Improved Low-Density Parity-Check Codes for Burst Erasure Channels

Enrico Paolini and Marco Chiani
DEIS, WiLAB
University of Bologna
via Venezia 52, 47023, Cesena, Italy
Email: {epaolini, mchiani}@deis.unibo.it

Abstract—In this work we deal with Low-Density Parity-Check (LDPC) codes under iterative message passing decoding algorithm, over channels introducing bursts of erasures. The burst erasure channel model we consider in this paper can be seen as an erasure channel based on a hidden Markov chain (HMC-EC). In order to characterize the channel, in the first part of the paper the expression of mutual information is recalled for any erasure channel with memory and with i.i.d. and equiprobable input symbols. In the second part of the paper an optimization algorithm is proposed which is able to heavily improve LDPC iterative decoder performance. This algorithm can be in principle applied to any given LDPC code. Simulation results relative to both random and IRA / eIRA codes are shown comparing the performance before and after the application of our optimization algorithm.

I. INTRODUCTION

A recent coding research field concerns the possibility to use long codes based on sparse graphs, above all LDPC codes, within packet erasure correcting protocols in many communication systems, including CCSDS (Consultative Committee for Space Data Systems) applications and the satellite broadcast [1]–[3]. The main reason for this investigation is that the upper layers in the communication stack have often to deal with packet erasures. In wireless communications, packet losses can be caused, depending on the scenario and on the transmission band, by weather, shadowing, loss of frame synchronization, or by detection and consequent discarding of erroneously received packets. Then, the packet erasure channel represents a suitable channel model for the upper layers. In packet erasure correcting LDPC codes, each information/encoded symbol (and each variable node) represents a packet of bits. For fixed length \( q \) packets, encoded packets are obtained by bitwise XOR of information packets. In this case, the packet oriented LDPC code is in principle equivalent to \( q \) binary LDPC codes working in parallel (with simple implementation).

The rising interest for LDPC codes in such applications is due to some their good properties. First, long codes can contrast long block fading. Second, the memory-less erasure channel capacity can be asymptotically achieved by LDPC codes under iterative message passing decoding, on condition that the bipartite graph has proper degree distribution pairs [4]. Third, the message passing decoder is very efficient. Fourth, the traditional problem concerning the encoding of LDPC codes has been overcome, as several solutions have been proposed which permit to perform the encoding operation with a complexity linearly increasing with the codeword length, while preserving good or very good decoding performance. In most of these solutions some constraints are imposed to the positions of the 1’s in the parity-check matrix \( H \) in order to permit an efficient encoding (structured or partially structured codes). On condition that these constraints are satisfied, the parity-check matrix can be constructed randomly. Examples considered in this paper are IRA [5] and eIRA [6] codes. Since eIRA are a generalization of IRA codes, in the following we will use the expression eIRA to refer both categories.

In the mentioned applications packet losses are usually correlated, and they often occur in bursts. Although this correlation should be considered within the channel model, the memory-less packet erasure channel, where packets are lost independently with equal erasure probability, is generally adopted. This choice is justified by assuming that a sufficiently long packet interleaving is present in the system. Despite this assumption can be appropriate in some contexts, it appears quite rough if the interleaving procedure (if available) only involves packets belonging to the same codeword, and the erasure burst length is not small respect to the codeword length \( n \) (where \( n \) is intended as number of encoded packets).

In this paper we remove the memory-less assumption and consider LDPC codes over channels where packet losses occur in bursts. We consider burst lengths which are considerable fractions of the codeword length, in a regime where each codeword is affected at most by one burst with high probability. This regime, though of notable practical interest, is scarcely treated in existing literature [7]. Burst erasure channels are modelled here as particular instances of erasure channels based on hidden Markov chains (HMC-EC), where the packet erasure probability depends on the state of a Markov chain which executes a state transition at each new packet. In order to characterize HMC-EC’s, the expression of mutual information is first recalled for any erasure channel with memory and with i.i.d. equiprobable input symbols. Then, in the second part of the paper an algorithm is proposed which is able to considerably improve the performance of LDPC codes over burst erasure channels.

The paper is organized as follows. In Section II we recall the mutual information for erasure channels with memory (and HMC-EC’s as a special case), and introduce the channel model...
used for simulations. In Section III we describe the LDPC codes optimization algorithm. Section IV shows numerical results about random and eIRA codes improved by our algorithm. Finally, some concluding remarks are presented in Section V.

II. MUTUAL INFORMATION FOR ERASURE CHANNELS WITH MEMORY

Consider a general erasure channel with memory, and with discrete input alphabet \( \mathcal{X} \) of size \(|\mathcal{X}|\). The corresponding output alphabet is \( \mathcal{Y} = \mathcal{X} \cup \{?\} \), where ? denotes an erased symbol. Let the input symbols be i.i.d. and equiprobable. Let \( X_\ell \) and \( Y_\ell \) be, respectively, the input and output channel symbols at time \( \ell \). Moreover, let \( X_\ell \) and \( Y_\ell \) be the input and output sequences \( \{X_1, X_2, \ldots, X_\ell\} \) and \( \{Y_1, Y_2, \ldots, Y_\ell\} \).

What we need to compute in order to characterize the channel is [8]

\[
I = \lim_{\ell \to \infty} \frac{1}{\ell} I(X_\ell; Y_\ell),
\]

where \( I(X_\ell; Y_\ell) \) is the average mutual information between the input and output sequences.

The expression (1) can be developed as:

\[
\lim_{\ell \to \infty} \frac{1}{\ell} I(X_\ell; Y_\ell) = \lim_{\ell \to \infty} \frac{1}{\ell} \left[ H(X_\ell) - H(X_\ell | Y_\ell) \right]
= \lim_{\ell \to \infty} \frac{1}{\ell} \left[ H(X_\ell) - \sum_{y_\ell} p(y_\ell) H(X_\ell | Y_\ell = y_\ell) \right]
= \lim_{\ell \to \infty} \frac{1}{\ell} \left[ H(X_\ell) - \sum_{e=0}^{\ell} \sum_{y_\ell} p(y_\ell^e) H(X_\ell | Y_\ell = y_\ell^e) \right],
\]

where \( e \) is the number of erasures in the received sequence of length \( \ell \) and \( y_\ell^e \) is a realization of the received sequence affected by \( e \) erasures.

Now \( H(X_\ell | Y_\ell = y_\ell^e) \) does not depend on the specific received sequence \( y_\ell^e \), but just on the specific number \( e \) of erasures in that sequence. In other words, \( H(X_\ell | Y_\ell = y_\ell^e) = H(X_\ell | N_\ell = e) \), where \( N_\ell \) is a random variable equal to the number of erasures up to time \( \ell \). For i.i.d. and equiprobable input symbols, expression (2) can be further developed as follows (\( \log() \) is the base 2 logarithm).

\[
\lim_{\ell \to \infty} \frac{1}{\ell} \left[ H(X_\ell) - \sum_{e=0}^{\ell} \sum_{y_\ell} p(y_\ell^e) H(X_\ell | Y_\ell = y_\ell^e) \right]
= \lim_{\ell \to \infty} \left[ \ell H(X) - \sum_{e=0}^{\ell} H(X_\ell | N_\ell = e) \sum_{y_\ell} p(y_\ell^e) \right]
= \lim_{\ell \to \infty} \left[ \ell H(X) - \sum_{e=0}^{\ell} e H(X) Pr(N_\ell = e) \right]
= \log|\mathcal{X}| \left[ 1 - \lim_{\ell \to \infty} \frac{\mathbb{E}[N_\ell]}{\ell} \right]
= \log|\mathcal{X}|(1 - \bar{p}),
\]

where \( \bar{p} \) is the average fraction of erasures.

Suppose that the erasure channel is based on an irreducible and aperiodic hidden Markov chain with \(|S|\) states belonging to the set \( S = \{S(1), S(2), \ldots, S(|S|)\} \) and time-invariant transition probabilities matrix \( P \) (\( P_{i,j} \) is the transition probability from state \( S(i) \) to state \( S(j) \)). The states of the process are considered to be ergodic, with stationary state probabilities vector \( \mathbf{u} = [u(1), u(2), \ldots, u(|S|)] \) (i.e. \( \mathbf{u} = \mathbf{u} P \) [9]. The generic state \( S(i) \) corresponds to a memory-less erasure channel with channel erasure probability \( p_i \) (see an example in Fig. 1 with \(|S| = 3\)). In this case \( \bar{p} = \sum_{j=1}^{|S|} u(i) p_j \).

Expression (3) is the mutual information in bit of information. For the applications briefly described in the introduction, where whole packets are either lost or correctly received, it is usual to refer to a packet (i.e. symbol) oriented mutual information, given by (for length \( q \) binary packets it is \(|\mathcal{X}| = 2^q\))

\[
I' = 1 - \bar{p}.
\]

This expression reduces to the usual \( C = 1 - p \) for the memory-less packet erasure channel capacity.

Example 1 (Gilbert-Elliott erasure channel): The Markov chain has only two states, namely Good and Bad states, with erasure probabilities \( p_G \) and \( p_B \), and \( p_G < p_B \). Let \( b \) and \( g \) be, respectively, the Good-to-Bad and Bad-to-Good transition probabilities. Then

\[
I' = g (1 - p_G) + b (1 - p_B).
\]

Example 2 (CLBuEC): For the constant length burst erasure channel (CLBuEC) the states of the Markov chain can be partitioned in a Good state with erasure probability \( p_G = 0 \), and \( L \) Bad states each with erasure probability \( p_B = 1 \) (see Fig. 2). The channel moves from the Good state to the Bad states with transition probability \( b \), thus generating a burst of erasures of length \( L \). After the generation of the last erasure in the burst, the channel returns in the Good state. The mutual information is given by

\[
I' = \frac{1}{1 + b L}.
\]

This simplified channel model has been adopted for simulations presented in Section IV, in order to show the performance improvement achievable by the optimization algorithm described in next Section. Much more realistic channel models can be obtained by letting \( L \) vary according to some distribution. In Appendix we also derive a lower bound on the decoding failure probability of finite length LDPC codes over CLBuEC for \( L > L_{\text{max}} \) (see next Section).

III. OPTIMIZATION ALGORITHM DESCRIPTION

In this Section we propose an optimization algorithm for LDPC codes over burst erasure channels. Before introducing the algorithm, some useful concepts have to be briefly reviewed. By definition, a stopping set is any set of variable nodes such that any check node connected to this set is connected to it at least twice [10]. The residual erased
variable nodes (after iterative decoding) constitute the maximal stopping set included in the original erasure pattern. Hence, a decoding failure takes place whenever the erasure pattern due to the channel contains a stopping set. Another important concept is the maximum guaranteed resolvable burst length of an LDPC code, $L_{\text{max}}$, defined as the maximal burst length such that the iterative decoder is able to successfully recover from the burst independently of its position within the codeword [11]. For large codeword lengths $n$ and properly permuted variable nodes, it results $L_{\text{max}} \approx p^* n$, where $p^*$ is the threshold over the memory-less erasure channel. In order to improve LDPC codes performance over burst erasure channels, in a regime where the codewords are prevalently affected by just one burst, parameter $L_{\text{max}}$ should be maximized. This is the aim of the proposed algorithm. From an equivalent point of view, decoding is successful if the burst contains no stopping sets. Hence, the algorithm tries to remove concentrated stopping sets (i.e. stopping sets made up of variable nodes which are near in the graph), and to increase their dispersion (the maximal distance between two variable nodes belonging to the stopping set). The stopping sets dispersion, of no interest over memory-less erasure channel, comes out to be a key parameter over burst erasure channels.

The $L_{\text{max}}$ optimization algorithm is described in the rest of the Section. It is a variable nodes permutation algorithm which receives in input a specific LDPC code from the ensemble $C^n(\lambda, \rho)$ [12] with maximum resolvable burst length $L'_{\text{max}}$ and returns a new LDPC code from the same ensemble with $L''_{\text{max}} \geq L'_{\text{max}}$. The algorithm simply executes permutations of variable nodes, without modifying neither their connections towards the check nodes nor the degree distribution pair $(\lambda, \rho)$.

Consider now an LDPC code from the ensemble $C^n(\lambda, \rho)$, let $L'_{\text{max}}$ be its maximum resolvable burst length, and $\mathcal{V}$ be the set of its variable nodes. At least one position exists within the codeword of length $n$ such that an erasure burst beginning in this position with length $L'_{\text{max}} + 1$ cannot be successfully decoded. Denoting by $j$ be the index corresponding to such position (see Fig. 3), the set of variable nodes $B = \{V_1, \ldots, V_{j+L'_{\text{max}}} \}$ contains a non-empty maximal stopping set. Conversely, the two erasures bursts of length $L''_{\text{max}}$ beginning in positions $j$ and $j+1$ can be successfully decoded. Hence, no stopping sets are included in the sets of variable nodes $\{V_j, \ldots, V_{j+L'_{\text{max}}+1}\}$ and $\{V_{j+1}, \ldots, V_{j+L'_{\text{max}}}\}$. The key observation here is that both variable nodes $V_j$ and $V_{j+L'_{\text{max}}}$ must belong to the maximal stopping set of $B$. In fact, removing any of them from $B$ leads to a new set of variable nodes which contains no stopping sets. This observation represents the basis of the proposed algorithm. The optimization algorithm tries to remove the maximal stopping set of $B$ by exchanging the position of $V_j$ (or $V_{j+L'_{\text{max}}}$) with the position of another variable node chosen in $\mathcal{V} \setminus B = \{V_0, \ldots, V_{j-1}\} \cup \{V_{j+L'_{\text{max}}+1}, \ldots, V_{n-1}\}$.

More specifically, the erasure burst set is set to 1 at the beginning of the algorithm. If a burst length $L_1$ is recognized as resolvable, then the new burst length $L_1 + 1$ is investigated. Suppose that a burst length $L_2$ is recognized as non-resolvable, since the decoder is not able to successfully recover from an erasure burst of length $L_2$ beginning in some position $j$. In this case, the position of variable node $V_j$ and the position of variable node $V_{j+L_2}$ are permuted, for a certain $V_i \in \mathcal{V} \setminus B$. Variable nodes $V_i$’s are considered in the following order: $V_0, \ldots, V_{j-1}, V_{j+L_2}, \ldots, V_{n-1}$. If, due to the permutation of $V_j$ and $V_i$, the erasure burst length $L_2$ becomes resolvable then the permutation is confirmed and the burst length $L_2 + 1$ is investigated. On the contrary, the permutation is rejected and the successive $V_i \in \mathcal{V} \setminus B$ is considered. If no permutation is confirmed, variable node $V_{j+L_2-1}$ is considered instead of $V_j$. If even for $V_{j+L_2-1}$ no permutation is confirmed, the algorithm returns the new LDPC code, with $L_{\text{max}} = L_2 - 1$.

The proposed algorithm has some good properties. First, in most cases the improvement in terms of $L_{\text{max}}$ achievable by the application of the algorithm is remarkable (see Section IV). Second, the algorithm is very general and can be in principle applied to any LDPC code, and also within LDPC-based hybrid ARQ schemes. Third, the algorithm is quite efficient.

### IV. Numerical Results

In this Section we present numerical results illustrating the improvement in terms of $L_{\text{max}}$ achievable by applying the optimization algorithm described in previous Section. Simulation results are also shown illustrating the improvement in terms of decoding failure rate over the simple CLBuEC model. We first consider the theoretically important, but of limited practical interest, case of random codes. Then we move to consider the much more practical case of eIRA codes. In order to relate our work to concrete applications, we consider two irregular degree distribution pairs with code rates $1/2$ and $9/10$ which have been recently selected for implementation for packet erasure correction in the Ku-Mobile system (distributions and related thresholds are available in [2]). More in detail, eIRA codes with these distributions have been selected for the system. Both random and eIRA codes are constructed according to these distributions, with information block length $k = 1000$ packets, which is a realistic value.

#### A. Random Codes

The proposed optimization algorithm has a very good behavior when applied to random LDPC codes. Consider for example a random $R = 1/2$ irregular $(2000, 1000)$ LDPC code. We found a maximum resolvable length $L'_{\text{max}} = 836$, which is quite high if compared with the maximal possible value $p^* n = 931$. Even for an input code with such good maximum resolvable length, the algorithm was able to considerably improve it. In fact, we found for the optimized code an $L''_{\text{max}} = 904$ which is an excellent value, especially considering the moderate codeword length $n = 2000$.

In Fig. 4 the decoding failure rate over a CLBuEC with $L = 890$ ($L'_{\text{max}} < L < L''_{\text{max}}$) is illustrated for both the original and the improved code, as a function of the Good-to-Bad transition probability $b$. The gain achieved in terms of $b$ is one order of magnitude at decoding failure rate $10^{-5}$ (even
better for smaller values of the target performance). The failure rate improvement is one order of magnitude for $b = 10^{-5}$ and about two orders of magnitude for $b = 10^{-6}$.

**B. eIRA Codes**

For eIRA codes, the part of the parity-check matrix corresponding to parity symbols is structured. This enables for efficient encoding, and makes them very attractive from a practical point of view.

We found for eIRA codes values of $L_{\text{max}}$ quite poor respect to the corresponding values for random codes. This comes as a consequence of their long sequence of weight-2 columns. For example, some $L_{\text{max}}$ values we found for $R = 1/2$ irregular (2000, 1000) eIRA codes were 272, 137, 141, 184, much smaller than $L_{\text{max}} = 836$ for the random code with the same distribution. However, if the structured weight-2 columns are uniformly spread in $\mathbf{H}$ to obtain a new, still efficiently encodable, parity-check matrix, much higher values of $L_{\text{max}}$ can be found. In all cases the $L''_{\text{max}}$ for the improved code resulted on the order of 900.

A suboptimal solution for eIRA codes is to apply the optimization algorithm only to the systematic symbols. We call *systematic improved code* the code obtained with this approach. For example, the algorithm applied only to the systematic symbols of the $L'_{\text{max}} = 272$ eIRA was able to generate a systematic improved code with $L''_{\text{max}} = 621$.

Fig. 5 shows the decoding failure rate for the original $L'_{\text{max}} = 272$ eIRA, the systematic improved code and the improved code, on CLBuEC with $L = 630$ and $L = 500$. For $L = 630$ ($L''_{\text{max}} < L < L''_{\text{max}}$) the improved code exhibits very good performance respect to both the original and the systematic improved codes. For $L = 500$ ($L < L''_{\text{max}} < L''_{\text{max}}$) the systematic improved code and improved code have similar performance, thus revealing that the optimization of just the systematic columns can be sufficient for small or moderate burst lengths.

The performance curves over a CLBuEC with $L = 70$ for the analogue $R = 9/10$ irregular (1111,1000) eIRA codes are shown in Fig. 6. The algorithm was able to increase the $L'_{\text{max}} = 46$ for the original eIRA code up to $L''_{\text{max}} = 80$ for the improved code and up to $L''''_{\text{max}} = 79$ for the systematic improved code. These values are not far from the maximal value $p^*n = 93$, despite the relatively small codeword length $n = 1111$. The very similar values of $L''_{\text{max}}$ and $L''''_{\text{max}}$ are due to the limited impact of redundant structured degree 2 variable nodes on the whole graph. This is confirmed by the very similar (and quite good, respect to the original eIRA) performance curves in Fig. 6, which reveals how for high code rates the the systematic improved approach guarantees the achievable gain respect to the original code.

**V. CONCLUSIONS**

In this paper, a solution has been investigated for improving the performance of LDPC codes under iterative decoding, over channels introducing bursts of erasures. These channels can be easily modelled as based on hidden Markov chains. For this reason, the mutual information has been first recalled for any general erasure channel with i.i.d. and equiprobable input symbols. Then, an LDPC codes optimization algorithm has been proposed and analyzed. The optimization algorithm simply performs permutations of variable nodes. Hence, it is very general and can be in principle applied to any given LDPC code. Numerical results have been shown for both random and eIRA codes about $L_{\text{max}}$ improvement and performance gain over a simple burst erasure channel model (CLBuEC). Our results reveal the effectiveness of the proposed algorithm.

**APPENDIX**

**LOWER BOUND ON THE DECODING FAILURE PROBABILITY**

Let $L > L_{\text{max}}$, $F$ be the event that the iterative decoder fails, $N_B$ be the (random) number of erasure bursts affecting a codeword (i.e. at least one symbol in the codeword is erased due to the burst) and $w_H(\cdot)$ be the Hamming weight function. The decoding failure probability can be bounded as

$$
\Pr\{F\} \geq \sum_{i \geq 1} \Pr\{F|N_B = i\} \Pr\{N_B = i\} = \Pr\{F|N_B = 1\} \left(1 - \Pr\{N_B = 0\}\right).
$$

The factor $\Pr\{F|N_B = 1\}$ is computed by introducing a vector $a^{(L)} = [a_{-L+1}^{(L)}, a_{-L+2}^{(L)}, \ldots, a_n^{(L)}]$ (of dimension $n + L - 1$) with the following structure: $a_j^{(L)} = 1$ if an erasure burst of length $L$ beginning in position $j$ cannot be successfully decoded, $a_j^{(L)} = 0$ otherwise. It follows

$$
\Pr\{F|N_B = 1\} = \frac{w_H(a^{(L)})}{n + L - 1}.
$$

By Markov properties it is easy to show that the probability that the generic codeword is not affected by erasure bursts is given by

$$
\Pr\{N_B = 0\} = \frac{1}{1 + b L} (1 - b)^{n-1}.
$$

Hence we finally obtain the following lower bound:

$$
\Pr\{F\} \geq \frac{w_H(a^{(L)})}{n + L - 1} \left[1 - \frac{1}{1 + b L} (1 - b)^{n-1}\right].
$$

For example, the systematic improved eIRA code of Fig. 5 has $w_H(a^{(L)}) = 40$ for $L = 630$ (see dotted line in Fig 5). The bound is very tight for small enough values of $b$.

**ACKNOWLEDGMENT**

The authors would like to thank W. E. Ryan and the anonymous reviewers for their valuable comments and constructive suggestions, which have improved the quality of this paper.

This work was supported by the European Commission under project FP6 IST-001812 PHOENIX.
REFERENCES


Fig. 1. An example of erasure channel based on a hidden Markov chain. "T": Transmitted symbol. "R": Correctly received symbol. "?": Erased symbol.

Fig. 2. Constant length burst erasure channel model.

Fig. 3. Non-resolvable erasures burst of length $L_{\text{max}}' + 1$ beginning on symbol $V_j$ within the codeword and resolvable bursts of length $L_{\text{max}}'$ beginning on symbols $V_j$ and $V_{j+1}$. 
Fig. 4. Decoding failure rate for the random (2000, 1000) irregular LDPC code before and after the application of the optimization algorithm (burst length $L = 890$).

Fig. 5. Decoding failure rate for the (2000, 1000) eIRA code, the systematic improved code and the improved code (burst lengths $L = 630$ and $L = 500$). Dotted line: lower bound on the improved IRA, for $L = 630$.

Fig. 6. Decoding failure rate for the (1111, 1000) eIRA code, systematic improved code and the improved code (burst length $L = 70$).