

# Asymptotic Optimality of Finite-Memory Detection

Gianluigi Ferrari, Giulio Colavolpe, and Riccardo Raheli

Dipartimento di Ingegneria dell'Informazione

Università di Parma

Parma, I-43100, Italy

e-mail: gianluigi.ferrari@unipr.it

**Abstract** — The subject of this paper is the asymptotic optimality of finite-memory detection for transmission over a channel characterized by a single multiplicative time-invariant stochastic parameter (e.g., block frequency non-selective fading). It is known that any finite-memory detection algorithm, either trellis-based or graph-based, is characterized by a single basic metric. We present a theorem which proves that this metric tends, asymptotically, to that of a receiver with perfect channel state information.

## I. INTRODUCTION AND PRELIMINARIES

In [1], a general framework for finite-memory detection is proposed. In particular, it is shown that statistical definitions of *causality* and *finite memory* conditions lead to formulations of Viterbi algorithm (VA), forward-backward (FB) and sum-product (SP) algorithms, based on a single basic metric.

Assume that the modulator/encoder can be modeled as a time-invariant finite state machine (FSM) with state  $\mu_k$  which evolves upon the arrival of a sequence of independent information symbols  $\{a_k\}$ , where  $k$  denotes the discrete-time epoch. This FSM can be characterized by a “next-state” function  $\text{ns}(\mu_k, a_k) = \mu_{k+1}$  and an “output” function  $\text{o}(\mu_k, a_k) = c_k$ . For simplicity, we refer to a discrete-time equivalent communication system, and we further consider a single discrete-time observation  $r_k$  in correspondence with each information symbol  $a_k$ —the case of multiple observations per information symbol can be dealt with by using an appropriate vector notation. The basic metric to be used in a finite-memory detection algorithm can be written as follows [1]:

$$\lambda_k(T_k) \triangleq \ln \left[ p(r_k | \mathbf{r}_{k-C}^{k-1}, T_k) P\{a_k\} \right] \quad (1)$$

where  $T_k \triangleq (\mu_{k-C}, \mathbf{a}_{k-C}^k)$  and the parameter  $C$  is defined as *finite-memory parameter*. In other words, the basic metric is given by the logarithm of the product of the conditional probability density function (pdf) of  $r_k$  given  $C$  previous observations  $\mathbf{r}_{k-C}^{k-1}$  with the a priori probability of symbol  $a_k$ .

This unified formulation is conceptually intriguing, since it suggests that *any* detection strategy appeared in the literature for a VA can be *systematically* extended to FB and SP algorithms, and vice versa. Finite-memory detection, based on the basic metric (1), is optimal only in the case of channels with strictly finite memory. In any other case, this approach is suboptimal for any given *finite* value of the parameter  $C$ —a very important example is given by a communication system with transmission over a *stochastic channel*.

In this paper, we investigate the *asymptotic optimality* of finite-memory detection based on the metric (1). In particular, we consider the case of transmission of a linearly modulated signal over a channel characterized by a single time-invariant stochastic parameter  $\xi$ .

## II. ASYMPTOTIC OPTIMALITY

**Theorem** Assume that the observable can be expressed as

$$r_k = \xi c_k + w_k \quad (2)$$

where  $\xi$  is an arbitrary random variable,  $\{c_k\}$  are transmitted (and possibly coded) symbols and  $\{w_k\}$  are independent and identically distributed noise samples with known statistics. The performance of the proposed finite-memory detection strategy, based on the metric (1), is asymptotically optimal, for  $C \rightarrow \infty$ , in the sense that it approaches the performance of an ideal receiver which perfectly knows the channel, i.e., the realization of  $\xi$ .

The proof is based on showing that

$$\lim_{C \rightarrow \infty} p(r_k | \mathbf{r}_{k-C}^{k-1}, \mathbf{a}_{k-C}^k, \mu_{k-C}) = p_w(r_k - \eta_\xi^{(\infty)} c_k) \quad (3)$$

where  $p_w(\cdot)$  is the pdf of a noise sample,  $\eta_\xi^{(\infty)} \triangleq \lim_{C \rightarrow \infty} \eta_\xi^{(C)}$ , and  $\eta_\xi^{(C)} \triangleq E\{\xi | \mathbf{r}_{k-C}^{k-1}, \mathbf{a}_{k-C}^k, \mu_{k-C}\}$ . Since one can show that  $\lim_{C \rightarrow \infty} \text{VAR}\{\xi | \mathbf{r}_{k-C}^{k-1}, \mathbf{a}_{k-C}^k, \mu_{k-C}\} = 0$ , it follows that  $p_w(r_k - \eta_\xi^{(\infty)} c_k)$  is equal to the pdf characterizing the basic metric of an ideally “coherent” receiver which perfectly knows the channel, i.e., it is asymptotically optimal.

According to the theorem, a finite-memory detection strategy is asymptotically optimal as long as  $\xi$  is a random variable, **regardless** of the particular distribution of  $\xi$  and of the particular (possibly coded) linear modulation. For example, this model includes a non-selective block fading channel, where the fading coefficient is time-invariant over the block. However, the theorem does not give any insight into the rate of *convergence* of the finite-memory detection strategy to the performance of a receiver with perfect channel state information. A special case of the general result in this theorem is considered in [2], where the convergence characteristics of the finite-memory detection strategy in the case of VA-based detection over an additive white Gaussian noise (AWGN) channel introducing a time-invariant channel phase rotation and PSK, were analyzed in detail.

The proof of the theorem suggests a detection strategy based on the following modified basic:

$$\tilde{\lambda}_k(T_k) \triangleq \ln \left[ p_w \left( r_k - c_k \frac{1}{C} \sum_{i=1}^C \frac{r_{k-i}}{c_{k-i}} \right) P\{a_k\} \right] \quad (4)$$

In the special case of white noise with zero mean (e.g., AWGN), it is possible to show that a receiver based on (4) is asymptotically optimal, **regardless** of the SNR. In particular, the lower the SNR, the slower the convergence, in terms of performance, of a receiver using (4) to the performance of a coherent receiver.

Although the asymptotic optimality of (4) is an interesting result from a theoretical viewpoint, simulation results (not reported here for lack of space) show that the convergence rate of a receiver using (4) is significantly lower than that of a receiver using (1), where statistical knowledge of the channel parameter is exploited.

## REFERENCES

- [1] G. Ferrari, G. Colavolpe, and R. Raheli, “Finite memory: optimality and reality,” in *Proc. Intern. Symp. on Turbo Codes & Relat. Topics*, Brest, France, September 2003, pp. 1–8.
- [2] G. Colavolpe and R. Raheli, “Theoretical analysis and performance limits of noncoherent sequence detection of coded PSK,” *IEEE Trans. Inform. Theory*, vol. 46, no. 4, pp. 1483–1494, July 2000.