decision errors.

We describe the received, discrete–time, complex baseband signal $y(n)$ as

$$y(n) = \left(\frac{B_0(q^{-1})}{A_0(q^{-1})} + \frac{\Delta B(q^{-1})}{A_1(q^{-1})}\right) d(n-k) + w(n) \quad (1)$$

with q^{-1} being the backward shift operator ($q^{-1}y(n) = y(n-1)$). The transmitted symbols $\{d(n)\}$ are assumed to be zero mean and white, with $E|d(n)|^2 = \sigma_d^2$.

[1] Direct adaptation corresponds to recursive minimization of $|\hat{d}(n) - \bar{d}(n)|^2$ w.r.t. filter parameters, using LMS or RLS. Indirect methods adjust a channel model, with output $\hat{y}(n)$, to minimize $|y(n) - \hat{y}(n)|^2$. Use of *a priori* information to improve tracking of time–varying FIR channels is described in [7].

0-7803-0946-4/93 $3.00

The noise $w(n)$ is described by

$$w(n) = \left(\frac{M_0(q^{-1})}{N_0(q^{-1})} + \frac{\Delta M(q^{-1})}{N_1(q^{-1})} \right) v(n) \qquad (2)$$

where $v(n)$ is zero mean and white, with (uncertain) standard deviation σ_v. In these time–invariant models, B_0/A_0 and M_0/N_0 represent stable and known ***nominal models*** of transfer functions, while $\Delta B/A_1$ and $\Delta M/N_1$ are members of model error classes. Coefficients of their numerator polynomials, eg. $\Delta B(q^{-1}) = \Delta b_0 + \Delta b_1 q^{-1} + \ldots + \Delta b_{\delta b} q^{-\delta b}$, are regarded as (time-independent) stochastic variables, with zero means and known covariance matrices. The stable denominators A_1 and N_1 are fixed. In [9], such representations are shown to be suitable for describing a wide range of model error types. They are related to the stochastic embedding approach of Goodwin [4], [5].

For example, a FIR channel with white noise is described by

$$y(n) = (B_0(q^{-1}) + \Delta B(q^{-1}))d(n-k) + v(n) \quad .$$

If $\hat{y}(n) = B_0(q^{-1})d(n-k)$ has been estimated by the least squares method and the order of B_0 is adequate, the LS covariance matrix can serve directly as an estimate of the covariances $E(\Delta b_i \Delta b_j^*)$.

Let us introduce an IIR decision feedback equalizer

$$\hat{d}(n-\ell|n) = \frac{S(q^{-1})}{R(q^{-1})} y(n) - \frac{Q(q^{-1})}{P(q^{-1})} \bar{d}(n-\ell-1) \quad (3)$$

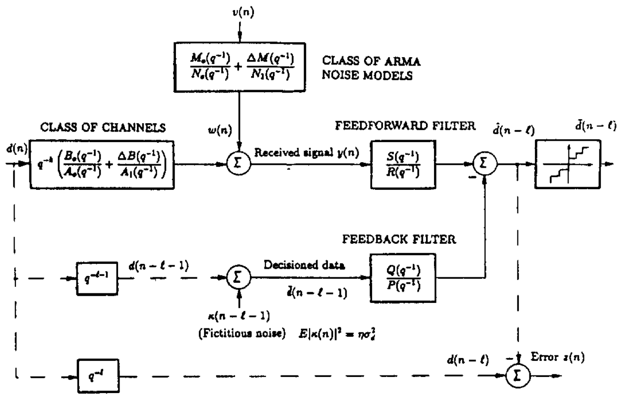
where ℓ is a user–chosen smoothing lag and $\bar{d}(n)$ is decisioned data. The denominator polynomials $R(q^{-1})$ and $P(q^{-1})$ are required to be monic and stable.

Errors in the decisioned data $\bar{d}(n)$ will be treated as uncertainty, and represented by ***additive white noise*** $\kappa(n)$, uncorrelated[2] with $d(n-j)$ and $v(n-j)$ for all j

$$\bar{d}(n) = d(n) + \kappa(n) \quad ; \quad E|\kappa(n)|^2 = \eta \sigma_d^2 \quad . \qquad (4)$$

The problem can then be solved with tools for linear quadratic design, since the nonlinear decision element is removed from the signal path to the error $d - \hat{d}$.

The scale factor $\eta \geq 0$ is used to trade off error propagation against theoretical performance, with $\eta = 0$ representing a belief in error-free decisioned data. The use of a small positive value of η often gives a lower bit error rate, in cases with severe error propagation for a design based on $\eta = 0$ [3].



3. FILTER DESIGN EQUATIONS

Now, a single equalizer is to be optimized with respect to the whole model error class. We minimize an *averaged MSE criterion*

$$J = \bar{E}E|d(n-\ell) - \hat{d}(n-\ell|n)|^2 \qquad (5)$$

where E represents expectation over d, v and κ and $\bar{E}$ is an expectation over the model error distribution in (1) and (2). This type of criterion has been used in connection to other filtering problems, e.g. by Chung and Bélanger [3]. Note that not only the *range* of uncertainties, but also their *likelihood* is taken into account by (5); common model deviations will have a greater impact on an estimator design than do very rare "worst cases". Compared to the use of a minimax design, the conservativeness is thus reduced.

In [10], design equations have been derived for minimizing the criterion (5) with respect to the filter coefficients of the DFE (3), for an ensemble of systems (1),(2), assuming (4). A novel derivation technique, described in [2] and in [9], has been used. For polynomials $P = p_0 + p_1 q^{-1} + \ldots$, let $P_* \triangleq p_0^* + p_1^* q + \ldots$. Define polynomials $H(q^{-1}), A(q^{-1})$ and $N(q^{-1})$ as

$$H \triangleq B_0 A_1 N_0 N_1 \quad ; \quad A \triangleq A_0 A_1 \quad ; \quad N \triangleq N_0 N_1 \quad .$$

Define double–sided polynomials $\tilde{B}\tilde{B}_*(q, q^{-1})$ and $\tilde{M}\tilde{M}_*(q, q^{-1})$ by

$$\begin{aligned} \tilde{B}\tilde{B}_* &\triangleq B_0 B_{0*} A_1 A_{1*} + \bar{E}(\Delta B \Delta B_*) A_0 A_{0*} \\ \tilde{M}\tilde{M}_* &\triangleq M_0 M_{0*} N_1 N_{1*} + \bar{E}(\Delta M \Delta M_*) N_0 N_{0*} \quad . \end{aligned} \qquad (6)$$

[2]This is, of course, a simplification. In reality, the error $\kappa(n)$ is non–stationary since decision errors tend to occur in bursts. There may also exist correlations to past noise samples, in particular to those that caused the error. These nonlinear and time–varying effects are neglected here, to obtain a tuning parameter which is simple to use.

[3]For a related suggestion, using a linear combination of a zero forcing linear equalizer and a zero forcing DFE, see [6].

Finally, define

$$\rho = \bar{E}(\sigma_v^2)/\sigma_d^2 \ .$$

Now, let the scalar r_1 and the stable monic polynomial $\beta(q^{-1})$ be the solution to the ***averaged spectral factorization***

$$\begin{aligned} r_1\beta\beta_* &= NN_*A_0A_{0*}\bar{E}(\Delta B\Delta B_*) \\ &+ \eta NN_*\tilde{B}\tilde{B}_* + (1+\eta)\rho AA_*\tilde{M}\tilde{M}_* \ . \end{aligned} \quad (7)$$

Let $\{Q(q^{-1}), S_1(q^{-1}), L_{1*}(q), L_{2*}(q)\}$ be the (unique) solution to the ***coupled polynomial equations***

$$\beta + q^{-1}(1+\eta)Q = q^{\ell-k}HS_1 + \beta L_{1*} \quad (8)$$

$$-q^{-\ell+k}\eta H_* + q(1+\eta)L_{2*} = -r_1\beta_*S_1 + q^{-\ell+k}H_*L_{1*} \quad (9)$$

Then, the equalizer (3), which minimizes (5), is given by Q from (8) and by

$$S = S_1NA \ ; \ R = \beta \ ; \ P = \beta \ . \quad (10)$$

This result can be derived by adding a variation $\nu(n) = \mathcal{T}_1y(n) + \mathcal{T}_2\bar{d}(n-\ell-1)$, with $\mathcal{T}_1$ and $\mathcal{T}_2$ arbitrary but rational and stable, to the estimate $\hat{d}(n-\ell|n)$. Optimality corresponds to orthogonality of the mean estimation error with respect to this variation. The equations (8) and (9) arise from requiring orthogonality with respect to the two terms $\mathcal{T}_1y$ and $\mathcal{T}_2\bar{d}$, separately. A detailed derivation can be found in [10], where generalization to correlated symbol sequences is also discussed.

The coupled equations (8),(9) can be solved in precisely the same way as the corresponding equations (3.3a,b) in [8]. Convert them to systems of linear equations in the coefficients. Then, a new system of linear equations is created by combining all equations with known left-hand sides. This system is solved with respect to the coefficients of S_1 and L_{1*}. Subsequently, Q is obtained from (8). The polynomial degrees are

$$\deg S_1 = \deg L_1 = \ell - k$$

$$\deg Q = \deg L_2 = \max\{\deg H, \deg\beta\} - 1 \ .$$

Solution of the spectral factorization (7) is straightforward. With a given right-hand side, it is just an ordinary polynomial (FIR) spectral factorization, for which there exist efficient iterative algorithms. The averaged factors in (7) can be evaluated as follows. For a stochastic polynomial $\Delta P(q^{-1})$, of degree δp, let the Hermitian parameter covariance matrix be

$$\mathbf{P}_{\Delta P} = \begin{bmatrix} \bar{E}|\Delta p_0|^2 & \dots & \bar{E}(\Delta p_0\Delta p_{\delta p}^*) \\ \vdots & \ddots & \vdots \\ \bar{E}(\Delta p_{\delta p}\Delta p_0^*) & \dots & \bar{E}|\Delta p_{\delta p}|^2 \end{bmatrix}$$

Denote the sum of the diagonal elements g_0, the sum of the elements in the i'th super–diagonal g_i, the sum of elements in the i'th subdiagonal g_{-i}. Note that $g_{-i} = g_i^*$. Then it becomes evident, by direct multiplication of $\Delta P(q^{-1})\Delta P_*(q)$, and taking expectations, that

$$\bar{E}(\Delta P\Delta P_*) = g_{dp}^*q^{-dp} + \ldots + g_1^*q^{-1} + g_0 + g_1q + \ldots + g_{dp}q^{dp}$$

Above, $dp \leq \delta p$, with $dp = 0$ for uncorrelated coeff.

Note that, apart from the nominal model and the variance ratios ρ and η, only second order moments of the model error distributions need to be known. If not known, A_1, N_1, and the covariance matrices of ΔB and ΔM can still be used as "robustness tuning knobs". In the case of no model uncertainty, we set $\Delta B = \Delta M = 0, A_1 = N_1 = 1$ in (6)–(10). An increase of the covariance matrix elements of ΔB or ΔM will result in more cautious feedback and feedforward filters, with lower gains and lower, broader, spectral peaks.

An increase of η reduces the gain of the feedback filter Q/P. This results in shorter, but more numerous, bursts of decision errors. (See section 5 below.)

4. THE CLASS OF EQUALIZERS

The equations (3)–(10) define a class of robust equalizers, with linear equalizers and DFE's as special cases:

- If $\eta = 0$ (perfect decisions assumed), and with no model uncertainty, the *IIR DFE* discussed in [8] is obtained [4]. In this (and only this) case, a solution of the spectral factorization (7) is not required. We directly obtain $\beta = A_0M_0$.
- If $\eta \to \infty$ (decisioned data are very unreliable), $\| Q(q^{-1}) \| \to 0$. Then, (7) and (9) reduce to the design equations for a ***robust linear equalizer*** S/R, derived in [9]. (Divide (7)–(9) by η and set $r \triangleq r_1/\eta$, which is finite.)
- When $\eta \to \infty$ and no model uncertainty is assumed, we obtain the ordinary linear recursive equalizer, discussed in e.g. [1],[8].

[4] With $\Delta B = \Delta M = 0, A_1 = N_1 = 1$, we get $\tilde{B}\tilde{B}_* = B_0B_{0*}$, $\tilde{M}\tilde{M}_* = M_0M_{0*}$, $r_1 = \rho$, $\beta = A_0M_0 \triangleq \gamma$ and $H = B_0N_0 \triangleq \tau$.

5. PERFORMANCE AND ROBUSTNESS

We are presently evaluating the performance of robust equalizers. The properties of a large number of linear equalizers ($\eta \to \infty$), based on randomly selected 5–tap FIR channel models, are summarized in Section 4.5 of [9]. In about 1/5 of those cases, the nominal channels had pronounced nulls. Equalizers designed without taking any model uncertainty into account then had spectral peaks. Their performance was very sensitive to model errors. Robust design eliminated the sensitivity, at the price of only a slight increase of the minimal MSE (for correct models).

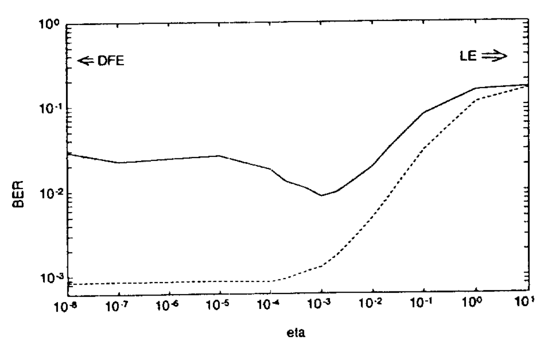
Example. The utility of the parameter η is investigated. Based on a model (without uncertainty)

$$y(n) = \frac{B_0}{A_0}d(n) + v(n) \;=\; \frac{(1+0.95q^{-1})}{(1-0.70q^{-1})^3}d(n) + v(t)$$

we designed DFE's for the signal to noise ratio 28dB $\Leftrightarrow$ $\rho = 0.553$. The smoothing lag was $\ell = 5$ and the data sequence was binary PAM ($d(n) \in [-1, 1]$). DFE's were designed for different values of the parameter η. Here, $H = B_0, \tilde{B}\tilde{B}_* = B_0B_{0*}, \tilde{M}\tilde{M}_* = 1$, and (7) reduces to

$$r_1\beta\beta_* = \eta B_0B_{0*} + (1+\eta)\rho A_0A_{0*} \ .$$

The bit error rate (BER) was estimated, from runs with 500000 symbols for each design. The diagram below shows the BER as a function of η (solid). Also shown is the BER without error propagation (dashed), i.e. the result we would obtain if correct past decisions could be substituted for $\bar{d}(n-\ell-1)$ in (3). In this difficult problem, a pure DFE has severe problems with error bursts, while a linear equalizer (LE) gives a large BER. However, the BER can be reduced by a factor of 3.5 by using $\eta = 0.001$ instead of $\eta = 0$ (pure DFE).



Furthermore, it is evident from the table below, that the use of a small $\eta \geq 0$ reduces the length of error propagation events[5] substantially. If a coding scheme is used which is sensitive to long consecutive sequences of errors, this property is valueable in itself.

Length:	Number of error propagation events 1-2	3-20	21-50	51-100	>100
$\eta = 0$:	328	197	75	66	133
$\eta = 10^{-3}$:	323	232	77	24	–
$\eta = 0.1$:	6100	4469	171	–	–

[5]Defined in this example as the length of sequences of $\bar{d}(n)$ which contain < 7 consecutive correct decisions.

References

[1] A. Ahlén and M. Sternad, "Optimal deconvolution based on polynomial methods". *IEEE Trans. Acoustics, Speech and Signal Processing*, vol 37, pp 217–226, 1989.

[2] A. Ahlén and M. Sternad, "Wiener filter design using polynomial equations". *IEEE Trans. Signal Processing*, vol 39, pp 2387–2399, 1991.

[3] R. C. Chung and P. R. Bélanger, "Minimum–sensitivity filter for linear time–invariant stochastic systems with uncertain parameters". *IEEE Trans. Aut. Control*, vol 21, pp 98–100, 1976.

[4] G. C. Goodwin and M. E. Salgado, "A stochastic embedding approach for quantifying uncertainty in the estimation of restricted complexity models". *International Journal of Adaptive Control and Signal Processing*, vol 3, pp 333–356, 1989.

[5] G. C. Goodwin, M. Gevers and B. Ninnes, "Quantifying the error in estimated transfer functions with application to model order selection". *IEEE Trans. Aut. Control*, vol 37, pp. 913–928, 1992.

[6] Y. C. Jenq, "A class of generalized decision feedback equalizers". *Journal of the Franklin Institute*, vol 306, pp 513–522, 1979.

[7] L. Lindbom, "Simplified Kalman estimation of fading mobile radio channels: high performance at LMS computational load." *Proc. ICASSP'93*, Minneapolis, April 1993.

[8] M. Sternad and A. Ahlén, "The structure and design of realizable decision feedback equalizers for IIR channels with colored noise". *IEEE Trans. Information Theory*, vol 36, pp 848–858, 1990.

[9] M. Sternad and A. Ahlén, "Robust filtering and feedforward control based on probabilistic descriptions of model errors". *Automatica*, May 1993.

[10] M. Sternad, A. Ahlén and E. Lindskog, "Robust decision feedback equalization". In preparation.