Optimal Time-Series Prediction over Galois Fields

Denis G. Fantinato[1], Daniel G. e Silva[1], Romis Attux[1], Rafael Ferrari[1], Leonardo T. Duarte[1], Ricardo Suyama[2], Aline Neves[2], Jugurta Montalvão Filho[3] and João M. T. Romano[1]

1 - Laboratory of Signal Processing for Communications (DSPCom) – FEEC / UNICAMP
2 - Federal University of ABC - UFABC
3 - Federal University of Sergipe - UFS

{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 prediction was developed predominantly under the aegis of the assumption that all signal samples and system parameters should be real or complex numbers. In this work, we provide an extension of this problem to the framework of finite (Galois) fields through use of the classical and elegant framework defined by stochastic models - AR, MA and ARMA. Simulation results in the context of prediction-based equalization are presented for a number of scenarios based on *GF*(*2*).
Keywords: *Prediction, Galois fields, information-theoretic learning.*

1. Introduction

The origins of adaptive filtering theory can be traced to the seminal contributions of Norbert Wiener and Andrey Kolmogorov, formalized in the first half of the last century [1][2]. From these contributions, it became possible to define in clear terms what is meant by optimal filter in the context of stationary information signals. In the last four decades, however, a new research branch sprang, entitled as unsupervised signal processing, being characterized, in simple terms, by an approach to identification and inverse problems based essentially on the explicit or implicit use of higher-order statistics.

In view of the nature of the problems that were tackled using filtering methods, the field was developed predominantly under the aegis of the assumption that all signal samples and system parameters should be real or complex numbers. This certainly suffices for many practical applications, but, as is indicated by the parallel development of coding theory, there remained a clear gap insofar as extensions to inherently discrete sources and distortion models are concerned. Presently, interest in sources of this type 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 also in practical terms – notice, for instance, the existence of enormous binary and genomic databases, as well as of inherently discrete problems related to bio-inspired computing [3].

Yeredor, Gutch, Gruber and Theis [4] took the initial steps in the direction of extending unsupervised signal methods to encompass the peculiarities of the formalism of finite or Galois fields, being this repertoire later enlarged by the contribution of Silva et al. [5]. These works, nonetheless, were restricted to spatial filtering. In view of this fact, the objective of this work is to formulate the optimal filtering problem – focusing on the task of prediction, in the context of finite fields, thus emphasizing the temporal aspects of this novel branch of information processing.

2. Optimal Prediction: Predictors and Prediction-Error Filters

The prediction problem is, in simple terms, that of estimating, from a signal $x(n)$ and delayed versions thereof, the value of a future signal sample, like $x(n+1)$. Such process can be defined within the optimal filtering framework in terms of the design of a suitable predictor, as illustrated in Fig. 1.

The predictor output is $y(n)$ – an estimate of $x(n+1)$ generated as the outcome of a mapping applied to the set of delayed samples. The need for estimating the future values of the time series in question can be important *per se*, but, in this work, it will be a path towards performing the deconvolution of a signal of interest. This is an essential step whenever information signals are corrupted by a process of superposition of time samples, which can be translated as an operation of convolution between these signals and a linear time-invariant system. The key to performing blind deconvolution is to use a prediction error filter (PEF) in the role of inverse filter or blind equalizer. Such approach is based on the property that a linear predictor designed in accordance with the second-order Wiener / Kolmogorov framework yields an output error $e(n)$ that can be significantly "white".

Assuming that the signal of interest $s(n)$ is a transmitted signal composed of independent and identically distributed (i.i.d.) - hence uncorrelated / white - samples, it is intuitive to expect that a white prediction error signal $e(n)$ – see Fig. 1 - might be a reliable estimate of $s(n)$.

In view of the difficulties arising from the need for adapting blind equalization paradigms to inverse problems defined over Galois fields, the idea of dealing with optimal prediction become very attractive. This point, which was raised in a preliminary work written by us [6], will be further investigated later on.

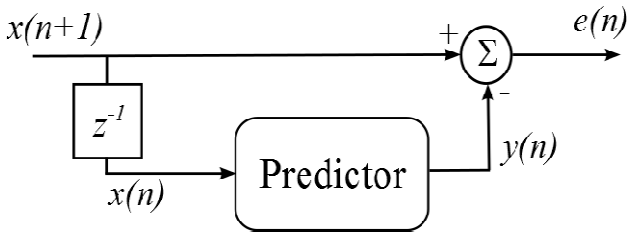

Figure 1. Prediction Setup

3. Prediction Models over Galois Fields

A proper analysis of the ideas outlined in the previous section will require from us the adaptation of classical prediction models to the idiosyncrasies of Galois fields. Finite fields with a number q of elements are commonly denoted as $GF(q)$. An illustrative example of Galois field is $GF(2)$: the set of numbers {0,1} with the "usual" product operation and the exclusive-OR (XOR) operation, in analogy to conventional summation.

Mathematically, it is perfectly possible to develop system models in harmony with the features of inherently finite / discrete sources and the operations selected to form the field. In view of this fact, it becomes possible, for instance, to define an optimal filter for $GF(q)$. Convolution can be duly adapted and the error signal can be defined, without significant modifications, as the difference between $d(n)$ and $y(n)$ within the field:

$$e(n) = d(n) \oplus \bar{y}(n) \tag{1}$$

where $\bar{y}(n)$ is the inverse element of $y(n)$ with respect to the sum i.e. $y(n) \oplus \bar{y}(n) = 0$. To distinguish the operations, we shall henceforth use the symbol $\oplus$ to denote addition over the finite field, while the traditional symbol '+' will denote the usual sum.

Proceeding with the extension procedure, to describe the filter structure, we will appeal to a classical and graceful approach of Stochastic Models. The representation of a stochastic process by a model was first performed by Yule, in 1927. There are three kinds of models that lead to a stochastic stationary signal: the autoregressive (AR), the moving-average (MA) and the mixed autoregressive-moving-average (ARMA). The three models are the filter structures that we shall analyze in the context of finite fields. Then it is necessary to define a criterion to determine the coefficients of the filter, in an optimal sense.

The adoption of Wiener's criterion cannot be carried on to filtering over $GF(q)$ because the product between the variable realization (in $GF(q)$) and the associated probability (in **R**) is not established. To replace this absence, a different criterion must be chosen, thus we propose the minimization of the entropy of the error:

$$H(e) = -\sum_{i \in GF(q)} p_e(i) \log[p_e(i)] \tag{2}$$

The maximum entropy distribution of a random variable over $GF(q)$ is the uniform distribution. There is a property, analogous to the central limit theorem in **R**, that states that the sum of independent random variables over $GF(q)$ results in a distribution that is "more uniform"[4]. If we assume that $s(n)$ is i.i.d. and non-uniform we can infer that the convolutional sum effect of the channel leads to a "uniformization". Then we conjecture that a filter that obtains a minimal entropy configuration in its output is capable to remove the channel effect and recover the original signal $s(n)$.

4. Results

Considering initially the models of a stochastic process, i.e., the AR, MA and ARMA models, we first generate a source $s(n)$ i.i.d. over $GF(2)$ and apply to it a linear combination or convolution, according to the models. We wish to recover the best estimate of $s(n)$ at the output of the stochastic model. Indeed, it is always possible to completely recover the original source. Since the solutions possibilities also belong to a finite field, we can exhaustively search for the best one. The resulting coefficients of the model that minimizes the entropy are given in Tab. 1, in which we use a vector notation for the Z-transform [6].

Table 1. Results – Stochastic Models.

Stochastic Model	Convolution	Coefficients		Minimum Entropy [bits]
		w	b	
AR	$\frac{1}{[1101]}$	[1010]	--	0.7163
	$\frac{1}{[1110]}$	[1100]	--	0.7419
MA	[1101]	--	[1010]	0.8713
	[1110]	--	[1100]	0.8661
ARMA	$\frac{[1011]}{[1110]}$	[110]	[011]	0.9044

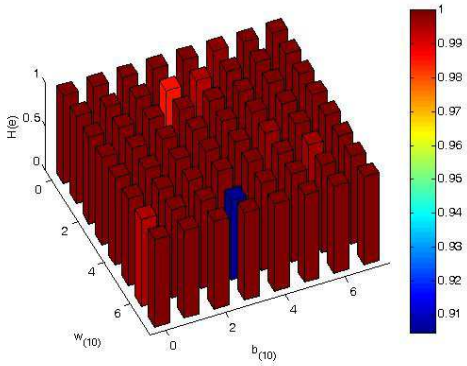

Figure 2. Entropy – ARMA model.

The minimum entropy indicated in Tab. 1 is related to the entropy of the model output. Since the source is totally recovered, the respective value is the entropy of $s(n)$, which was chosen to be different from the worst case, i.e., with maximum entropy.

In order to analyze the soundness of the metric estimator, we take the ARMA case – last one described in Tab. 1 - and present, in Fig. 2, the value of the measured entropy of the output signal for all possible model coefficients ($\boldsymbol{w}$ and $\boldsymbol{b}$ ranging from [0 0 0] to [1 1 1]). For the sake of simplicity, the three-bit combination is shown in terms of the equivalent decimal. It is clear that there is a single optimal solution ($\boldsymbol{w}$ = #6 – [1 1 0] and $\boldsymbol{b}$ = #3 – [0 1 1]).

5. Conclusion

In this work, we proposed an extension to the problem of optimal prediction to signals and systems over Galois fields through use of the classical AR, MA and ARMA models. The obtained results show the viability of obtaining the zero forcing condition in all cases, which motivate us for further research efforts, including the extension to the problem of convolutive mixtures. It would be also interesting to consider, for noisy and underdetermined cases, possible connections with coding theory.

References

[1] S. Haykin, *Adaptive Filter Theory*. Prentice- Hall, 1996, 3rd ed.

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

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

[4] 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.

[5] 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.

[6] D. Fantinato, D. G. e Silva, R. Attux, R. Ferrari, L. T. Duarte, R. Suyama, J. Montalvão Filho, Aline O. Neves, J. M. T. Romano, *Optimal Linear Filtering over a Galois Field: Equalization and Prediction*, Encontro dos Alunos e Docentes do DCA (EADCA), Apr. 2012.