Optimal Linear Filtering over a Galois Field: Equalization and Prediction

Denis Fantinato[1], Daniel Guerreiro e Silva[1], Romis Attux[1], Rafael Ferrari[2], Leonardo Tomazeli Duarte[3], Ricardo Suyama[4], Jugurta Montalvão Filho[5], Aline de Oliveira Neves[4] e João Marcos Travassos Romano[2]

1 – DCA/FEEC – University of Campinas (UNICAMP)
2 – DMO/FEEC – University of Campinas (UNICAMP)
3 – FCA – University of Campinas (UNICAMP)
4 – UFABC – Federal University of ABC
5 - UFS – Federal University of Sergipe

{denisgf,danielgs,attux}@dca.fee.unicamp.br,
rferrari@decom.fee.unicamp.br, leonardo.duarte@fca.unicamp.br,
{ricardo.suyama,aline.neves}@ufabc.edu.br, jmontalvao@ufs.br,romano@dmo.fee.unicamp.br

Abstract – The problem of optimal linear filtering and prediction has been so far typically formulated and studied in the context of real- or complex-valued signals. In this article, we provide an extension of this problem to the framework of finite (Galois) fields. Simulation results encompassing supervised and unsupervised prediction-based equalization are presented for a number of scenarios based on *GF(2)*.

Keywords – *Equalization, prediction, Galois fields, information-theoretic learning.*

"Le savant n'étudie pas la nature parce que cela est utile; il l'étudie parce qu'il y prend plaisir et il y prend plaisir parce qu'elle est belle. Si la nature n'était pas belle, elle ne vaudrait pas la peine d'être connue, la vie ne vaudrait pas la peine d'être vécue"

Henri Poincaré, "Science et méthode"

1. Introduction

The problems of optimal filtering, in a generic context, and, more specifically, of equalization and prediction are regarded as fundamental topics within the field of signal processing. These subjects have been deeply analyzed from several standpoints in the last decades, but, beyond any doubt, it remains imperative to highlight the analytical framework developed in the 1940s by Wiener and Kolmogorov [1][2], which encompasses both discrete- and continuous-time stationary stochastic signals.

In simple terms, this framework may be understood, in the single-input single-output (SISO) discrete-time case, as providing the necessary means to find the optimal parameters of a filter that attempts to approximate a desired signal *d(n)* from a set of samples of the input signal *x(n)*. These signals are assumed to be real-valued and, when the filtering structure is assumed to be a finite impulse response (FIR) filter, the ensuing optimization task can be solved in closed form, thereby yielding the so-called *Wiener solution* [2].

The two aforementioned particular problems, those of equalization (or deconvolution) and prediction, can be straightforwardly formulated within this framework. In the case of equalization, the transmitted signal *s(n)* suffers a distortion, e.g., by means of a finite impulse response (FIR) channel *h* with *L* real-valued coefficients, thus giving rise to the signal *x(n)*. This signal – *x(n)* – is expressed as the convolution between the channel impulse response and *s(n)*. Hence, from a filtering standpoint, the objective is to recover, with maximum precision, the original signal *s(n)* from the samples of *x(n)*. Usually, in a supervised approach, this is achieved using the structure presented in Figure 1 [2].

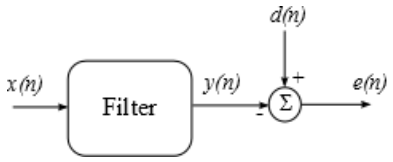

Figure 1. Filtering Setup - Equalization

In Figure 1, *y(n)* is the filter output signal i.e. an estimate of *s(n)*, *d(n)* is the desired signal – the role of which will be played by *s(n)* in this case[1] – and *e(n)* is the error signal, generated as:

$$e(n) = d(n) - y(n). \quad (1)$$

For the *prediction* problem, a similar line of reasoning is pursued. Now, the objective is to estimate from *x(n)* and its delayed versions the value of a future signal sample, e.g., *x(n+1)*. Such process can be defined within the Wiener filtering framework in terms of the structure illustrated in Figure 2.

[1] *A more rigorous definition requires that the desired signal be s(n) or a delayed version thereof, but the issue of choosing the equalization delay will not be dealt with here.*

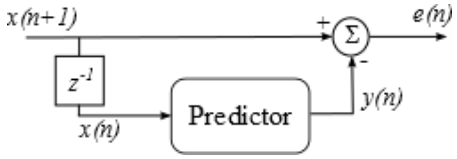

Figure 2. Prediction Setup

In this case, *y(n)* is the predictor output – an estimate of x*(n+1)* generated by a linear combination of delayed samples, and Eq. (1) holds for *d(n) = x(n+1)*.

Classically, the equalization and prediction structures are considered to be linear filters. The most usual option is for FIR structures, but, in some cases, the use of infinite impulse response (IIR) devices is either attractive or necessary [3]. For the more general IIR case, the filter output can be written as:

$$y(n) = \sum_{j=0}^{M-1} a_j x(n-j) + \sum_{k=1}^{N} b_k y(n-k) \qquad (2)$$

where M is the number of coefficients a_j in the feedforward part and N is the number of coefficients b_k in the feedback part. To obtain the input-output relationship of a FIR filter, it suffices to consider all coefficients of the feedback part as being null.

Within the Wiener-Kolmogorov theory, the optimal parameters a_j and b_k are chosen so as to minimize the second moment of the error signal *e(n)*, which is also referred to as mean-squared error (MSE). As mentioned earlier, for the FIR case, the optimal solution can be obtained in closed form. This is done by solving a linear system of equations that usually bears the names of Norbert Wiener and Eberhard Hopf [3].

These and other strategies that form the core of modern optimal filtering theory were developed under the aegis of the assumption that all signal samples and system parameters are real or complex numbers. This certainly is enough for many practical applications, but there remained a clear gap insofar as extensions to inherently discrete sources and distortions are concerned. Interest in sources of this kind can be justified in theoretical terms – e.g. in terms of building a more complete view of the potentialities inherent to the very idea of filtering and of establishing connections with the vast *corpus* of coding theory – and also in practical terms – in view, for instance, of the existence of enormous binary and genomic databases, as well as of inherently discrete problems related to bio-inspired information systems based, for instance, on the use of molecules [4].

Yeredor, Gutch, Gruber and Theis [5] took the initial steps in the direction of extending key signal processing methods to encompass the peculiarities of entities defined in accordance with the formalism of finite fields. Their efforts were essential to establishing the bases of a blind source separation (BSS) theory applicable to Galois fields. They also proposed algorithms for solving the BSS problem via ICA, being this repertoire later enlarged by the contribution of Silva et al. [6]. These contributions, nonetheless, were restricted to spatial filtering, i.e., instantaneous mixtures of sources.

Having this in view, we propose, in this work, to formulate the optimal filtering problem – in the particular case of equalization and prediction – in the context of finite fields, attempting thus to deal with the temporal aspects (convolutive mixtures) of this novel branch of information processing. The proposed framework will be tested considering *GF(2)* for several distinct scenarios, and its performance will be evaluated establishing, whenever possible, parallels with the canonical theory for real-valued variables.

The rest of this article is organized as follows: in Section 2, we describe the problems of performing equalization and prediction over finite fields; in Section 3, the evaluation metrics used, together with an exhaustive search, to find the optimal equalizers and predictors are presented; in Section 4, we present the results of simulations and the corresponding performance analysis; finally, in Section 5, we discuss the main results, some key issues and future perspectives.

2. Finite Fields: Equalization and Prediction

Although the efficient theoretical approach to the problem of ICA over Galois fields established by Gutch et al. [5] encourages extensions to temporal filtering, the classical framework outlined in Section I is not suitable for handling finite fields. Nevertheless, we will present in the following possible ways to circumvent these difficulties, thus establishing viable formulations of linear equalization and prediction over a Galois field.

Considering *d(n)* as a sequence of symbols associated with a finite field that, in this case, will be *GF(2)*, it is possible to consider it as a sequence of bits i.e. a Boolean sequence. Under the canonical GF(2) operations, Eq. (1) can be rewritten as:

$$e(n) = d(n) \oplus y(n) \qquad (3)$$

where $\{\oplus\}$ is the modulo-2 sum (equivalent to the logical operation XOR). Analogously, the general linear filter described in Eq. (2) can be reformulated in terms of the *GF(2)* sum and product operations:

$$y(n) = \sum_{j=0}^{M-1} a_k \cdot x(n-j) \oplus \sum_{k=1}^{N} b_k \cdot y(n-k) \qquad (4)$$

where the operator $\{\cdot\}$ corresponds to the product (which leads to the same results, for "0" and "1", obtained with the "conventional" product). Also, it is important to emphasize that the summation symbols correspond to the modulo-2 operation and that *x(n)*, a_k, b_k and *y(n)* assume binary values ("0" and "1").

Similar considerations are valid for the distortion caused by channel.

Having thus defined the general linear filtering structure and the error signal for *GF(2)*, we are ready to address the two problems of interest, equalization and prediction. The remaining steps are to build an evaluation metric based on the error signal and to propose a suitable optimization approach.

3. Evaluation Metrics

Whenever one deals with finite fields, all signal samples can be modeled in terms of a discrete probability distribution, the parameters of which can be estimated using a frequency-based methodology. In our case, the samples of the signal *e(n)* can be considered as a discrete random variable (RV) with two possible values, "0" and "1". Hence, the probability of *e(n)* being equal to zero can be estimated as:

$$P_e(0) = \frac{N_0(\{e(n)\})}{n} \qquad (5)$$

where $N_0(\cdot)$ is the number of times in which *e(n)* is equal to zero for a total number of samples equal to *n*. As the aim of optimal filtering is, intuitively, to reach an error signal as close to zero as possible, one may consider $P_e(0)$ as an evaluation metric to be maximized.

This metric is reasonable if an ideal solution is reachable, but, if this is not the case, it can be misleading. For instance, if one deals with blind equalization based on linear prediction, the error signal will never be a sequence of zeros (see Fig. 2); instead, it will ideally be equal to *s(n)* [7]. Consequently, if the probability that *s(n)* be equal to zero is, say, 0.3, by maximizing the metric shown in Eq. (5), one will not actually be able to equalize the channel, as a signal with residual intersymbol interference will have a frequency of zeros closer to 0.5 (i.e. higher) in view of the "central limit theorem" for *GF(2)* mixtures [5].

To overcome this issue, we resort to an information-theoretic criterion – the minimization of the entropy of the error signal:

$$H(e(n)) = -\sum_{i=0}^{1} P_e(i)\log_2(P_e(i)) \qquad (6)$$

With these two evaluation metrics, it is possible to test these ideas for different kinds of scenarios.

4. Results

First we will consider the problem of supervised equalization over *GF(2)*. A source *s(n)* is generated with independent and identically distributed (i.i.d.) binary samples, with $P_s(0) = p$ and $P_s(1) = 1 - p$. The signal is submitted to the effect of a FIR channel *h* with *L* taps. The resulting signal *x(n)* can be equalized with a FIR or an IIR filter.

Let us first consider a FIR equalizer, which obeys Eq. (4) with $b_k = 0$ for all *k*. Unfortunately, an equalizer of this kind is severely limited as the combination terms of *x(n)* always lead to a residual intersymbol interference element that cannot be eliminated. For example, for a channel $h = 1 + z^{-1}$, or in vector notation, $h = [1\ 1]$, a two-tap FIR filter can use the samples $x(n) = s(n) \oplus s(n-1)$ and $x(n-1) = s(n-1) \oplus s(n-2)$. If the equalizer selects only either *x(n)* or *x(n-1)*, there will be clearly no equalization. The same is valid if the equalizer combines both terms: *s(n-2)* will remain, causing a non-negligible degree of error. Indeed, for real-valued FIR equalizers, the residual interference also exists, although in tolerable values. Thus, satisfactory equalization over a Galois Field strongly demands general (IIR) filters.

Given a probability *p* and a channel *h*, all possible solutions for an IIR filter of sufficient order were tested and the best according to the maximum probability error metric, estimated with Eq. (5), were chosen. The results are presented in Table 1.

	Channel *h*		Optimum Filter		error
p	a	b	a	b	(%)
0.5	[1 0 0 1 1]	-	[1]	[0 0 1 1]	0
0.2	[1 1 0 0 1]	-	[1]	[1 0 0 1]	0
0.9	[1 1 0 1 1]	-	[1]	[1 0 1 1]	0
0.1	[1 1 0 1 1]	-	[1]	[1 0 1 1]	0

Table 1. Results - Supervised Equalization

Remembering that the first coefficient b_1 is a logical operation over *y(n-1)*, it is possible to see that the channel inverse was always obtained, which allowed a perfect estimate of the transmitted signal to be reached. Notice that the performance, as expected, does not depend on the probability *p*.

In order to analyze the soundness of the metric estimator shown in Eq. (5) for a perfect inversion case – the last one described in Table 1 – we present, in Fig. 3, the value of the relative frequency of zeros in the error signal for all possible equalizer denominator coefficients (ranging from [0 0 0 1] to [1 1 1 1]). For the sake of simplicity, the four-bit combination is shown in terms of the equivalent decimal. It is clear that there is a single optimal solution (#11 – [1 0 1 1]), which is exactly the channel inverse.

For the prediction problem, simulation results are presented in Table 2. In this case, we perform the prediction in order to use the prediction-error filter in the role of blind equalizer [7]. In this case, ideally, the prediction error should be identical to *s(n)*. In Table 2, error (%) is relative frequency of ones in *e(n)* while error PEF(%) refers to the relative frequency of errors in the comparison between *e(n)* and *s(n)*. Notice that, in all cases, error (%) is approximately equal to $1 - p$ and PEF(%) is null, which reveals that the prediction task was successfully fulfilled. In this case,

as discussed in the end of Section 3, we had to use the entropy of the error signal in the role of filtering criterion. The associated cost function for all 12 possible combinations (numerator ranging from [0 1] to [1 1] and denominator ranging from [0 0] to [1 1]) associated with the third case presented in Table 2 is shown in Fig. 4. Notice that the minimum value, as expected, is very close to the entropy for $p = 0.1$, which is 0.469 bits.

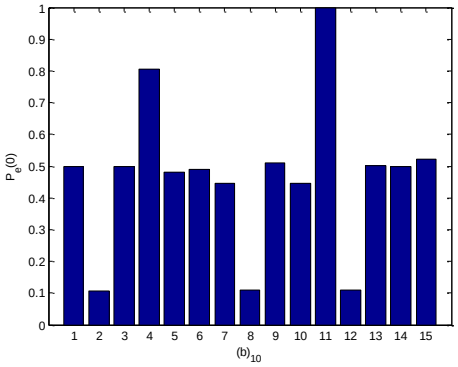

Figure 3. Cost Function – Supervised Equalization

	Channel *h*		Optimum Filter		error	error
p	a	b	a	b	(%)	PEF(%)
0.8	[1 1]	-	[1]	[1]	0.189	0
0.2	[1 1 1]	-	[1 0 1]	[0 0 1]	0.823	0
0.1	[1 0 1]	-	[0 1]	[0 1]	0.899	0
0.4	[1 0 1 1]	-	[0 1 1]	[0 1 1]	0.620	0

Table 2. Results - Prediction

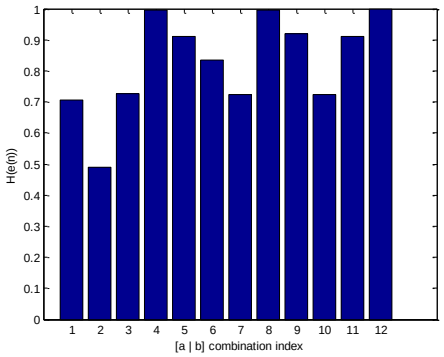

Figure 4. Cost Function - Prediction

5. Conclusions

In this work, we proposed an extension to the problem of optimal linear filtering – specifically in the contexts of equalization and prediction – to signals and systems over Galois fields. The obtained results show the viability of this extension for a number of scenarios, thus encouraging further research efforts, which may include (i) a deeper analysis on the evaluation metrics and their validity, (ii) building general (ARMA) forecasting models for binary and genomic data and (iii) building extensions of the Karhunen-Loève transform. It would be also interesting to consider, for noisy and underdetermined cases, possible connections with coding theory.

Acknowledgements

The authors thank CNPq for the financial support.

References

[1] N. Wiener, *Extrapolation, Interpolation, and Smoothing of Stationary Time Series*, MIT Press (Cambridge, MA), 1949.

[2] J. M. T. Romano, R. Attux, C. C. Cavalcante, R. Suyama, *Unsupervised Signal Processing: Channel Equalization and Source Separation*, CRC Press, 2010.

[3] S. Haykin, *Adaptive Filter Theory*, Prentice Hall, 1996.

[4] L. M. Adleman, *Molecular computation of solutions to combinatorial problems*, Science, vol. 266, pp. 1021-1024, Nov. 1994.

[5] H. W. Gutch, P. Gruber, A. Yeredor, F. J. Theis, *ICA over Finite Fields – Separability and Algorithms*, Signal Processing, vol. 92, no. 8, pp. 1796–1808, 2012.

[6] D. G. e Silva, R. Attux, E.Z. Nadalin, L.T. Duarte, R. Suyama, *An Immune-Inspired Information-Theoretic Approach to the Problem of ICA over a Galois Field*, IEEE Information Theory Workshop (ITW), pp. 618-622, Oct. 2011.

[7] R. Ferrari, C. M. Panazio, R. Attux, C. C. Cavalcante, L. N. de Castro, F. J. Von Zuben, J. M. T. Romano, *Unsupervised Channel Equalization Using Prediction-Error Filters*, IEEE Workshop on Neural Networks for Signal Processing, pp. 869-878, 2003.