On Reliable Communications over Channels Impaired by Bursty Impulse Noise

Dario Fertonani, *Member, IEEE,* and Giulio Colavolpe, *Member, IEEE*

Abstract—Digital communications over channels impaired by impulse noise are addressed. We adopt a two-state Markov model that allows to describe the typical bursty nature of the impulse noise, in contrast to the memoryless models generally considered in the literature. For this channel, we evaluate the achievable information rate and propose a couple of practical communication systems based on powerful codes and iterative receivers. Moreover, we discuss the effectiveness of the considered receivers in terms of performance/latency tradeoff as well as in terms of robustness to erroneous channel estimations. The proposed schemes are shown to perform fairly close to the theoretical limits, and significantly better than the conventional schemes employing memoryless detection.

Index Terms—Impulse noise, Markov channels, maximum-a-posteriori symbol detection, achievable information rate, low-density parity-check codes.

I. INTRODUCTION

THE power delivery networks and some mobile radio scenarios are often characterized by interference that exhibits a significant impulsive nature and cannot be properly described by the conventional additive white Gaussian noise (AWGN) model. Such phenomena, referred to as "impulse noise", are generally described by means of the Class-A model [1] or the Bernoulli-Gaussian model [2]. Since these models are memoryless, they cannot describe one of the main features of the actual channel, i.e., the occurrence of impulsive bursts (see [3] and references therein). Hence, given that the actual channel is characterized by a significant amount of memory, it is interesting to evaluate which performance gain can be achieved when the memory is exploited in the system design, in terms of ultimate theoretical limits as well as performance of practical systems. To address these issues, we follow the guidelines in [3] and consider a channel model that modifies the Bernoulli-Gaussian model [2] such that the channel state is, instead of a Bernoulli process, a two-state Markov process [4]. A two-state Markov process indeed provides a simple and effective way for describing a bursty evolution of the channel state [5]. The considered model, yet able to account for the actual channel memory, is still as manageable as the memoryless ones in [1], [2], so that we can derive an

algorithm for optimal maximum-a-posteriori (MAP) symbol detection.

We first carry out an information-theoretical analysis on the performance limits imposed by the channel, in terms of information rate [6] of systems employing linear modulations. Although the state process underlying the channel is the same as in the Gilbert-Elliott model (see [5] and references therein), whose information rate can be analytically computed [5], the same analytical arguments do not lead to a closed-form expression here, since the channel-output alphabet is not finite nor discrete [7]. Hence, we compute the information rate by means of the simulation-based method described in [8], exploiting the derived algorithm for MAP symbol detection. Such investigations definitely show that the ultimate performance limit improves as the memory of the impulse noise becomes more significant, motivating us to design practical schemes able to exploit the channel memory. Moreover, we compare the information rate achievable in conditions of ideal channel estimation with that achievable in conditions of mismatched decoding [9], that is, with errors in the channel estimation. Interestingly, the system results fairly robust to possible estimation errors.

Aimed at approaching the theoretical performance limits as close as possible, we focus on systems employing powerful channel codes, such as low-density parity-check (LDPC) codes or turbo-like codes [10], [11], and propose two receivers based on MAP detection. One receiver exploits the exchange of soft information between the MAP detector and the decoder, as in the turbo-equalization schemes [12], while the other receiver is simpler and does not perform iterative detection. We show that both receivers can perform close to the ultimate limit, and significantly better than the conventional schemes that neglect the channel memory.

The remainder of this paper is organized as follows. In Section II, we describe the considered channel model and compare it with the Bernoulli-Gaussian model. In Section III, we derive an algorithm for optimal MAP symbol detection, which is then exploited for the evaluation, in Section IV, of the achievable information rate. In Section V, we describe a couple of practical communication schemes, comparing their performance and complexity. Finally, some conclusions are drawn in Section VI.

II. CHANNEL MODEL

A sequence of M-ary complex-valued symbols $\{c_k\}_{k=1}^K$ belonging to a suitable constellation, e.g., phase-shift keying (PSK) or quadrature amplitude modulation (QAM), is transmitted over a discrete-time channel that introduces ad-

ditive noise. The received signal can be written as

$$y_k = c_k + n_k \quad , \quad k \in \{1, 2, \dots, K\} \tag{1}$$

where $\{n_k\}_{k=1}^K$ are noise samples. At each time epoch k, the statistical properties of the sample n_k are completely defined by the channel state s_k, which belongs to the binary alphabet $\{G, B\}$ — G for the good channel and B for the bad channel, according to the notation in [5]. We assume that n_k is a complex circularly-symmetric Gaussian random variable with variance depending on s_k, so that the probability density functions (PDFs) of n_k conditioned to s_k are

$$p(n_k|s_k = G) = \frac{1}{2\pi\sigma^2} \exp\left\{-\frac{|n_k|^2}{2\sigma^2}\right\} \tag{2}$$

$$p(n_k|s_k = B) = \frac{1}{2\pi R\sigma^2} \exp\left\{-\frac{|n_k|^2}{2R\sigma^2}\right\} \tag{3}$$

where the parameter $R \geq 1$ gives the ratio between the average noise power in the bad channel and that in the good channel. This model can describe a channel that, in the good state, is impaired only by background Gaussian noise, whereas, in the bad state, is impaired also by impulsive interferers [2], [3].

The statistical description of the state process s_1^K completes the characterization of the channel. A widely adopted model, proposed in [2], assumes that s_1^K is a stationary Bernoulli process [4], fully described by the probability that the bad state occurs

$$P_B = P(s_k = B) . \tag{4}$$

Here, we consider a more general model for the state process, assuming that s_1^K is a stationary first-order Markov process [4], that is,

$$P(s_1^{K+1}) = P(s_1) \prod_{k=1}^K P(s_{k+1}|s_k) \tag{5}$$

for each realization of the process. The evolution of the channel state is thus described by the transition probabilities

$$P_{IJ} = P(s_{k+1} = J|s_k = I) \quad , \quad I, J \in \{G, B\} \tag{6}$$

which here are all assumed non-zero. Since the memoryless channel model recalled above is often referred to as Bernoulli-Gaussian channel due to the underlying Bernoulli state process, we will refer to the considered model as Markov-Gaussian channel due to the underlying Markov state process. The latter model reduces to the former when the transition probabilities (6) depend on the arrival state only. We also point out that the Markov-Gaussian channel turns out to be a particular instance of the general model proposed in [3] for describing channels impaired by bursty impulse noise.

Note that the pair (P_{GB}, P_{BG}) suffices for a complete description of the state process [4]. The probabilities of being in a given state are

$$P_G = P(s_k = G) = \frac{P_{BG}}{P_{GB} + P_{BG}} \tag{7}$$

$$P_B = P(s_k = B) = \frac{P_{GB}}{P_{GB} + P_{BG}} \tag{8}$$

and the average number of consecutive samples of persistence in a given state are

$$T_G = \frac{1}{P_{GB}} \tag{9}$$

$$T_B = \frac{1}{P_{BG}} \tag{10}$$

for the good state and the bad state, respectively [4]. In all typical scenarios, the bad state is relatively infrequent, that is, $P_G > P_B$ or, equivalently, $T_G > T_B$. Equations (9) and (10) clarify how to derive the statistical parameters (P_{GB}, P_{BG}) from the knowledge of the physical parameters (T_G, T_B), which can be measured on field [3].

Let us define the parameter

$$\gamma = \frac{1}{P_{GB} + P_{BG}} \tag{11}$$

which, as explained later, quantifies the channel memory. Based on (7), (8), and (11), we obtain the following constraint on the value of γ

$$\gamma > \max\{P_B, 1 - P_B\} \tag{12}$$

by imposing that both P_{GB} and P_{BG} belong to $(0, 1)$. When (12) is satisfied, the pair (P_B, γ) is in one-to-one correspondence with the pair (P_{GB}, P_{BG}), and thus completely describes the state process. Hence, with respect to the Bernoulli-Gaussian model, where the state process is completely described by P_B, the Markov-Gaussian model requires the further parameter γ, whose meaning is now discussed. First, it can be shown that the channel is memoryless if and only if $\gamma = 1$. Moreover, for any given value of P_B, the parameter γ equals the ratio between the actual value of T_I and the value of T_I as if the channel were memoryless, where I can be either G or B. Hence, according to the notation in [5], we can state that, when $\gamma > 1$, the channel has a *persistent* memory, that is, the average permanence in a state is longer with respect to the memoryless case, whereas, when $\gamma < 1$, the channel has an *oscillatory* memory, that is, the average permanence in a state is shorter with respect to the memoryless case.[1] Since we are interested in modeling interferers that occur in bursts, thus with persistent memory by physical nature, we will focus on scenarios where $\gamma \geq 1$. The impact of the parameter γ on the noise process, when all other channel parameters are fixed, is shown in Fig. 1. It is clear that, while the classical memoryless model ($\gamma = 1$) fails in describing the bursts typically affecting the actual channels, the Markov model properly does, allowing to set the average duration of such bursts simply by modifying the value of γ.

III. MAXIMUM-A-POSTERIORI SYMBOL DETECTION

We now consider the computation, for each time epoch k and each symbol c_k belonging to the modulation alphabet, of the *a posteriori* probability $P(c_k|y_1^K)$ that the actual transmitted symbol is equal to c_k, given the received sequence y_1^K. Based on information-theoretical arguments, we

---

[1]In [5], the parameter $\mu = 1 - 1/\gamma$ is adopted for the characterization of the channel memory. We prefer the parameter γ because of the direct proportionality between its value and the average duration of the permanence in a state. For example, a doubled value of γ readily calls for a doubled average duration of the impulse bursts.
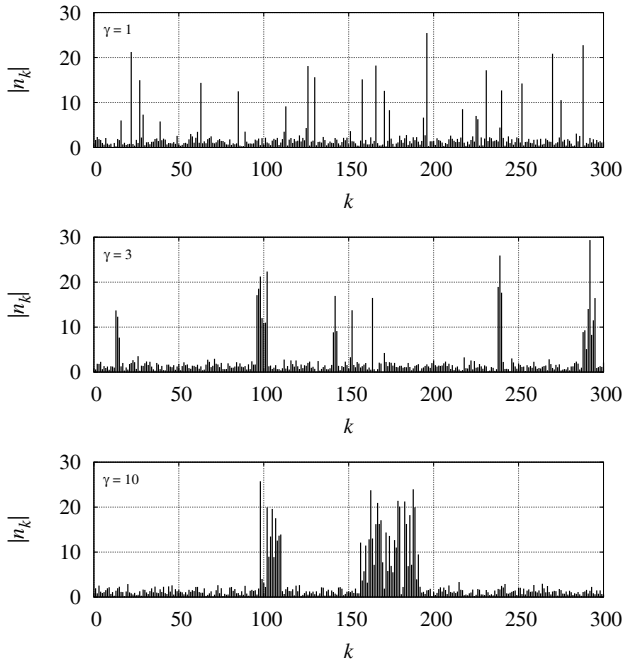
Fig. 1. Magnitude of the noise process for three different values of γ, when the other channel parameters are $\sigma^2 = 1$, $P_B = 10^{-1}$, and $R = 10^2$.

will first exploit the relevant algorithm for computing the ultimate performance limits imposed by the channel. Then, we will explain how this algorithm can be employed as soft-input soft-output (SISO) module for iterative detection/decoding. The algorithm, whose derivation is omitted being basically an instance of the general forward-backward algorithm (FBA) discussed in [13], can simply be obtained as an application of the factor graphs (FGs) and the sum-product algorithm (SPA) [13], but the same result can be achieved by means of the probabilistic arguments leading to the SISO algorithm in [14].

Defining

$$F_k(c_k, s_k, s_{k+1}) = P(s_{k+1}|s_k)p(n_k = y_k - c_k|s_k) \qquad (13)$$

and denoting by $P(c_k)$ the *a priori* probability that the source emits, at time epoch k, the symbol c_k, the algorithm is based on the following forward and backward recursions

$$\alpha_{k+1}(s_{k+1}) = \sum_{s_k, c_k} F_k(c_k, s_k, s_{k+1}) P(c_k) \alpha_k(s_k) \qquad (14)$$

$$\beta_k(s_k) = \sum_{s_{k+1}, c_k} F_k(c_k, s_k, s_{k+1}) P(c_k) \beta_{k+1}(s_{k+1}) \quad (15)$$

and on the completion stage

$$U_k(c_k) = \sum_{s_k, s_{k+1}} F_k(c_k, s_k, s_{k+1}) \alpha_k(s_k) \beta_{k+1}(s_{k+1}) . \quad (16)$$

The recursions are initialized as follows [13]

$$\alpha_1(I) = P_I \qquad , \qquad I \in \{G, B\} \qquad (17)$$
$$\beta_{K+1}(I) = 1 \qquad , \qquad I \in \{G, B\} \qquad (18)$$

and the *a posteriori* probabilities are obtained from

$$P(c_k|\boldsymbol{y}_1^K) \propto P(c_k) U_k(c_k) \qquad (19)$$

according to the SPA, where the proportionality symbol $\propto$ denotes that the sides may differ for a positive multiplicative factor irrelevant for the detection process [13]. Equation (19) also clarifies that the terms $\{U_k(c_k)\}$ constitute the "extrinsic information" produced by the algorithm [15]. Although implementation-related details are not a major concern here, we point out that the algorithm results numerically safer when implemented in the logarithmic domain [16]. Also, note that a simpler implementation of the algorithm consists of propagating, at each time epoch k, not the values of $\alpha_k(G)$ and $\alpha_k(B)$ (or $\beta_k(G)$ and $\beta_k(B)$), but their ratio only.

It is worth to notice that the described algorithm degenerates into a symbol-by-symbol algorithm when the channel is memoryless, that is, when $\gamma = 1$. In this case, no recursion should be performed, and the completion stage simply becomes

$$U_k(c_k) = \sum_{s_k} p(n_k = y_k - c_k|s_k) P(s_k) . \qquad (20)$$

Finally, we notice that the described algorithm produces, as a by-product output, the *a posteriori* probability $P(s_k|\boldsymbol{y}_1^K)$ that the state s_k is the actual channel state at time epoch k, since

$$P(s_k|\boldsymbol{y}_1^K) \propto \alpha_k(s_k) \beta_k(s_k) \qquad (21)$$

according to the SPA. Hence, the detection algorithm also provides a Bayesian state identifier, which is expected to be much more effective than the threshold-based identifiers generally considered in the literature.

IV. ULTIMATE PERFORMANCE LIMITS

Before describing practical communication systems, we analyze the ultimate performance limits imposed by the considered channel. In particular, we evaluate the information rate between the sequence $\boldsymbol{c}_1^K$ and the received samples $\boldsymbol{y}_1^K$ [6]. We recall that it is possible to design communication schemes with arbitrarily low error rate only if the number of information bits transmitted per channel use does not exceed the information rate [6]. We will restrict ourselves to the case of memoryless sources that emit equally likely symbols. Following the arguments in [7], we do not expect that Markov sources and/or biased sources can improve the achievable information rate.

A simulation-based method for the computation of the information rate of finite-state channels is exploited in the following. Since the theoretical details of the method are described in [8], only the resulting algorithm is reported here. First, we generate, according to the statistical properties of the source and the channel, a very long sequence of symbols $\boldsymbol{c}_1^K$, as well as the corresponding sequence of received samples $\boldsymbol{y}_1^K$ — the meaning of "very long" is discussed in [8]. Then, for each time epoch $k \in \{1, 2, \ldots, K+1\}$ and for each value s_k of the state, we compute the coefficients

$$\mu_k(s_k) = p(s_k, \boldsymbol{y}_1^{k-1}) \qquad (22)$$
$$\nu_k(s_k) = p(s_k, \boldsymbol{y}_1^{k-1}|\boldsymbol{c}_1^{k-1}) \qquad (23)$$

where $p(s_k, \boldsymbol{y}_1^{k-1})$ is the joint PDF of s_k and $\boldsymbol{y}_1^{k-1}$, while $p(s_k, \boldsymbol{y}_1^{k-1}|\boldsymbol{c}_1^{k-1})$ is the same joint PDF conditioned to the knowledge of the transmitted symbols $\boldsymbol{c}_1^{k-1}$. A computationally-efficient way for the evaluation of the terms $\{\mu_k(s_k)\}$ and $\{\nu_k(s_k)\}$ consists of resorting to the forward
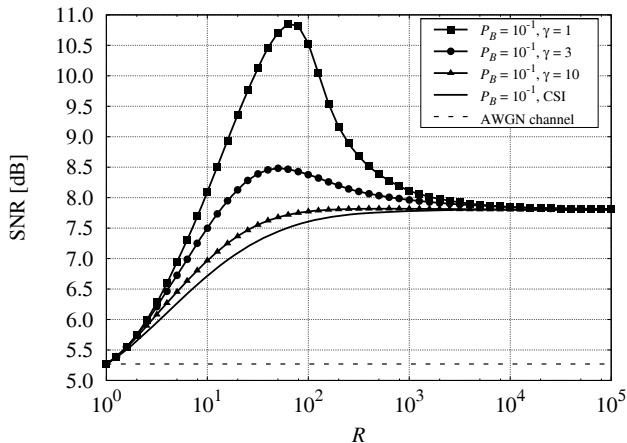
Fig. 2. Signal-to-noise ratio required to achieve an information rate equal to 1.75 bits per channel use. A QPSK modulation is adopted and matched decoding is considered.
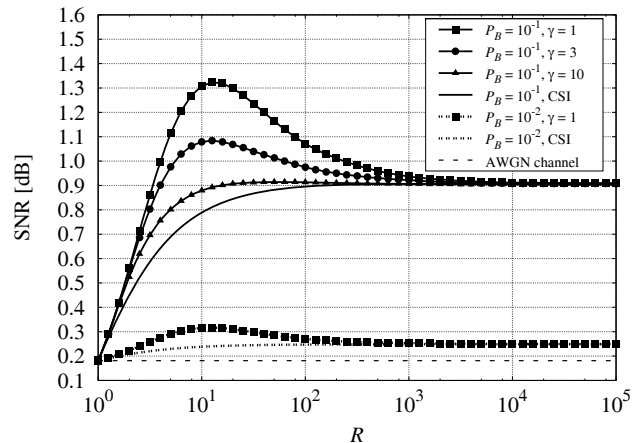

Fig. 3. Signal-to-noise ratio required to achieve an information rate equal to one bit per channel use. A QPSK modulation is adopted and matched decoding is considered.

recursion (14), by properly setting the *a priori* probabilities $\{P(c_k)\}$ [8]. Namely, the coefficients $\{\mu_k(s_k)\}$ are evaluated by setting each term $P(c_k)$ equal to $1/M$, since equally likely modulation symbols are considered, whereas the coefficients $\{\nu_k(s_k)\}$ are evaluated by setting, at each time epoch k, the term $P(c_k)$ equal to one if c_k is equal to the actual transmitted symbol, and to zero otherwise [8]. Finally, the simulation-based estimation of the information rate is evaluated as [8]

$$\frac{1}{K} \log_2 \left(\frac{\nu_{K+1}(G) + \nu_{K+1}(B)}{\mu_{K+1}(G) + \mu_{K+1}(B)} \right). \qquad (24)$$

We point out that the implementation of the FBA, including the forward recursion (14) required for the evaluation of the information rate, assumes that the values of P_B, γ, R, and σ^2 are known. Since, in most practical scenarios, these parameters are actually unknown, it is interesting to evaluate how the system performance is impaired by erroneous assumptions on their values. As explained in [8], the information rate achievable in this condition can still be evaluated by means of the simulation-based method described above. In particular, the channel output should be simulated according to the actual channel parameters, whereas the evaluation of the terms $\{\mu_k(s_k)\}$ and $\{\nu_k(s_k)\}$ should be based on the erroneous channel parameters.

Some of the obtained results are reported and discussed in the following, first for the condition of *matched decoding*, that is, ideal knowledge of the channel parameters, then for the condition of *mismatched decoding*, that is, erroneous assumptions on the channel parameters.

A. Matched Decoding

In Fig. 2, it is shown how the value of the signal-to-noise ratio (SNR) required to achieve an information rate equal to 1.75 bits per channel use varies when different channels are considered and a quaternary PSK (QPSK) is adopted. The SNR is here defined with respect to the background Gaussian noise only, that is,

$$\text{SNR} = \frac{E\left\{|c_k|^2\right\}}{2\sigma^2} \qquad (25)$$

where $E\{\cdot\}$ is the expectation operator. Together with the curves related to Markov-Gaussian channels characterized by $P_B = 10^{-1}$ and different values of γ and R, we also reported the corresponding curves related to an AWGN channel and a system with ideal channel-state information (CSI), that is, a system that knows the actual realizations of the state process. Note that the information rate of systems with CSI equals the statistical average of the information rates over the channel states, so that the value of γ is irrelevant [7]. It is clear that, for a given value of R, the theoretical power efficiency monotonically improves as the memory increases, up to the value related to the system with CSI. This was expected, since the memory helps to track the state process, up to the limit given by the ideal CSI tracker. On the other hand, when the value of R is very large, all curves tend to the value related to the system with CSI, irrespectively of the amount of memory. This also was expected, since the stronger the impulsive interferers, the easier to detect them even when the state tracking is not helped by the memory. It is also interesting to notice that, for a given value of γ, the theoretical power efficiency exhibits a non-monotonic behavior with respect to R. In particular, there exists a threshold value for R depending on γ such that, above that value, the larger the value of R the better the performance. Such a behavior, which tends to vanish as the value of γ increases and vanishes at all for the system with CSI, is somehow surprising since it implies that the system can take advantage of a larger power of the impulsive interferers. The key point is, again, that the stronger the interferers, the easier to detect their presence.

Although similar considerations *qualitatively* hold for all values of P_B and all values of the target information rate, the performance gain provided by the presence of memory *quantitatively* depends on such values. This is proved by comparing the simulation results reported in Fig. 2 and those reported in Fig. 3, which refer to QPSK transmissions with a target information rate equal to one bit per channel use. Namely, in the former case the memory can provide performance gains in the order of 3 dB, while in the latter case the beneficial effect of the memory is very limited when $P_B = 10^{-1}$ and practically negligible when $P_B = 10^{-2}$.
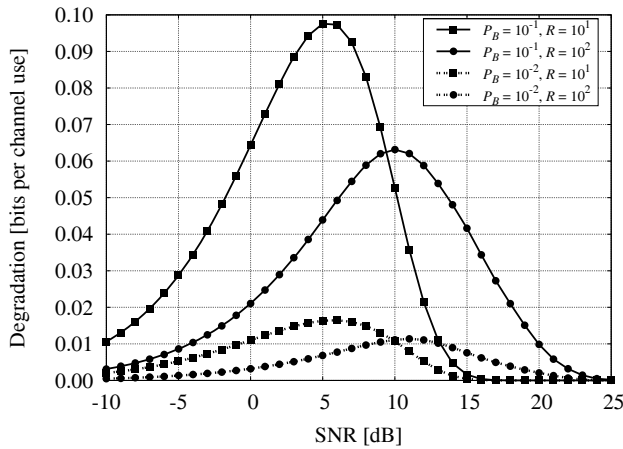
Fig. 4.   Degradation in information rate between systems with CSI and the corresponding channels with $\gamma = 1$. A QPSK modulation is adopted and matched decoding is considered.
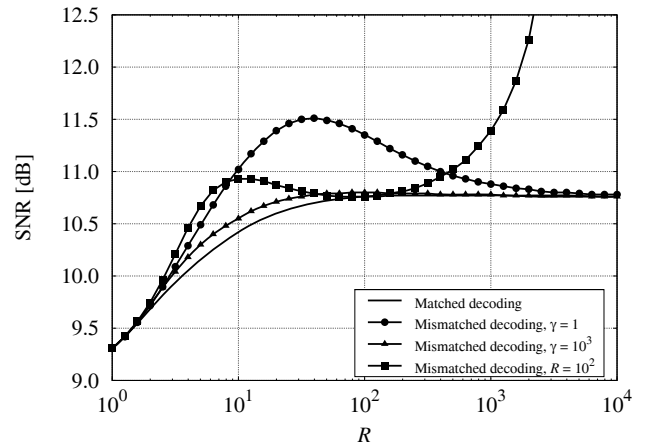

Fig. 5.   Signal-to-noise ratio required to achieve an information rate equal to 3.0 bits per channel use. A 16-QAM modulation is adopted and various conditions of mismatched decoding are considered. The actual channel is characterized the $P_B = 10^{-1}$ and $\gamma = 10$.

The impact of the memory can also be characterized by evaluating how the achievable information rate varies with γ when all other channel parameters are fixed. In this case, as discussed above, the information rate increases with the value of γ, and takes values in the range whose lower limit is given by the memoryless channel with $\gamma = 1$ and whose upper limit is given by the corresponding system with CSI. The span of such ranges is reported in Fig. 4 for QPSK transmissions and different channel parameterizations. These results definitely show that the impact of the memory on the achievable information rate is limited, and tends to vanish when the value of P_B is very low, that is, when the presence of impulsive interferers is very infrequent. Hence, the presence of memory can provide a significant gain in terms of SNR when we fix the target information rate (see Fig. 2), but it does not provide a significant gain in terms of achievable information rate when we fix the target SNR (see Fig. 4).

B. Mismatched Decoding

In Fig. 5, it is shown how the value of the SNR required to achieve an information rate equal to 3 bits per channel use varies when different channels are considered and a 16-QAM is adopted. In all cases, the channel is characterized the $P_B = 10^{-1}$ and $\gamma = 10$, while the parameter R is the abscissa of the plot. The performance in different conditions of mismatched decoding are compared with the performance in the ideal condition of matched decoding. For each curve reported in Fig. 5, the mismatched parameter and the related (erroneous) value are pointed out in the key. Let us first comment on the impact of an erroneous assumption on the memory parameter γ. According to Fig. 5, a receiver that assumes $\gamma = 1$, that is, a receiver that neglects the channel memory, may experience a degradation of almost 1 dB with respect to the optimal receiver. On the other hand, a receiver that strongly overestimates the value of γ, assuming $\gamma = 10^3$, is much more effective. Many other simulation results, which are not reported here due to space limitations, confirm this behavior, showing that the classical memoryless assumption may cause degradations even of 3 dB or more, thus motivating the design of practical communication schemes able to exploit

the channel memory — this issue is addressed in Section V. As a rule of thumb, it turns out that we should design the system assuming the presence of some memory, and the value of γ can be safely overestimated.

Other results suggest that the estimate of the parameter R is critical for the performance of the FBA (for example, see Fig. 5), while the estimate of the parameter P_B is not. We point out that the estimate of the parameter R may be much less critical, or even not critical at all, when blind detection schemes are used instead of the FBA — this point is addressed in [17], [18] for channels with $\gamma = 1$. In the following, since we are here mainly interested in the impact of the parameter γ, it is assumed that correct information on all other channel parameters is available.

V. PRACTICAL COMMUNICATION SYSTEMS

We are interested in designing communication systems that can approach the ultimate performance limits evaluated in Section IV. Hence, we focus on systems employing powerful channel codes, such as LDPC codes or turbo-like codes [10], [11]. The block scheme of the considered system is depicted in Fig. 6, and is briefly described in the following. At the transmitter side, a sequence of equally likely information bits is first encoded, then interleaved, and finally mapped onto the modulation sequence c_1^K. The channel corrupts the sequence c_1^K, generating the noisy samples y_1^K according to the Markov-Gaussian model. At the receiver side, we consider a scheme based on the exchange of extrinsic information [15] between the FBA detector described in Section III and the SISO decoder, which iteratively refine the quality of their outputs similarly to the receivers employing turbo equalization [12] — in this case, the term "turbo detection" may be preferred, but the underlying idea is the same, that is, the application of the SPA to a FG with cycles [13]. In particular, at each iteration, the detection algorithm updates the extrinsic information $\{U_k(c_k)\}$ based on the last vector $\{P(c_k)\}$ produced by the SISO decoder, which then updates the vector $\{P(c_k)\}$ based on the novel terms $\{U_k(c_k)\}$. The exchange of extrinsic information between these two blocks is managed by SISO mapping/demapping, as explained in [19].
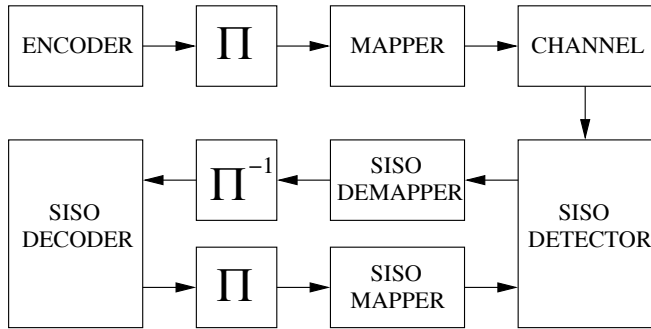
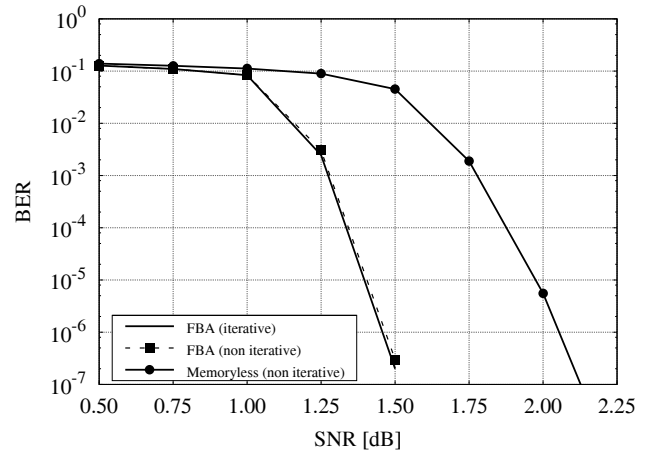Fig. 6. Block scheme of a system employing iterative detection/decoding.

Fig. 7. Performance of various detection schemes over a Markov-Gaussian channel characterized by $P_B = 10^{-1}$, $\gamma = 10$, and $R = 10$. A system employing a rate-1/2 LDPC code and a QPSK modulation is considered.

As a reference benchmark for the performance of the proposed scheme, we adopt the conventional approach that consists of neglecting the channel memory. This leads to a simpler receiver with respect to that depicted in Fig. 6, since, once assumed the absence of memory, the detection algorithm works symbol-by-symbol producing the metrics (20), and does not require to exchange extrinsic information with the SISO decoder — the terms $\{P(c_k)\}$ are indeed not exploited in (20). Hence, this approach provides a significant improvement in terms of latency with respect to the scheme based on the FBA. On the other hand, the information-theoretical results reported in Section IV show that, in many scenarios, the potential performance gain justifies the increase of latency. A "hybrid" solution, aimed at reducing the latency that characterizes the proposed scheme without neglecting the channel memory, consists of executing the FBA only once per received codeword, without further iterations with the SISO decoder.

The performance of the considered schemes was assessed by means of computer simulations in terms of bit-error rate (BER) versus SNR. In Fig. 7, the BER curves related to different detection schemes are reported, for the case of a system employing QPSK transmissions over a Markov-Gaussian channel characterized by $P_B = 10^{-1}$, $\gamma = 10$, and $R = 10$. The adopted channel code is a rate-1/2 irregular LDPC code with codewords of 20,000 bits, whose degree distribution, optimized for AWGN channels, is taken from [20] — the choice of this code is motivated at the end of the section. Due to the pseudo-random structure of the LDPC code, no interleaver is used, and two consecutive code bits are directly Gray-mapped onto a QPSK symbol. At the receiver side, the decoder performs a maximum of 100 iterations before producing the hard decisions on the information bits [10]. We remark that the proposed scheme employing iterative detection/decoding provides a performance gain of about 0.6 dB with respect to the conventional memoryless one. Interestingly, if the FBA is applied without iterative detection/decoding, the performance is practically the same — this fact is discussed at the end of the section. Note that the obtained results can be directly compared with the ultimate limits reported in Fig. 3, since the use of a rate-1/2 binary code combined with a QPSK modulation implies an actual information rate equal to one bit per channel use. For the considered channel parametrization, the proposed scheme achieves a BER of 10^{-6} when the SNR is about 0.5 dB worse than the ultimate limit shown in Fig. 3. This result, yet satisfactory, can be improved by adopting longer codewords and/or increasing the

number of decoding iterations [10]. Hence, given the excellent agreement between the theoretical results and the performance of the practical systems, we can definitely confirm that the conventional approach of neglecting the channel memory, which is actually present in any real channel due to the bursty nature of the interferers, can be significantly outperformed.

Finally, we discuss the fact that iterative detection/decoding does not significantly improve the performance of the FBA in the scenario considered above. This suggests that the FBA cannot take advantage of the soft information produced by the LDPC decoder. It is interesting to investigate if this behavior is peculiar of the considered scenario or if it is a general property, irrespectively of the noise parameterization. A simple tool for characterizing how much a detection algorithm can take advantage of iterating with a SISO decoder consists of the extrinsic information transfer (EXIT) charts [20]. We carried out extensive simulations for various parameterizations of the Markov-Gaussian channels, and found that, in all of them, the detection algorithm is expected not to significantly gain from iterating with a SISO decoder — in the context of [20], it means that the detection EXIT chart is nearly flat. Some EXIT charts related to different detection strategies and channel parameterizations, pointed out in the key and in the caption, are reported in Fig. 8, where the standard notation for labelling the plot axes is adopted [20]. Note that the EXIT charts related to the FBA are as flat as those related to the memoryless detector, which definitely shows that there is no need for iterative detection/decoding — this fact also implies that LDPC codes optimized for AWGN channels are good for the Markov-Gaussian channels too. Interestingly, the same conclusion does not hold for more general Markov models, for which the code design should thus be based on ad-hoc techniques as those presented in [21] and [22].

VI. CONCLUSIONS

The performance of communication systems over channels impaired by Markov-Gaussian impulse noise has been analyzed. We have first approached this issue from an information-theoretical viewpoint, showing that the information rate of such systems improves as the channel memory
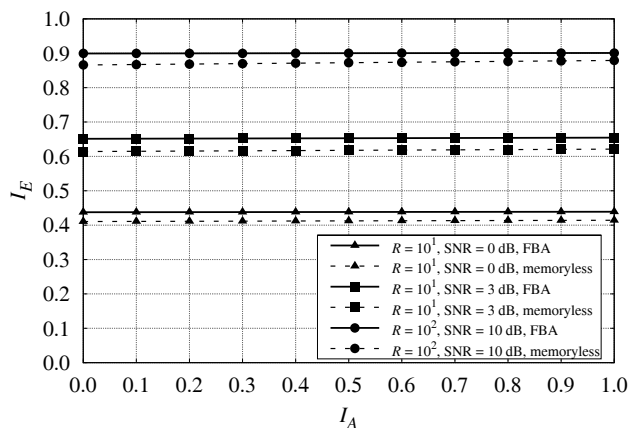
Fig. 8.   EXIT charts related to different detection strategies and different channel parameterizations. In all cases, we adopted a Gray-mapped QPSK modulation and we set $P_B = 10^{-1}$ and $\gamma = 10$.

becomes more significant. Then, motivated by these results, we have described a couple of practical detection schemes able to exploit the channel memory. The proposed schemes have been shown to perform fairly close to the ultimate limit and to provide, in various scenarios, a significant gain with respect to the conventional memoryless ones.

Acknowledgment

References

[1] D. Middleton, "Statistical-physical model of electromagnetic interference," *IEEE Trans. Electromagn. Compat.*, vol. 19, no. 3, pp. 106-126, Aug. 1977.

[2] M. Ghosh, "Analysis of the effect of impulse noise on multicarrier and single carrier QAM systems," *IEEE Trans. Commun.*, vol. 44, pp. 145-147, Feb. 1996.

[3] M. Zimmermann and K. Dostert, "Analysis and modeling of impulse noise in broad-band powerline communications," *IEEE Trans. Electromagn. Compat.*, vol. 44, pp. 249-258, Feb. 2002.

[4] A. Papoulis, *Probability, Random Variables and Sthocastic Processes*. New York: McGraw-Hill, 1991.

[5] M. Mushkin and I. Bar-David, "Capacity and coding for the Gilbert-Elliott channels," *IEEE Trans. Inform. Theory*, vol. 35, pp. 1277-1290, Nov. 1989.

[6] T. M. Cover and J. A. Thomas, *Elements of Information Theory*. New York: John Wiley & Sons, Inc., 1991.

[7] A. Goldsmith and P. Varaiya, "Capacity, mutual information and coding for finite-state Markov channels," *IEEE Trans. Inform. Theory*, vol. 42, pp. 868-886, May 1996.

[8] D. M. Arnold, H.-A. Loeliger, P. O. Vontobel, A. Kavčič, and W. Zeng, "Simulation-based computation of information rates for channels with memory," *IEEE Trans. Inform. Theory*, vol. 52, no. 8, pp. 3498-3508, Aug. 2006.

[9] N. Merhav, G. Kaplan, A. Lapidoth, and S. S. Shitz, "On information rates for mismatched decoders," *IEEE Trans. Inform. Theory*, vol. 40, pp. 1953-1967, Nov. 1994.

[10] T. Richardson and R. Urbanke, "The capacity of low density parity check codes under message passing decoding," *IEEE Trans. Inform. Theory*, vol. 47, pp. 599-618, Feb. 2001.

[11] S. Benedetto, D. Divsalar, G. Montorsi, and F. Pollara, "Serial concatenation of interleaved codes: performance analysis, design, and iterative decoding," *IEEE Trans. Inform. Theory*, vol. 44, pp. 909-926, May 1998.

[12] M. Tüchler, R. Koetter, and A. C. Singer, "Turbo equalization: principles and new results," *IEEE Trans. Commun.*, vol. 55, pp. 754-767, May 2002.

[13] F. R. Kschischang, B. J. Frey, and H.-A. Loeliger, "Factor graphs and the sum-product algorithm," *IEEE Trans. Inform. Theory*, vol. 47, pp. 498-519, Feb. 2001.

[14] L. R. Bahl, J. Cocke, F. Jelinek, and J. Raviv, "Optimal decoding of linear codes for minimizing symbol error rate," *IEEE Trans. Inform. Theory*, vol. 20, pp. 284-287, Mar. 1974.

[15] G. Colavolpe, G. Ferrari, and R. Raheli, "Extrinsic information in iterative decoding: a unified view," *IEEE Trans. Commun.*, vol. 49, pp. 2088-2094, Dec. 2001.

[16] P. Roberston, E. Villebrun, and P. Hoeher, "Optimal and sub-optimal maximum a posteriori algorithms suitable for turbo decoding," *European Trans. Telecommun.*, vol. 8, pp. 119-125, Mar./Apr. 1997.

[17] D. Fertonani and G. Colavolpe, "A simplified metric for soft-output detection in the presence of impulse noise," in *Intl. Symp. Power Line Commun. Its Appl.*, Pisa, Italy, Mar. 2007.

[18] D. Fertonani and G. Colavolpe, "A robust metric for soft-output detection in the presence of class-A noise," *IEEE Trans. Commun.*, vol. 57, pp. 36-40, Jan. 2009.

[19] X. Li and J. A. Ritcey, "Bit-interleaved coded modulation with iterative decoding using soft feedback," *Electron. Lett.*, vol. 34, pp. 942-943, May 1998.

[20] S. ten Brink, G. Kramer, and A. Ashikhmin, "Design of low-density parity-check codes for modulation and detection," *IEEE Trans. Commun.*, vol. 52, pp. 670-678, Apr. 2004.

[21] A. W. Eckford, F. R. Kschischang, and S. Pasupathy, "On designing good LDPC codes for Markov channels," *IEEE Trans. Inform. Theory*, vol. 53, pp. 5-21, Jan. 2007.

[22] H. Lou and J. Garcia-Frias, "Low-density generator matrix codes for indoor and Markov channels," *IEEE Trans. Wireless Commun.*, vol. 6, pp. 1436-1445, Apr. 2007.

Dario Fertonani (S'06) was born in Mantua, Italy, in 1980. He received the Dr.Ing. degree (cum laude) in telecommunications engineering in 2004 and the Ph.D degree in Information Technology in 2008, both from the University of Parma, Italy. Then, he held a post-doc position at the Arizona State University, Tempe, AZ. Currently, he is a Research Associate at Scuola Superiore Sant'Anna, Pisa, Italy.

His research interests include various topics in digital communications and information theory, with particular interest on design issues related to iterative detection/decoding over channels with memory and channels with impulse noise. He is coauthor of publications on international journals and proceedings of international conferences, and was one of the recipients of the student travel grant for IEEE ICC 2008. Also, he participated in research projects funded by the European Space Agency and the Jet Propulsion Laboratory.

Giulio Colavolpe (S'96-M'00) received the Dr. Ing. degree in Telecommunications Engineering (cum laude) from the University of Pisa, Italy, in 1994 and the Ph.D. degree in Information Technology from the University of Parma, Italy, in 1998. Since 1997, he has been at the University of Parma, Italy, where he is now an Associate Professor of Telecommunications. In 2000, he was Visiting Scientist at the Institut Eurécom, Valbonne, France.

His main research interests include digital transmission theory, adaptive signal processing, channel coding and information theory. His research activity has led to numerous scientific publications in leading international journals and conference proceedings and a few industrial patents. He is also co-author of the book *Detection Algorithms for Wireless Communications, with Applications to Wired and Storage Systems* (New York: John Wiley & Sons, 2004). He received the best paper award at the 13th International Conference on Software, Telecommunications and Computer Networks (SoftCOM'05) and at the IEEE International Conference on Communications (ICC 2008).