IRWIN AND JOAN JACOBS
CENTER FOR COMMUNICATION AND INFORMATION TECHNOLOGIES

Performance Bounds for Erasure, List and Feedback Schemes with Linear Block Codes

Eran Hof, Igal Sason,
Shlomo Shamai

CCIT Report #728
May 2009

Electronics
Computers
Communications

DEPARTMENT OF ELECTRICAL ENGINEERING
TECHNION - ISRAEL INSTITUTE OF TECHNOLOGY, HAIFA 32000, ISRAEL

Performance Bounds for Erasure, List and Feedback Schemes with Linear Block Codes

Eran Hof Igal Sason Shlomo Shamai

Department of Electrical Engineering
Technion – Israel Institute of Technology
Haifa 32000, Israel
E-mails: {hof@tx, sason@ee, sshlomo@ee}.technion.ac.il

May 6, 2009

Abstract

A message independence property and some new performance upper bounds are derived in this work for erasure, list and decision-feedback schemes with linear block codes transmitted over memoryless symmetric channels. Similarly to the classical work of Forney, this work is focused on the derivation of some Gallager-type bounds on the achievable tradeoffs for these coding schemes, where the main novelty is the suitability of the bounds for both random and structured linear block codes (or code ensembles). The bounds are applicable to finite-length codes and to the asymptotic case of infinite block length, and they are applied to low-density parity-check (LDPC) code ensembles.

Index Terms

Automatic repeat request (ARQ), erasures, error exponents, feedback, linear codes, list decoding, low-density parity-check (LDPC) codes.

I. Introduction

Exponential error bounds were derived and studied by Forney [13], referring to the following two situations:

1) A decoder is allowed not to make a decision on a received signal, or rejecting all estimates; this output is called an *erasure*. The event where the decoder makes in this case a decision on the transmitted message, and it is wrong, is called an *undetected error*.
2) A decoder is allowed to make more than one estimate of the received signal. The output of this decoder forms a list of codewords, and the event where the transmitted message is not on the list is called a *list error event*.

Throughout this paper, decoding rules for these two situations are called *generalized decoding rules* since they apply to the general setting where the decoder does not necessarily need to make a single decision about the codeword which was sent. As explained in [13], erasure and list options may be useful when the transmitted data contains some redundancy, when a feedback channel is available, or when several stages of coding (e.g., concatenation) are used.

The size of the decoded list in [13] is allowed to vary according to the received signal. This decoding rule has to be distinguished from [12], and [34] where the size of the list is predetermined and fixed.

By allowing a decoder to increase the probability of erasures in the first case, the undetected error probability can be reduced. In the second case, by allowing the decoder to increase the size of the list, the list error probability

This research was supported by the Israel Science Foundation (grant no. 1070/07), and by the European Commission in the framework of the FP7 Network of Excellence in Wireless Communications (NEWCOM++).

Igal Sason is the corresponding author (E-mail: sason@ee.technion.ac.il).

can be reduced. The optimum decoding rules with respect to these tradeoffs were provided in [13] and they were analyzed via the derivation of exponential bounds for random codes.

This paper is focused on the derivation of some Gallager-type bounds on the achievable tradeoff between these quantities, where the new bounds are useful for both random and structured linear block codes (or ensembles). These new bounds are applied to expurgated ensembles of regular low-density parity-check (LDPC) code ensembles.

Performance analysis of specific codes is often prohibitively complex. As a result, various upper and lower bounds on the decoding error probability are provided in the literature. A significant part of this analysis is devoted to the error performance under maximum-likelihood (ML) decoding (see, e.g., [24] and references therein). Lower bounds on the error exponents for fully-random block codes under generalized decoding rules, are derived in [2], [13], [21], and [31]. Achievable error exponents are provided in [28] and [29] for random codes with constant composition under some suboptimal decoding rules (note that the upper bound in [29] concerns the moments of the decoded list size). An upper bound on the error exponent under fixed-size list-decoding is provided in [26]. In contrast to the vast literature available on the error performance under ML decoding, few results are available for error performance of structured codes under generalized decoding rules. The error performance under fixed-size list-decoding is studied for specific codes in [1], [4] and [20] where the communication is assumed to take place over an AWGN channel.

The analysis of error probabilities under generalized decoding rules with erasures, enables the study of coded communications with a noiseless decision feedback. Specifically, it is assumed that decoding erasures are followed by a repeat request over a noiseless and immediate feedback channel. Such schemes are often referred to as hybrid automatic repeat request (ARQ) systems. Unlike channel capacity for discrete memoryless channels, which is not affected by feedback (see for example [7, p. 216]), a significant improvement is demonstrated in [13] for the achievable error exponents. In this respect, the reader is also referred to [16] where the error exponents of hybrid ARQ schemes with limited retransmissions are studied. The effect of feedback was also considered in [6], and it was shown to significantly reduce the block error probability for discrete memoryless channels.

In this paper, upper bounds on the error probabilities under generalized decoding rules are provided for linear block codes over memoryless symmetric channels. Both optimal and suboptimal decoding rules are considered. When variable-size list-decoding is considered, upper bounds on the expected size of the decoded list is provided. In addition, upper bounds on the list error probability are introduced for linear block codes when the size of the list is fixed. The bounds derived in this work are applicable to the performance analysis of specific codes, and code ensembles, via their (average) distance spectra. Moreover, these bounds are applicable to finite block lengths and to the asymptotic case of an infinite block length. The provided results are exemplified for two coding schemes: Fully-random linear block codes, and regular, binary and non-binary, LDPC code ensembles with finite block lengths. Applications to coded hybrid-ARQ schemes, are also studied.

This paper is structured as follows: The definitions of channel symmetry, generalized decoding rules, and some of their basic properties, are provided in Section II. New upper bounds under the generalized decoding rules in [13] are derived in Section III. Applications of these bounds to the performance analysis of coded hybrid-ARQ schemes are provided in Section IV. Error performance of suboptimal decoding and fixed-size list-decoding, are provided in Sections V and VI, respectively. Some technical details are relegated to the appendices.

II. Channel Symmetry, Generalized Decoding, and Message Independence

In this section we introduce some definitions, examples, and statements related to channel symmetry, Forney's generalized decoding rule [13], and sub-optimal versions ([2] and [13]), as well as list decoding rules ([12] and [34]). A message independence property is stated for these decoding rules, which is used for the simplification of the analysis.

Let $\mathcal{X} = \{x_0, x_1, \ldots, x_{q-1}\}$ be a given alphabet with a cardinality q. We assume an addition operation $(+)$ over the alphabet $\mathcal{X}$ for which $\{\mathcal{X}, +\}$ forms an Abelian group. Let $x_0 = 0$ be the additive identity of this group. In addition, let $\mathcal{Y}$ be a given discrete (or continuous) alphabet. We assume a memoryless channel, and denote the channel transition probability (or probability density, respectively) function by $p(y|x)$, where $x \in \mathcal{X}$ and $y \in \mathcal{Y}$.

Definition 1 (**Channel symmetry).** A memoryless channel which is characterized by a transition probability p, an input-alphabet $\mathcal{X}$ and a discrete output alphabet $\mathcal{Y}$ is *symmetric* if there exists a function $\mathcal{T} : \mathcal{Y} \times \mathcal{X} \to \mathcal{Y}$ which satisfies the following properties:

1) For every $x \in \mathcal{X}$, the function $\mathcal{T}(\cdot, x) : \mathcal{Y} \to \mathcal{Y}$ is bijective.
2) For every $x_1, x_2 \in \mathcal{X}$ and $y \in \mathcal{Y}$, the following equality holds:

$$p(y|x_1) = p(\mathcal{T}(y, x_2 - x_1)|x_2). \tag{1}$$

Remark 1. For channels whose output alphabet is continuous, an additional requirement on the mapping $\mathcal{T}$ is that its Jacobian is equal to 1.[1] In this case, the condition in (1) implies that

$$\int p(y|x_1)\, dy = \int p(\mathcal{T}(y, x_2 - x_1)|x_2)\, dy.$$

Example 1 (Memoryless binary-input output symmetric channels). Consider a memoryless binary-input output-symmetric (MBIOS) channel. Setting

$$\mathcal{T}(y, x) = \begin{cases} y & x = 0 \\ -y & x = 1 \end{cases}$$

then Definition 1 coincides with the standard definition of MBIOS channels.

Let $\mathcal{C} = \{\mathbf{x}_m\}_{m=1}^{q^k}$ be a linear block code whose generator matrix is a $k \times n$ full-rank matrix with entries over $\mathcal{X}$. The decoding rules studied in this paper are specified in terms of decision regions Λ_m, $1 \leq m \leq q^k$, which are all subsets of $\mathcal{Y}^n$. The conditional error probability of the m-th message is given by

$$P_{\text{e}|m} = \sum_{\mathbf{y} \in \Lambda_m^{\text{c}}} p(\mathbf{y}|\mathbf{x}_m) \tag{2}$$

where Λ_m forms the decision region for the m-th codeword, and the superscript 'c' stands for the complementary set. The decision region of the m-th codeword under ML decoding gets the form

$$\Lambda_m = \big\{\mathbf{y} : p(\mathbf{y}|\mathbf{x}_m) > p(\mathbf{y}|\mathbf{x}_{m'}),\ \forall\ m' \neq m\big\} \tag{3}$$

where ties are resolved randomly with equal probability. Assuming equal a-priori probabilities for the transmitted messages, the ML decoding rule minimizes the error probability given in (2). A well-known result for binary linear block codes operating over MBIOS channels is that their error probability under ML decoding is independent of the transmitted codeword. This enables a great simplification in the analysis by assuming that the all-zero codeword is transmitted. This result is generalized in [18] for non-binary linear block codes whose transmission takes place over memoryless symmetric channels with discrete input alphabet.

When generalized decoding rules are considered, the decision regions Λ_m are not necessarily disjoint nor they include all the possible received vectors. The former case corresponds to decoding rules with a possibly *variable* list-size, and the latter case corresponds to decoding with erasures. A list is produced by the decoder where the received vector may possibly belong to more than one decision region. An erasure event is declared by the decoder when the received vector does not belong to any decision region. These concepts were first introduced in [13]. When generalized decoding rules are allowed, the conditional block error probability $P_{\text{e}|m}$ in (2) stands for the probability of either an undetected error or an erasure. When the decision regions are disjoint, the conditional undetected error probability is given by

$$P_{\text{ue}|m} = \sum_{m' \neq m} \sum_{\mathbf{y} \in \Lambda_{m'}} p(\mathbf{y}|\mathbf{x}_m). \tag{4}$$

In addition, let $P_{\text{x}|m}$ denote the conditional probability of an erasure event given that $\mathbf{x}_m$ is transmitted. Then

$$P_{\text{x}|m} = P_{\text{e}|m} - P_{\text{ue}|m}.$$

In the case where list decoding is considered, the decision regions are not disjoint, and $P_{\text{ue}|m}$ as given in (4) is no longer a probability. However the RHS of (4) equals the conditional expectation of the number of incorrect codewords in the list (the same notation, $P_{\text{ue}|m}$, is used in both cases to simplify the statement of the following

[1] It is possible to use a generalized definition for both discrete and continuous output alphabets using the notion of unitary functions, as done for example in [33, Section III-A].

results). The optimum decoding rule with respect to the tradeoff between the error and the undetected error event is derived in [13].

Definition 2 (**Forney's generalized decoding).** Consider a block code over an alphabet $\mathcal{X}$, and let $\{\mathbf{x}_m\}$ denote its codebook. The generalized decoding rule is defined by the following decision regions:

$$\Lambda_m = \left\{ \mathbf{y} \in \mathcal{Y}^n : \ \frac{\Pr(\mathbf{y}, \mathbf{x}_m)}{\sum_{m' \neq m} \Pr(\mathbf{y}, \mathbf{x}_{m'})} \geq e^{nT} \right\} \tag{5}$$

where m is the index of the codeword, $T \in \mathbb{R}$ is a parameter, $\Pr(\mathbf{y}, \mathbf{x}_m)$ denotes the joint probability that $\mathbf{x}_m$ is the transmitted codeword and $\mathbf{y}$ is the received vector, and the summation is over all codewords except of $\mathbf{x}_m$.

Remark 2. The decision region in (5) can be expressed equivalently in the form

$$\Lambda_m = \left\{ \mathbf{y} \in \mathcal{Y}^n : \ \Pr(\mathbf{x}_m | \mathbf{y}) \geq \frac{e^{nT}}{1 + e^{nT}} \right\}. \tag{6}$$

Note that for $T = 0$, this decision region includes all the vectors $\mathbf{y} \in \mathcal{Y}^n$ for which $\Pr(\mathbf{x}_m|\mathbf{y}) \geq \frac{1}{2}$. The a-posteriori probability of $\mathbf{x}_m$, given that $\mathbf{y} \in \Lambda_m$ is received, is therefore larger than the a-posteriori probability for any other codeword. Hence, if a codeword is selected according to the decoder with the decision regions in (6) with $T = 0$, then the same decision is made by a MAP decoder (as no other codeword can get an a-posteriori probability larger than $\frac{1}{2}$). This implies that the undetected error exponent for the decoder in (6) with $T = 0$ cannot be smaller than the error exponent of an ML decoder with equally-likely codewords. Interestingly, as will be shown later, we get the same lower bound on the error exponents for both decoders.

Remark 3. The threshold parameter T in (5) controls the tradeoff between erasures and undetected errors (or average list size and decoding error). Setting $T > 0$ guarantees that the decision regions Λ_m are disjoint.

Proposition 1 (**Forney's generalized decoding [13]).** Assume that the decoding of a block code is carried according to the generalized decoding rule in Definition 2. Then, there is no other decoding rule that simultaneously gives a lower error probability and an undetected error probability (or an average number of incorrect codewords when list decoding is considered).

Remark 4 (**On optimal generalized decoding of convolutional codes**). Optimal generalized decoding of convolutional codes, whose transmission takes place over memoryless channels, is provided in [19]. This algorithm is based on the decision regions in (5). Specifically, the algorithm is based on a modification of the standard Viterbi algorithm, where the denominator in (5) is evaluated recursively. The optimality of the algorithm in [19] is based on the optimality in Proposition 1.

The following proposition generalizes the message independence property for the case of generalized decoding:

Proposition 2 (**Message independence property for optimal generalized decoding).** Let $\mathcal{C}$ be a linear block code whose transmission takes place over a memoryless and symmetric channel. Then, the block error probability and the undetected error probability, under the generalized decoding rule in Definition 2, are independent of the transmitted codeword.

Proof: See Appendix A. ■

Remark 5. In the case where list decoding is considered (i.e., the decision regions are not disjoint), then Proposition 2 holds when we refer to the conditional expectation of the number of incorrect messages in the list produced by the generalized decoding rule, instead of the undetected error probability.

The following suboptimal decoding rule is suggested in [13] for the case of decoding with erasures:

Definition 3 (**Likelihood Ratio (LR) Decoding).** Consider a block code over the alphabet $\mathcal{X}$, and let $\{\mathbf{x}_m\}$ denote its codebook. The LR decoding rule is defined by the following decision regions:

$$\Lambda_m^{\text{LR}} = \left\{ \mathbf{y} \in \mathcal{Y}^n : \ \frac{\Pr(\mathbf{y}, \mathbf{x}_m)}{\Pr(\mathbf{y}, \mathbf{x}_{m_2})} \geq e^{nT} \right\} \tag{7}$$

where m is a codeword index, $T > 0$ is a parameter, $\Pr(\mathbf{y}, \mathbf{x}_m)$ denotes the joint probability that $\mathbf{x}_m$ is the transmitted codeword and $\mathbf{y}$ is the received vector, and $m_2 = m_2(\mathbf{y})$ denotes the second most probable codeword for each received vector $\mathbf{y}$.

Remark 6. It is observed in [13] that the LR decoding rule may be a good approximation to the optimal regions in (5), since the second most likely codeword is usually much more probable than the rest of the codewords (excluding the most probable codeword). It is also noted in [13] that this suboptimal decoding rule is of practical utility.

Example 2 (**Suboptimal generalized decoding).** Consider the transmission of a binary linear block code over a BSC. Given a received vector $\mathbf{y} \in \{0,1\}^n$, the decoded codeword is $\mathbf{x}$ if and only if

$$d_{\text{H}}(\mathbf{x}', \mathbf{y}) - d_{\text{H}}(\mathbf{x}, \mathbf{y}) > 2\tau n \tag{8}$$

for all codewords $\mathbf{x}' \neq \mathbf{x}$, where $d_{\text{H}}(\mathbf{x}, \mathbf{y})$ denoted the Hamming distance between $\mathbf{x}$, and $\mathbf{y}$, and $\tau \geq 0$ is an arbitrary parameter. Otherwise, an erasure is declared. It is easily verified that this rule is a particular case of (7). The error exponents for this setting are studied in [2].

The following proposition obtains a message independence property for the suboptimal decoding rule in Definition 3:

Proposition 3 (**Message independence property for (suboptimal) LR decoding).** Let $\mathcal{C}$ be a linear block code whose transmission takes place over a memoryless and symmetric channel. Then, the block error probability and the undetected error probability, under the suboptimal decoding rule in (7), are independent of the transmitted codeword.

Proof: See Appendix B. ■

The following definition considers list decoding with a fixed size. Such a decoding rule is based on a fixed size of the list (instead of a variable list size which characterizes the decoding rule in Definition 2 with $T < 0$).

Definition 4 (**Fixed-size list-decoding).** Consider a block code over an alphabet $\mathcal{X}$, and let $\{\mathbf{x}_m\}$ denote its codebook. Given a fixed list size L, the list-decoder is a mapping from the set of all possible received vectors $\mathcal{Y}^n$ to the set of all possible lists of L codewords. This mapping produces the list whose likelihoods are the highest among all other codewords. That is, given a received vector $\mathbf{y}$, a codeword $\mathbf{x}_m$ is in the list if $p(\mathbf{y}|\mathbf{x}_m) > p(\mathbf{y}|\mathbf{x}_{m'})$ for all $m' \neq m$ except for at most $L-1$ other possible codewords.

Assuming that the codeword $\mathbf{x}_m$ is transmitted, a block error event is occurred by the fixed-size list-decoding rule in Definition 4, if the list produced by the decoder does not include the transmitted codeword $\mathbf{x}_m$. The following proposition is analogous to the message independence property in Propositions 2 and 3:

Proposition 4 (**Message independence property for fixed-size list-decoding).** Let $\mathcal{C}$ be a linear block code whose transmission takes place over a memoryless and symmetric channel. Then, the block error probability, under the fixed-size list-decoding is independent of the transmitted codeword.

Proof: See Appendix C. ■

III. Upper Bounds under optimal generalized decoding

The transmission of block codes (not necessarily linear) is first considered. In addition, throughout the paper, all codewords are assumed to have a uniform a-priori probability.

Proposition 5. Consider the transmission of a code $\mathcal{C}$ with a block length n and M codewords, and let $p(\mathbf{y}|\mathbf{x})$ designate the transition probability of the channel where $\mathbf{x} \in \mathcal{C}$ is the transmitted codeword and $\mathbf{y} \in \mathcal{Y}^n$ is the received vector. Then, the conditional block error probability $(P_{\text{e}|m})$ and the average undetected error probability (P_{ue}) under the generalized decoding rule in (5) satisfy

$$P_{\text{e}|m} \leq e^{nsT} D_{\text{B}}(m, G_n^m, s, \rho) \tag{9}$$

$$P_{\text{ue}} \leq e^{n(s-1)T} \frac{1}{M} \sum_{m=1}^{M} D_{\text{B}}(m, G_n^m, s, \rho) \tag{10}$$

where $0 \leq s \leq \rho \leq 1$ are real-valued parameters, G_n^m is an arbitrary non-negative function over $\mathcal{Y}^n$ which possibly depends on the codeword $\mathbf{x}_m$, $1 \leq m \leq M$, and

$$\begin{aligned} D_{\mathrm{B}}(m, G_n^m, s, \rho) \triangleq & \left(\sum_{\mathbf{y}} G_n^m(\mathbf{y}) p(\mathbf{y}|\mathbf{x}_m) \right)^{1-\rho} \\ & \left(\sum_{m' \neq m} \sum_{\mathbf{y}} p(\mathbf{y}|\mathbf{x}_m) G_n^m(\mathbf{y})^{1-\frac{1}{\rho}} \left(\frac{p(\mathbf{y}|\mathbf{x}_{m'})}{p(\mathbf{y}|\mathbf{x}_m)} \right)^{\frac{s}{\rho}} \right)^{\rho} . \end{aligned} \tag{11}$$

Proof: See Appendix D. ■

Remark 7. Bounds (9) and (10) in Proposition 5 may be considered as a generalization of the DS2 bound ([8], [25], [24]). In fact, setting $T = 0$ in (9) reproduces the DS2 bound under ML decoding. Note however that for $T = 0$, the decision regions in (5) do not coincide with those under ML decoding (e.g., in the former case there are erasures).

The following corollary is a particularization of Proposition 5 for fully random block codes whose transmission takes place over memoryless channels. The corollary reproduces the exponential upper bounds as in [13, Th. 2].

Corollary 1 (Random coding error exponents under optimum generalized decoding). Consider the transmission of block codes over a memoryless communication channel with a transition probability law p. Then, under the notation in Proposition 5, there exists a block code which simultaneously satisfies

$$P_{\mathrm{e}} \leq e^{-nE_1(R,T)} \tag{12}$$

$$P_{\mathrm{ue}} \leq e^{-nE_2(R,T)} \tag{13}$$

where $R = \ln M/n$ is the code rate (in nats per channel use),

$$E_1(R,T) \triangleq \max_{0 \leq s \leq \rho \leq 1,\ q_X} \Big(E_0(s, \rho, q_X) - \rho R - sT \Big) \tag{14}$$

$$E_2(R,T) \triangleq E_1(R,T) + T$$

$$E_0(s, \rho, q_X) \triangleq -\ln \sum_{y \in \mathcal{Y}} \left\{ \left(\sum_{x \in \mathcal{X}} q_X(x) p(y|x)^{1-s} \right) \left(\sum_{x \in \mathcal{X}} q_X(x) p(y|x)^{\frac{s}{\rho}} \right)^{\rho} \right\} \tag{15}$$

and q_X is a probability distribution over $\mathcal{X}$.

Proof: See Appendix E. ■

The bounds in Corollary 1 are derived in [13] without relying on tilting measures. The current derivation relies on the DS2 bound which makes use of tilting measures and Jensen's inequality. It is noted in [13] that setting $T = 0$ in Corollary 1, provides the random coding error exponent of Gallager [15]. Hence, as is mentioned in [13], the random coding error exponent is attainable not only under ML decoding, but also under the generalized decoding rule in (5) with $T = 0$. The following proposition is a particularization of Proposition 5 for linear block codes.

Proposition 6. Consider an (n, k) linear block code $\mathcal{C}$ whose transmission takes place over a memoryless symmetric channel. Assume that the channel input and output alphabets are $\mathcal{X}$ and $\mathcal{Y}$, respectively, and let p be the transition probability of the channel. Then, the block error probability P_{e} and the undetected error probability P_{ue} under the generalized decoding rule in (5), satisfy

$$P_{\mathrm{e}} \leq e^{nsT} D(g, s, \rho) \tag{16}$$

$$P_{\mathrm{ue}} \leq e^{-n(1-s)T} D(g, s, \rho) \tag{17}$$

where $g : \mathcal{Y} \to \mathbb{R}$ is an arbitrary non-negative real-valued function, $0 \leq s \leq \rho \leq 1$ are arbitrary parameters, and

$$D(g, s, \rho) \triangleq \left(\sum_{y \in \mathcal{Y}} g(y) p(y|0) \right)^{n(1-\rho)} \left(\sum_{m' \neq 0} \prod_{i=1}^{n} \sum_{y \in \mathcal{Y}} g(y)^{1-\frac{1}{\rho}} p(y|0) \left(\frac{p(y|x_{m',i})}{p(y|0)} \right)^{\frac{s}{\rho}} \right)^{\rho} . \tag{18}$$

Proof: See Appendix F. ■

Remark 8. When the decision regions are not disjoint (i.e., a list decoder is considered), P_{ue} in (17) does not denote a probability but the expected number of incorrect codewords in the decoded list. The block error probability P_e in (16) refers, in this case, to the list decoding error probability.

Remark 9. The parameters s and ρ in Proposition 6 may be chosen separately for the bounds in (16) and (17). However, the optimized choice of the two parameters is identical in both bounds (since they only differ in the multiplicative term e^{-nT}).

The mathematical structure of the bound provided in the following corollary is similar to the Shulman-Feder bound (SFB) in [27]. Because of this reason, this bound may be considered as a generalization of the SFB for the generalized decoding rule in (5). To simplify the notation, the corollary is provided for the case of a binary linear block code whose transmission takes place over an MBIOS channel (the generalization of the bounds to non-binary linear block codes is performed similarly to the approach in the proof of [18, Theorem 2]).

Corollary 2. Consider an (n, k) binary linear block code $\mathcal{C}$ whose transmission takes place over an MBIOS channel with a transition probability law p. Then, the block error probability P_{e} and the undetected error probability P_{ue} under the generalized decoding rule in (5) satisfy

$$P_{\text{e}} \leq e^{-n\left(E(\rho,R,\mathcal{C})-\frac{\rho T}{1+\rho}\right)} \tag{19}$$

$$P_{\text{ue}} \leq e^{-n\left(E(\rho,R,\mathcal{C})+\frac{T}{1+\rho}\right)} \tag{20}$$

where $0 \leq \rho \leq 1$ is an arbitrary real-valued parameter, $R \triangleq \left(\frac{k}{n}\right) \cdot \ln 2$ is the code rate (in nats per channel use),

$$E\left(\rho, R, \mathcal{C}\right) \triangleq E_0(\rho) - \rho\left(R + \frac{\ln(\alpha(\mathcal{C}))}{n}\right) \tag{21}$$

$$E_0\left(\rho\right) \triangleq -\ln\left(\sum_y \left(\frac{1}{2}p(y|0)^{\frac{1}{1+\rho}} + \frac{1}{2}p(y|1)^{\frac{1}{1+\rho}}\right)^{1+\rho}\right) \tag{22}$$

$$\alpha(\mathcal{C}) \triangleq \max_{1\leq i\leq n} \frac{|\mathcal{C}_i|}{2^{-(n-k)}\binom{n}{i}} \tag{23}$$

and $|\mathcal{C}_i|$ denotes the number of codewords whose Hamming weight is i.

Proof: Setting $s = \frac{\rho}{1+\rho}$, and

$$g(y) = \left(\frac{1}{2}p(y|0)^{\frac{1}{1+\rho}} + \frac{1}{2}p(y|1)^{\frac{1}{1+\rho}}\right)^{\rho} p(y|0)^{-\frac{\rho}{1+\rho}} \tag{24}$$

in the bounds of Proposition 6, the proof follows in the same way as in [24, Ch. 4.4.1]. ■

Remark 10. In the case where the performance of an ensemble of linear block codes is of interest, repeating the derivation of Corollary 2 leads to the same upper bounds as in (19) and (20), where the cardinality $|\mathcal{C}_i|$ in (23) is replaced with its statistical expectation over the considered ensemble, and the codebooks of this ensemble are chosen uniformly at random.

Example 3 (Error exponents of fully random binary linear block codes). Consider the transmission of fully random binary linear (n, k) block codes over a memoryless symmetric channel. For this particular case, the term $\alpha(\mathcal{C})$ in (23) equals 1. As a result, it follows from Corollary 2 that the exponent of the block error probability (including erasures and undetected errors), denoted by E_{e}, satisfies

$$E_{\text{e}} \geq \max_{0\leq\rho\leq 1}\left(E_0(\rho) - \rho R - \frac{\rho T}{1+\rho}\right) \tag{25}$$

where $E_0(\rho)$ is defined in (22), R is the code rate (in nats per channel use), and T is the parameter of the generalized decoding rule in Definition 2. Setting $T = 0$ in (25) reproduces the (non-expurgated) random coding error exponent of Gallager [15]. This observation was first made by Forney for the ensemble of fully random block codes [13]. The undetected error exponent, denoted by E_{ue}, satisfies

$$E_{\text{ue}} \geq T + \max_{0 \leq \rho \leq 1} \left(E_0(\rho) - \rho R - \frac{\rho T}{1+\rho} \right).$$

The lower bounds on the two error exponents are shown in Figs. 1 and 2 for the case of transmission over a BSC with a crossover probability of $p = 0.11$, and for a binary-input AWGN channel with $E_{\text{s}}/N_0 = -2.8$ dB, respectively (both values refer to the capacity limit for a rate of one-half bits per channel use). The bounds are sketched as a function of the code rate (in nats per channel use). The lower bounds on the error exponents for the case of decoding with erasures ($T \geq 0$) are provided in Figs. 1(a) and 2(a) for $T = 0, 0.025, 0.05, 0.1$ and 0.15. For the case of decoding with a variable list-size ($T < 0$), the lower bounds on the error exponents are provided in Fig. 1(b) and 2(b) for $T = 0, -0.05$, and -0.1. In addition, lower bounds on the exponent $E_N \triangleq -(\ln N)/n$, where N is the number of incorrect codewords in the decoded list, are also provided for this case. Note that the exponent E_N is negative above some rate. The figures show the region for which the exponent E_N is non-negative; the negative part of E_N, for which an upper bound on the size of the decoded list grows exponentially with the block length, is removed from these figures.

Definition 5 (Composition of a vector). Let $\mathbf{c}$ be a vector whose components are symbols in an alphabet $\mathcal{X}$ of size q. Let us assume without loss of generality that $\mathcal{X} = \{0, \ldots, q-1\}$. The composition of $\mathbf{c}$, denoted by $\mathbf{t} = \mathbf{t}(\mathbf{c})$, is a vector $\mathbf{t} = (t_0, t_1, \ldots, t_{q-1})$ where t_x (for $x \in \mathcal{X}$) denotes the number of symbols in $\mathbf{c}$ that are equal to x.

Definition 6 (Complete composition spectrum). Let $\mathcal{C}$ be a linear block code of length n over an alphabet $\mathcal{X}$. The complete composition spectrum is the sequence $\{|\mathcal{C}_{\mathbf{t}}|\}$ where $|\mathcal{C}_{\mathbf{t}}|$ is the number of codewords whose composition is $\mathbf{t}$, and $\mathbf{t}$ ranges over the set $\mathcal{H}$ of all possible compositions over $\mathcal{X}^n$.

Corollary 3. Consider an ensemble $\mathcal{E}$ of (n, k) linear block codes having the property that the average composition spectrum over all the codes $\mathcal{C}$ which are uniformly selected at random from this ensemble satisfies

$$\mathsf{E}\Big[|\mathcal{C}_{\mathbf{t}}|\Big] = P(n - t_0) \binom{n}{\mathbf{t}} \tag{26}$$

where $P(l)$ denotes the probability that a vector whose Hamming weight is l, forms a codeword in a randomly selected codebook $\mathcal{C}$. Assuming that the transmission takes place over a memoryless symmetric channel, then under the notation in Proposition 6, the block error probability P_{e} and the undetected error probability P_{ue}, satisfy

$$P_{\text{e}} \leq e^{\frac{n\rho T}{1+\rho}} \cdot D_{\text{s}}(\rho, \mathcal{C}) \tag{27}$$

$$P_{\text{ue}} \leq e^{-\frac{nT}{1+\rho}} \cdot D_{\text{s}}(\rho, \mathcal{C}) \tag{28}$$

where $0 \leq \rho \leq 1$, and

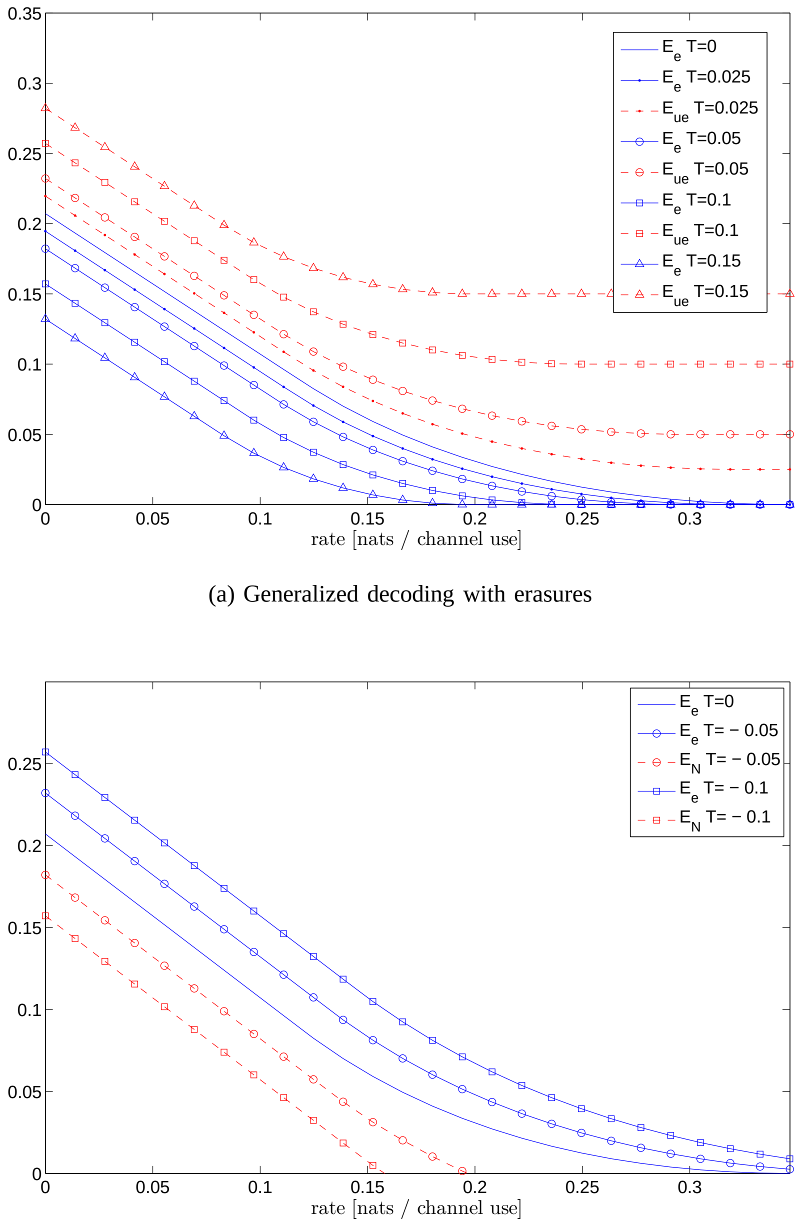
$$D_{\text{s}}(\rho, \mathcal{C}) \triangleq A(\rho)^{n(1-\rho)} \left(\sum_{1 \leq l \leq n} P(l) \binom{n}{l} B(\rho)^{n-l} C(\rho)^l \right)^{\rho} \tag{29}$$

$$A(\rho) \triangleq \sum_{y \in \mathcal{Y}} \left(\frac{1}{q} \sum_{x \in \mathcal{X}} p(y|x)^{\frac{1}{1+\rho}} \right)^{1+\rho} \tag{30}$$

$$B(\rho) \triangleq \sum_{y \in \mathcal{Y}} \left(\frac{1}{q} \sum_{x \in \mathcal{X}} p(y|x)^{\frac{1}{1+\rho}} \right)^{\rho-1} \left(\frac{1}{q} \sum_{x \in \mathcal{X}} p(y|x)^{\frac{2}{1+\rho}} \right) \tag{31}$$

$$C(\rho) \triangleq qA(\rho) - B(\rho). \tag{32}$$

Proof: Setting $s = \frac{\rho}{1+\rho}$ and choosing the tilting measure g in (24), the proof follows from Proposition 6 in the same way as in [18, Theorem 3]. ■



(a) Generalized decoding with erasures

(b) Generalized decoding with a variable-size list

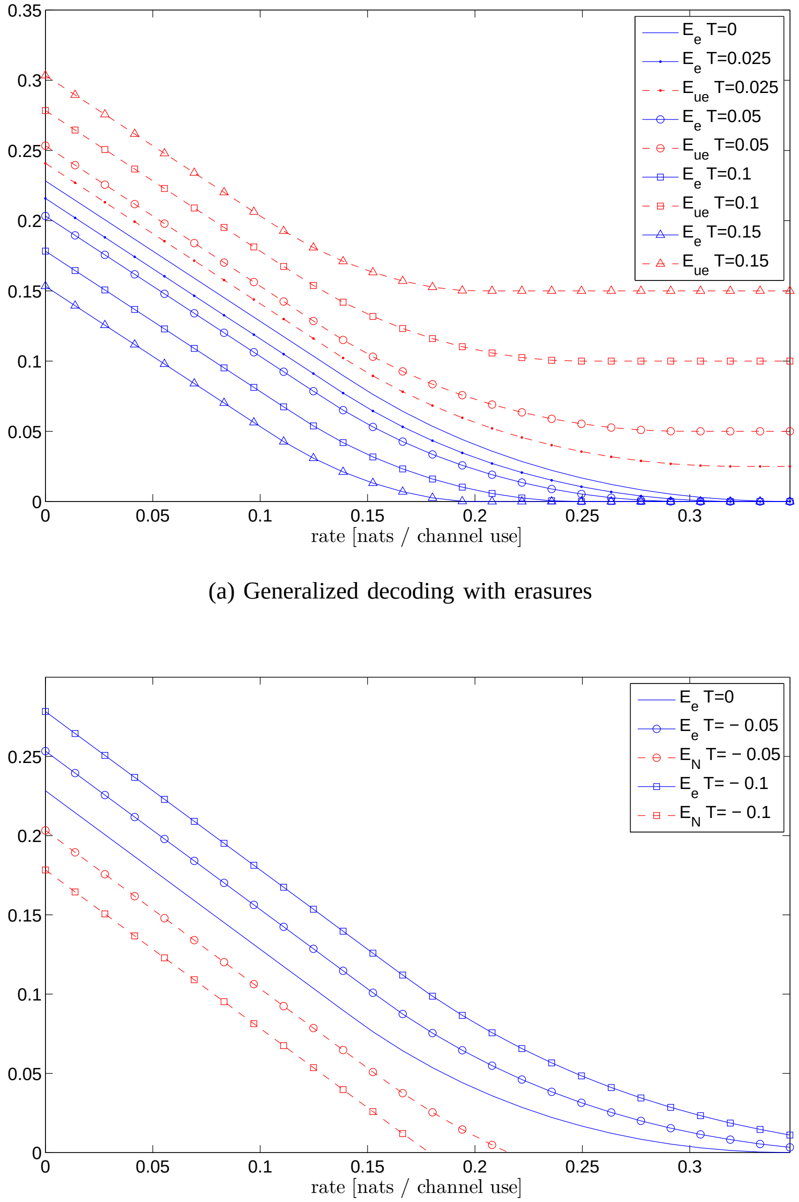
Fig. 1: Lower bounds on the error exponents and list-size exponents for the ensemble of fully-random binary linear block codes whose transmission takes place over a BSC with a crossover probability of $p = 0.11$. The lower bounds in Corollary 2 are sketched in plots (a) and (b), for the generalized decoding rule in (5) with erasures (i.e., $T \geq 0$) and with a variable list-size (i.e., $T < 0$), respectively.

Remark 11. For an ensemble of *binary* linear block codes, the condition in (26) is not mandatory. Repeating the derivation results in the same bounds as in Corollary 3 where the term $P(l)\binom{n}{l}$ in (29) is replaced with the expected complete composition spectrum of the ensemble.

Remark 12. The bounds in Corollary 3 are tighter than those in Corollary 2. Hence, for a finite block length, the bounds in Corollary 3 are more attractive even though they lack the appealing exponential structure of the bounds in Corollary 2.

Remark 13. As a particular case of Remark 7, setting $T = 0$ in (27) reproduces the upper bound on the decoding error probability of non-binary linear block codes under ML decoding in [18, Theorem 3].

The following comments concerns the numerical results shown in the examples throughout paper:



(a) Generalized decoding with erasures

(b) Generalized decoding with a variable-size list

Fig. 2: Lower bounds on the error exponents and list-size exponents for the ensemble of fully-random binary linear block codes whose transmission takes place over a binary-input AWGN channel with $E_s/N_0 = -2.8$ dB. The lower bounds in Corollary 2 are sketched in plots (a) and (b), for the generalized decoding rule in (5) with erasures (i.e., $T \geq 0$) and with a variable list-size (i.e., $T < 0$), respectively.

1) *Expurgation of codebooks*: The examples presented in this paper consider the performance of some expurgated ensembles of regular LDPC codes under generalized decoding rules. Specifically, an expurgation of the codebooks whose minimum Hamming distance is not larger than a specific value D_n is assumed. As a result, the expected complete composition spectrum $\mathsf{E}\left[|\mathcal{C}_{\mathbf{t}}|\,|d_{\min} > D_n\right]$ of a codebook which is chosen uniformly at random from the expurgated ensemble, satisfies the following upper bound:

$$\mathsf{E}\Big[|\mathcal{C}_{\mathbf{t}}|\,|d_{\min} > D_n\Big] \leq \frac{\mathsf{E}\Big[|\mathcal{C}_{\mathbf{t}}|\Big]}{1-\epsilon_n} \tag{33}$$

where $\mathsf{E}\Big[|\mathcal{C}_{\mathbf{t}}|\Big]$ is the expected composition spectrum of the original (non-expurgated) ensemble, and

$$\sum_{\mathbf{t}:\; n-t_0 \leq D_n} \mathsf{E}\Big[|\mathcal{C}_{\mathbf{t}}|\Big] \leq \epsilon_n. \tag{34}$$

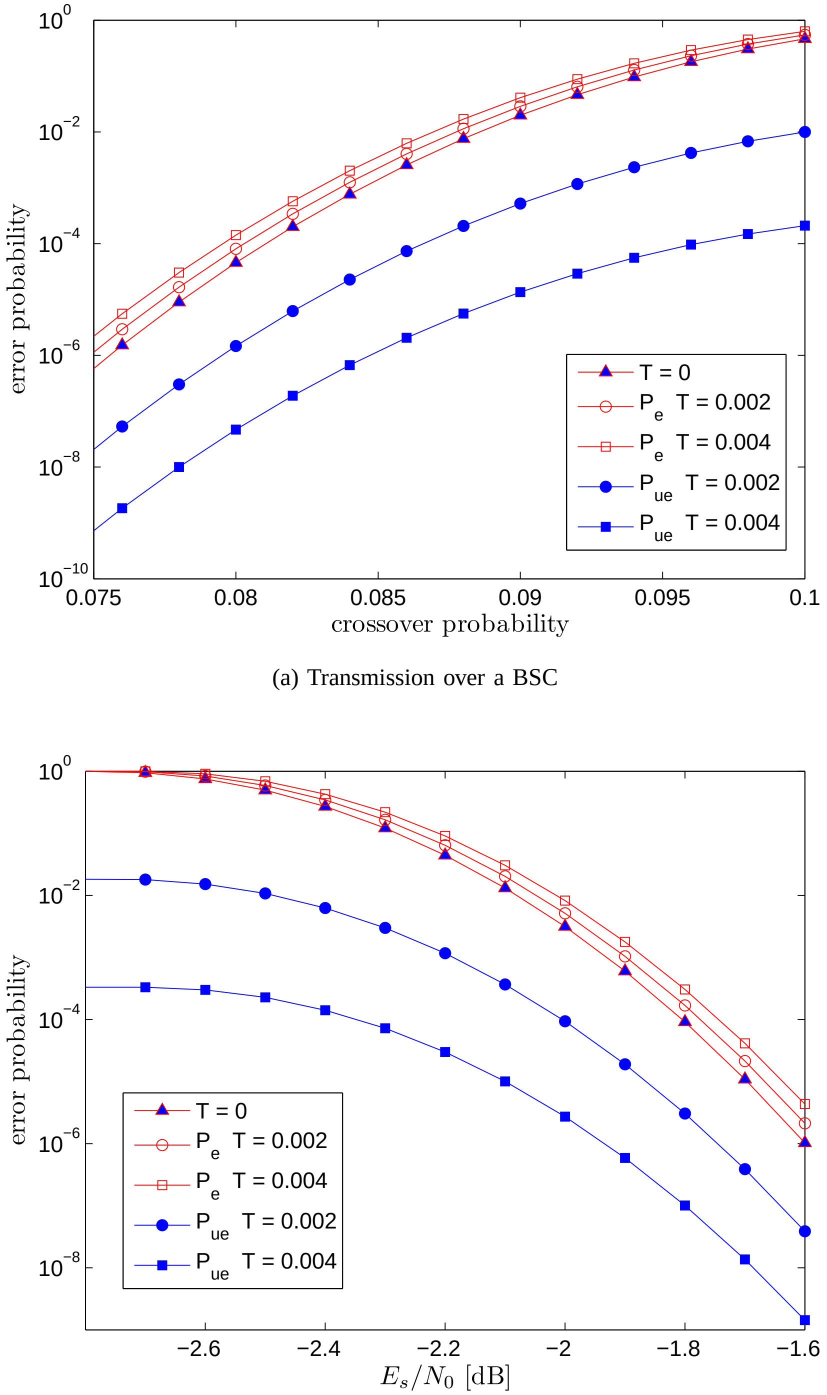
The fraction of the removed codebooks is upper bounded by ϵ_n. In the following examples, the value of ϵ_n is negligible. For the (6,12) regular binary ensemble with block lengths of $n = 504$ and 2004 bits, $\epsilon_n = 3.6002 \cdot 10^{-5}$, and $5.5058 \cdot 10^{-8}$, for $D_n = 40$ and 160 bits, respectively. For the (8,16) regular octal alphabet ensemble with a block length of $n = 1008$ symbols and $D_n = 80$ symbols, ϵ_n is around 10^{-14}.

2) *Performance over the AWGN channel*: For the AWGN channel, the results in this paper are provided as function of the signal-to-noise ratio $\frac{E_s}{N_0}$ where E_s is the energy per transmitted coded symbols, and $\frac{N_0}{2}$ is the two-sided power spectral density of the additive white noise. This comment concerns both binary and non-binary codes.

Example 4 (**Error performance of binary regular LDPC code ensembles under generalized decoding with erasures).** Consider an expurgation of the binary and regular (6,12) LDPC code ensemble of Gallager [14] with a block length of $n = 2004$ bits. In this expurgated ensemble, all the codebooks whose minimum distance is not larger than $D_n = 160$ are removed. Upper bounds on the block error probability and the undetected error probability, under Forney's generalized decoding with erasures, are studied based on Corollary 3. The composition spectrum is upper bounded via (33) and (34), where the composition spectrum of the original (non-expurgated) regular LDPC code ensemble is evaluated using the method provided in [5], [30]). The bounds are provided for several non-negative values of T in Figs. 3(a) and 3(b), assuming that the transmission takes place over a BSC and a binary-input AWGN channel, respectively. Note that if $T = 0$, the resulting bounds on the block error probability and the undetected error probability coincide, and they also provide an upper bound on the ML decoding error probability. The results indicate that by allowing an error probability that may be slightly higher than the upper bound on the error probability under ML decoding, significant improvement is guaranteed for the undetected error probability. Consider for example the error performance where the transmission takes place over a BSC with a crossover probability of 0.088. The upper bound on the error probability under ML decoding is around $7.5 \cdot 10^{-3}$ (see Figs. 3(a)). By allowing the total error probability to be less than $2 \cdot 10^{-2}$, the undetected errors are guaranteed to be less than $2 \cdot 10^{-4}$ and $5 \cdot 10^{-6}$ for $T = 0.002$ and 0.004, respectively.

Example 5 (**Error performance of binary regular LDPC code ensembles under generalized decoding with a variable-size list).** The performance of the same expurgated ensemble as in Example 4 is studied here under Forney's generalized decoding with a variable list-size. Upper bounds on the block error probability and the expected number of incorrect codewords in the list, are evaluated based on the bounds in Corollary 3 for several non-positive values of T. These bounds are provided in Figs. 4(a), and 4(b), assuming a transmission over a BSC or a binary-input AWGN channels, respectively. It is evident that only a slight improvement in the error performance is possible by using the generalized decoding rule. Take for example the case of transmission over a BSC: for crossover probabilities where the block error probability under ML decoding is below 0.09, the expected number of incorrect codewords is low. In fact, the upper bound on the expected number of incorrect codewords for such crossover probabilities, is less than one which implies that the list is likely to include only the correct codeword. However, for crossover probabilities for which the probability of the list error event is larger, the upper bound on the size of the decoded list grows considerably above 1 (see Fig. 4(a)).

Example 6 (**Generalized decoding of non-binary regular LDPC code ensembles).** Consider an expurgation of Gallager's ensemble of (8,16) regular LDPC codes [14] with an octal alphabet, and a block length of 1008 symbols. Consider the case where the expurgated ensemble excludes all the codebooks whose minimum distance is not larger than $D_n = 80$. The upper bounds on the error probabilities, under the generalized decoding rule in (5), are studied based on the upper bounds provided in Corollary 3. The (average) composition spectrum is upper bounded via (33) and (34), and the composition spectrum of the original ensemble is evaluated using the method provided in [18]. For the case of decoding with erasures, upper bounds on the block error and undetected error probabilities are provided, whereas for decoding with a variable list size, an upper bound on the expected number of incorrect codewords in the list and an upper bound on the block error probability are provided. These bounds are shown in Fig. 5(a) and 5(b), assuming that the transmission takes place over an 8-ary discrete memoryless symmetric channel, and an AWGN channel with 8-PSK modulation, respectively. It is evident that the upper bound on the block error probability for the case of decoding with erasures, referring to $T = 0.01$ in Fig. 5(a) and 5(b), slightly deteriorates as compared to the block error probability under ML decoding (where the bound presented for $T = 0$

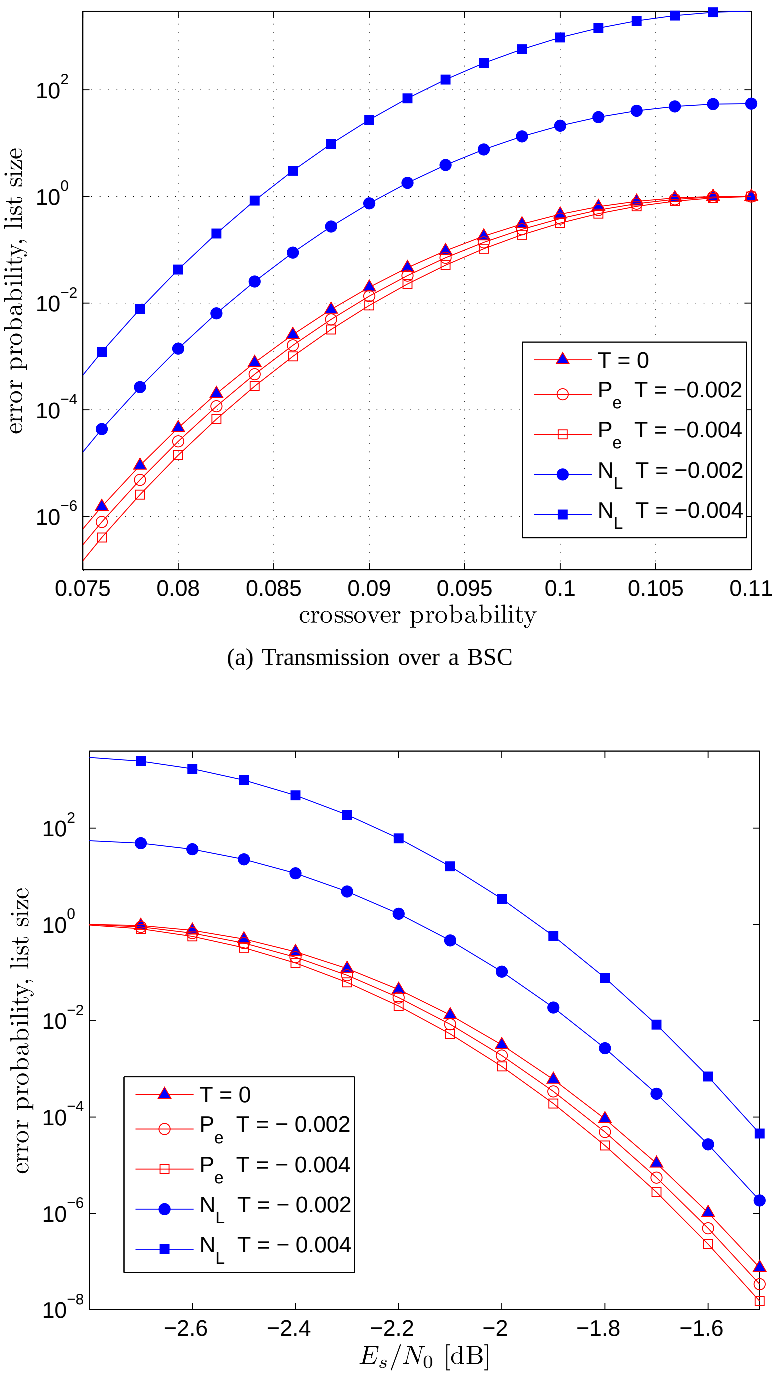


(a) Transmission over a BSC

(b) Transmission over a binary-input AWGN channel

Fig. 3: Upper bounds on the block error and undetected block error probabilities under the generalized decoding rule in (5) with erasures ($T \geq 0$). An expurgation of the binary and regular (6,12) LDPC code ensemble of Gallager is considered, where the block length is 2004 bits, and the parameter D_n which refers to the expurgation is set to 160 (see Example 4). The transmission in plots (a) and (b) is assumed to take place over a BSC, and a binary-input AWGN channels, respectively.

coincides with the bound under ML decoding). However, a remarkable improvement is shown in these figures with resect to the undetected error probability (referring to P_{ue} for $T = 0.01$ in both figures). For the variable-size list decoding which refers to $T = 0.01$ in (5), only a slight improvement is provided in the probability of error.



(a) Transmission over a BSC

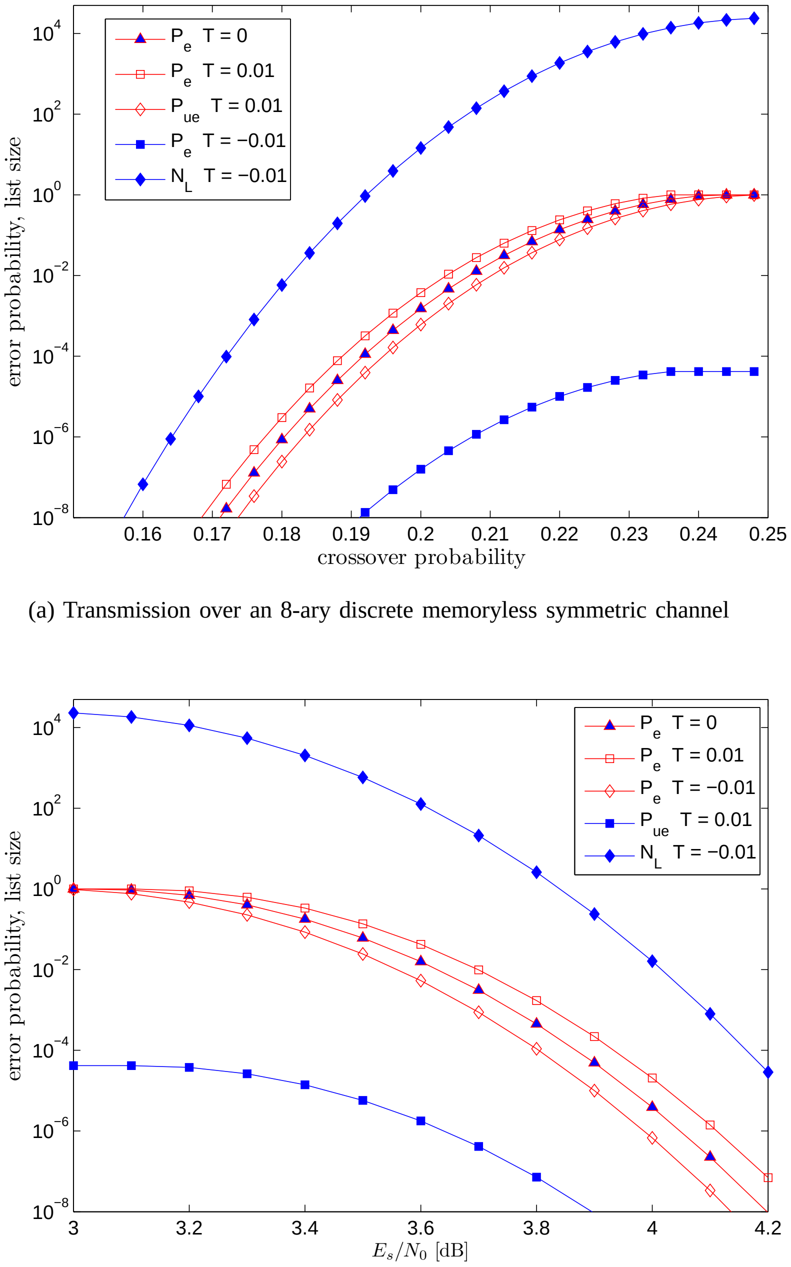
(b) Transmission over a binary-input AWGN channel

Fig. 4: Upper bounds on the block error probability and expected size of incorrect codewords in the decoded list, under the generalized decoding rule in (5) with variable-size list $(T \leq 0)$. An expurgation of the binary and regular (6,12) LDPC code ensemble of Gallager is considered, where the block length is 2004 bits, and the parameter D_n which refers to the expurgation is set to 160 (see Example 4). The transmission in plots (a) and (b) is assumed to take place over a BSC, and a binary-input AWGN channels, respectively.

IV. Applications to performance analysis of hybrid-ARQ systems

A. Preliminaries

Coded communication systems with one-bit noiseless feedback are considered where a generalized decoding rule with erasures is applied at the receiver. Each decoding erasure is communicated via the feedback to the transmitter,



(a) Transmission over an 8-ary discrete memoryless symmetric channel

(b) Transmission over an AWGN with 8-PSK modulation

Fig. 5: Upper bounds on the decoding error probabilities and number of incorrect codewords in the decoded list for an expurgated ensemble of LDPC codes. The considered ensemble refers to the octal-alphabet regular (8,16) LDPC code ensemble of Gallager with a block length of 1008 symbols, and where the parameter D_n which refers to the expurgation is set to 80 (see Example 6). The upper bounds in Corollary 3 are provided in plots (a) and (b), assuming that the transmission takes place over an 8-ary discrete memoryless symmetric channel, and an AWGN channel with 8-ary PSK modulation, respectively.

which then retransmits its message. It is first assumed that each transmitted block is decoded separately. Such a hybrid-ARQ system is described and studied in [13], where the error exponents for random coding are provided. For the case where deadlines are assumed, the error exponents for random coding are provided in [16]. The following discussion is provided in [13] and [16], and it is surveyed here for the sake of completeness.

Since Forney's generalized decoding rule (5) with a positive value of T is used in the context of erasures, the resulting decision regions at the receiver are disjoint, and the erasure probability P_{x} for a single block transmission is given by

$$P_{\text{x}} = P_{\text{e}} - P_{\text{ue}}$$

where P_{e} and P_{ue} are, respectively, the (total) block error probability and undetected error probability for a single block transmission. The erasure probability is studied via an upper bound on the error probability P_{e}. Assuming a noiseless and immediate feedback, for the case where no deadlines are considered, the expected rate of the considered system equals

$$(1 - P_{\text{x}})R \tag{35}$$

where R is the rate of the codebook used (in units of bits per channel use) for a single block transmission. The error probability of this scheme is given by

$$\frac{P_{\text{ue}}}{1 - P_{\text{x}}}. \tag{36}$$

Note that the replacement of P_{x} in (35) and (36) with an upper bound on P_{e}, provides a lower bound on the expected rate and an upper bound on the error probability.

For the case where deadlines are considered, let Q ($Q \geq 1$) be the maximal number of block retransmissions (including the first transmitted block). Each transmitted block is decoded separately using Forney's generalized decoding rule with erasures. Such a scheme is termed memoryless in [16] (note that the ARQ scheme without deadlines, studied in [13], is also memoryless in this sense). In cases where Q consequent block transmissions occur, then the generalized decoding rule is replaced for the last (Q-th) retransmitted block with an ML decoder. As a result, the expected rate and error probability, denoted by $R(Q)$ and $P_{\text{e}}(Q)$, respectively, satisfy

$$\begin{aligned} R(Q) &= \frac{R}{\sum_{k=0}^{Q-1} (P_{\text{x}})^k} \\ &= \frac{R\,(1 - P_{\text{x}})}{1 - (P_{\text{x}})^Q} \end{aligned} \tag{37}$$

and

$$\begin{aligned} P_{\text{e}}(Q) &= \sum_{k=1}^{Q-1} (P_{\text{x}})^{k-1}\, P_{\text{ue}} + (P_{\text{x}})^{Q-1}\, P_{\text{e}}^{\text{ML}} \\ &= \frac{\left(1 - (P_{\text{x}})^{Q-1}\right)\, P_{\text{ue}}}{1 - P_{\text{x}}} + (P_{\text{x}})^{Q-1}\, P_{\text{e}}^{\text{ML}} \end{aligned} \tag{38}$$

where P_{e}^{ML} is the block error probability under ML decoding for the considered code (while referring to the decoding of the last retransmitted block separately). Note that in the limit where $Q \to \infty$ (no deadlines), then (37) and (38) tend asymptotically to (35) and (36), respectively. Replacing P_{x} in (37) and (38) with an upper bound on the (total) error probability P_{e}, results in a lower bound on the expected rate, and an upper bound on the error probability, respectively.

In hybrid incremental-redundancy ARQ schemes, a repeat request triggers the transmission of a new block of n coded symbols which is not necessarily equal to the former block (even though the transmission of the same message is concerned). The decoder, instead of processing only the last block, decodes the message by observing the entire blocks received so far for the concerned message. For such cases, the expected rate, denoted by $R^{\text{IR}}(Q)$, satisfies the following lower bound [16, Eq. (24)]:

$$R^{\text{IR}}(Q) \geq \frac{R}{1 + (Q-1)P_{\text{x}}}. \tag{39}$$

This bound coincides with (37) if $Q = 2$. However, for $Q > 2$, the bound in (39) is loosened because of the specific derivation used in [16]. Assuming that an ML decoder is used after the last retransmitted block, the error probability for the IR-ARQ scheme, denoted by $P_{\text{e}}^{\text{IR}}(Q)$, is upper bounded by [16, Eq. (25)]:

$$P_{\text{e}}^{\text{IR}}(Q) \leq \sum_{k=1}^{Q-1} P_{\text{ue}}(k) + P_{\text{e}}^{\text{ML}}(Q) \tag{40}$$

where $P_{\text{ue}}(k)$ denotes the undetected error probability of the generalized decoding rule, which operates on the received observations of k consequent transmitted blocks ($1 \leq k \leq Q-1$), and $P_{\text{e}}^{\text{ML}}(Q)$ denotes the error probability under ML decoding, based on the entire transmission of Q blocks (the ML decoder is used only if Q blocks are needed to be transmitted for the same message). Note that the dominant summand in (40) is $P_{\text{ue}}(1)$, i.e., the undetected error probability of the first transmitted block.

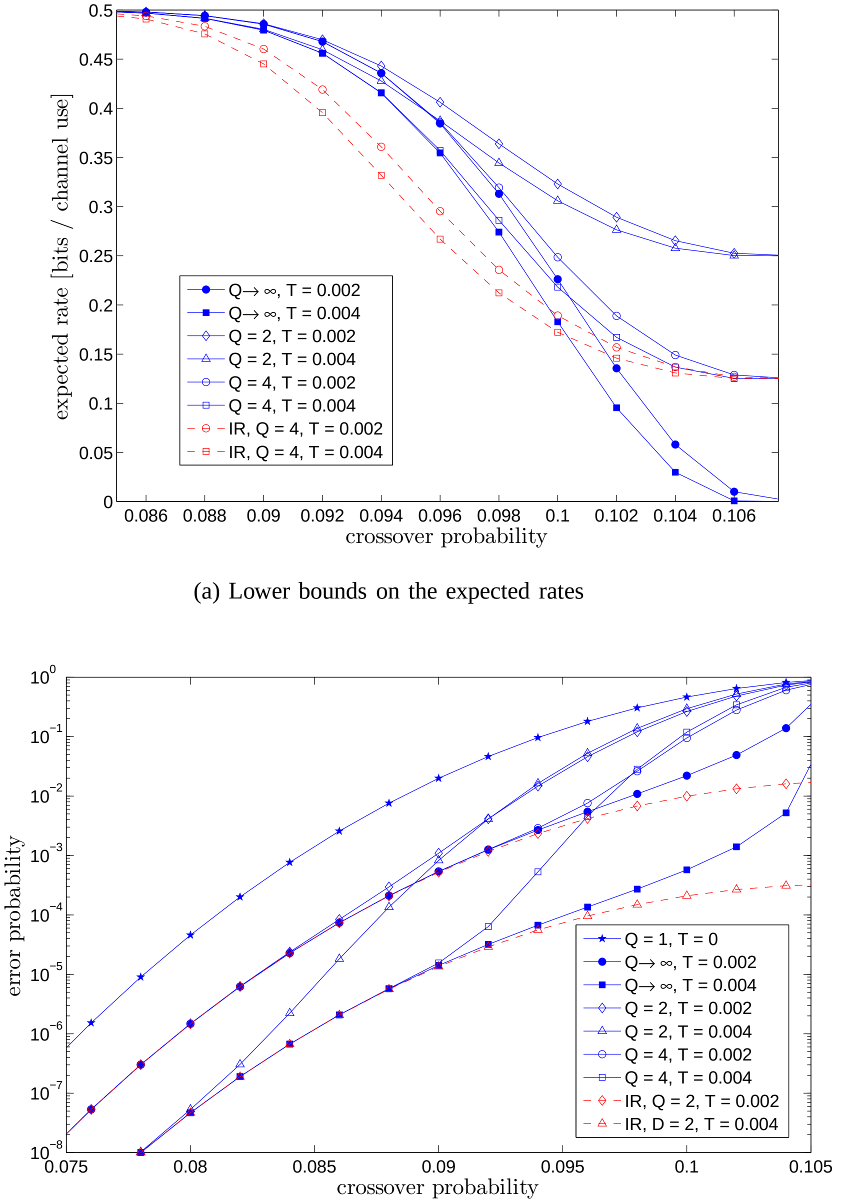
B. Examples

In the following examples, upper bounds on the error performance and lower bounds on the expected rates of some hybrid-ARQ systems are studied. These bounds are based on the bounds in Corollary 3 and the results in Section IV-A. As mentioned, each block of coded symbols in the IR-ARQ scheme may include new coded symbols. Nevertheless, for all examples in this section where IR-ARQ schemes are considered, a retransmission of equal coded blocks is assumed.

Example 7 (**Hybrid-ARQ schemes over BSC).** Consider the expurgated ensemble of binary regular LDPC codes in Example 4, whose transmission takes place over a BSC. Lower bounds on the expected rates are presented for several values of the decoding parameter T in Fig. 6(a). For memoryless systems without deadlines, the provided lower bound on the expected rate in (35) drops to zero as the crossover probability of the BSC approaches the capacity limit (which is 0.11 for a design rate of $R = \frac{1}{2}$ bits per channel use). For schemes with deadlines of $Q = 2$ and 4 transmissions, the lower bounds on the expected rate in (37) drop to $\frac{R}{Q} = \frac{1}{4}$ and $\frac{1}{8}$, respectively, as the crossover probability of the BSC approaches the capacity limit (which is the limit of (37) when we let P_{x} tend to 1). Schemes with incremental redundancy are also considered. Note that the lower bound on the expected rates for memoryless schemes with deadline of $Q = 2$, also applies to schemes with incremental redundancy, the lower bound in (39) coincides with the equality in (37) for $Q = 2$. For the case of $Q = 4$, the loosened lower bound on the expected rate for incremental redundancy schemes in (39) is also provided. Upper bounds on the decoding error probabilities for the considered schemes are provided in Fig. 6(b). The upper bound for a block error probability with $T = 0$ and where no feedback is available (a single transmission, $Q = 1$) is also provided. Note that this bound is valid for the block error probability under ML decoding. Comparing this upper bound (for $T = 0$ and $Q = 1$), with the upper bounds for $T = 0.002$ and 0.004, shows that the introduction of one-bit immediate and noiseless feedback allows for a considerable improvements in the error performance. This improvement is achieved while maintaining reasonable rate drops (at least for crossover probabilities below the threshold for which the rate starts dropping considerably). Moreover, the improvement is of interest even for the simplified memoryless-ARQ schemes with moderate deadlines (of $Q = 2$ and 4 block transmissions).

Example 8 (**Hybrid-ARQ schemes over binary-input AWGN channels).** Consider the expurgated, binary, and regular LDPC code ensemble in Example 4, and the hybrid-ARQ scheme used in Example 7. Lower bounds on the expected rates, and upper bounds on the error probabilities for such schemes are provided in Figs. 7(a), and 7(b), respectively, assuming that transmission takes place over a binary-input AWGN channel. The results show that if the SNR is above a threshold for which the expected rate does not deteriorate considerably, a substantial improvement in the decoding error probability is possible. This improvement is achieved while maintaining a negligible rate loss, even for the simplified memoryless schemes with moderate deadlines (e.g., $Q = 2$ and 4). Take for example the case where $E_{\text{s}}/N_0 = -2.1$ dB. For this setting, the upper bound on the error probability under ML decoding without retransmissions ($T = 0$, $Q = 1$) is slightly above 10^{-2}. By introducing a one bit noiseless feedback, the upper bounds on the error probability for all considered schemes with $T = 0.004$ are in the range of $10^{-4} - -10^{-5}$ while maintaining a small rate loss (the rate loss for the memoryless scheme with deadlines of $Q = 2$ transmissions is below 3.2%).

Example 9 (**Hybrid-ARQ schemes over AWGN channels with non-binary LDPC codes).** Hybrid ARQ schemes over the AWGN channel with 8-PSK modulation is considered where the expurgated and octal-alphabet LDPC code ensemble in Example 6 is used. Lower bounds on the expected rate and upper bounds on the decoding error probability are shown in Figs. 8(a) and 8(b), respectively. Schemes with and without deadlines are considered. The results show that the lower bounds on the expected rates drop considerably, below $E_{\text{s}}/N_0 = 3.6$ dB. However, above this SNR, the introduction of a single-bit, noiseless and immediate feedback allows to achieve remarkable



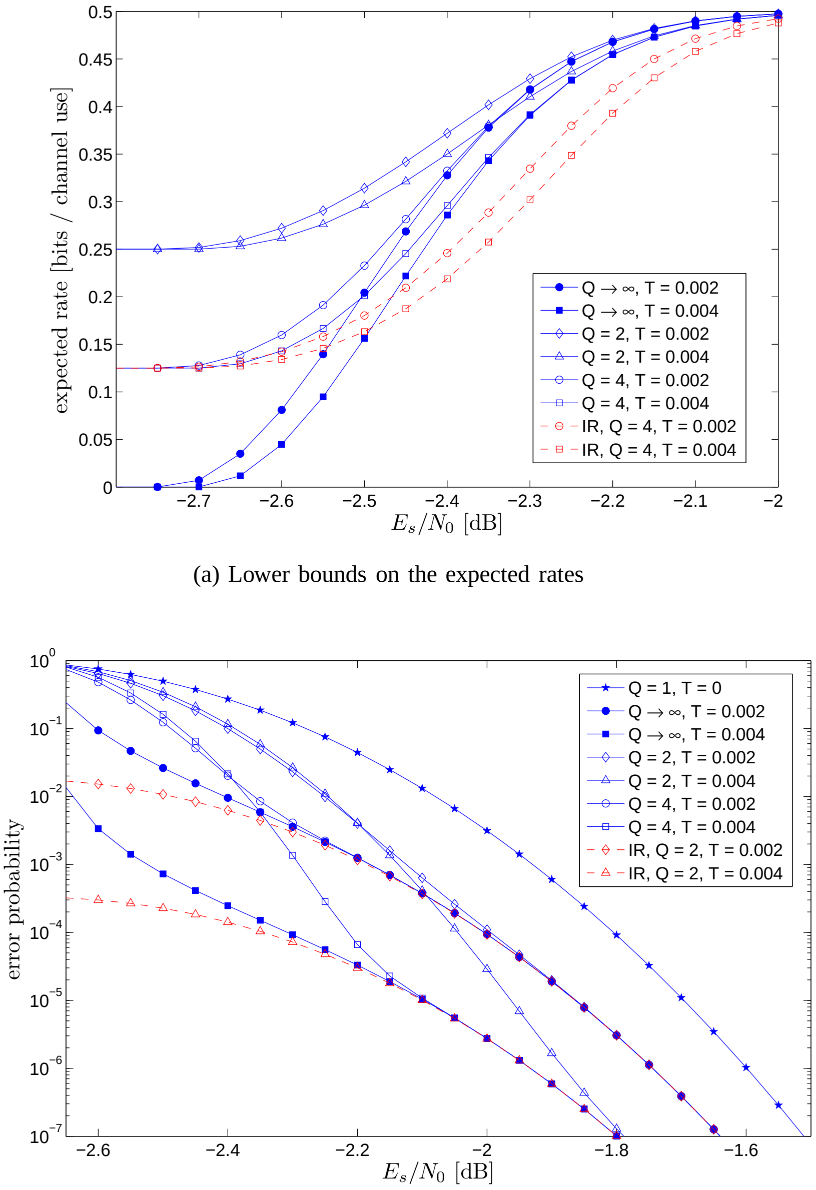
(a) Lower bounds on the expected rates

(b) Upper bounds on the error probability

Fig. 6: Performance bounds of hybrid-ARQ schemes for the expurgated, binary and regular (6,12) LDPC code ensemble of Gallager with a block length of $n = 2004$ bits (see Example 4). The transmissions are assumed to take place over the BSC. In plot (a), lower bounds on the expected rates for memoryless hybrid-ARQ schemes with and without deadlines (see (37), and (35), respectively) are shown for $T = 0.002$ and 0.004 (and deadlines of $Q = 2$ and 4 transmissions). In plot (b), upper bounds on the error probability are provided for the considered schemes. For the case of $Q = 2$, lower bounds on the expected rate and upper bounds on the decoding error probability are also provided in plots (a) and (b), respectively, assuming incremental-redundancy ARQ at the decoder (see (39)).

improvements in the error performance. Take for example the case where $E_{\text{s}}/N_0 = 3.62$ dB where the upper bound on the error probability under ML decoding without feedback (see the curve for $T = 0$ and $Q = 1$) is around 10^{-2}. For the same channel, if no deadlines are assumed, the upper bounds on the error probability are around $2 \cdot 10^{-6}$. When deadlines of $Q = 2$ and 4 total retransmissions (including the first transmission) are assumed, the upper bounds on the error probability for the same channel are $6 \cdot 10^{-4}$ and $3 \cdot 10^{-6}$, respectively. For all considered schemes, the expected rate deteriorates at this point by no more than 4%.

Immediate and noiseless one-bit feedback is assumed in Examples 7-9. The restriction to immediate feedback is loosened in most network applications where some sort of a multiple-access protocol is introduced. As a result of



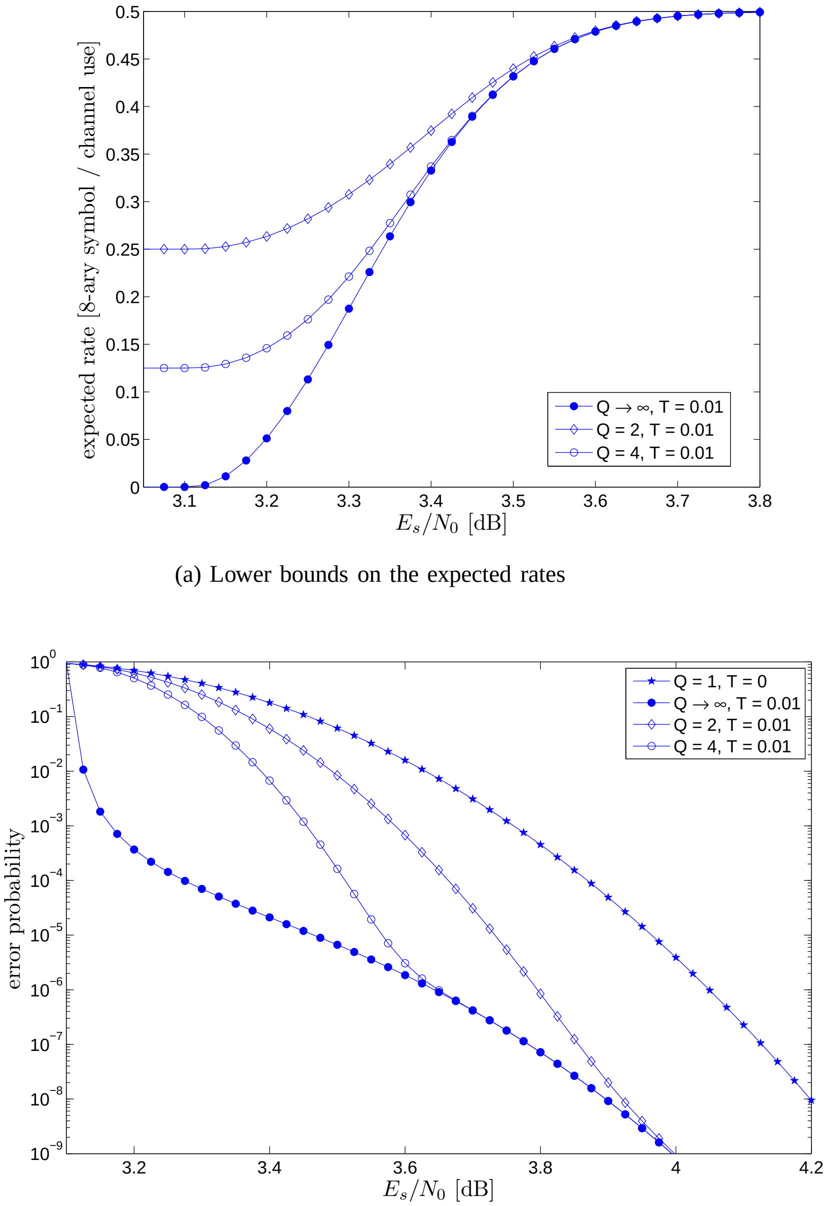
(a) Lower bounds on the expected rates

(b) Upper bounds on the error probability

Fig. 7: Performance bounds of hybrid-ARQ schemes for the expurgated, binary and regular (6,12) LDPC code ensemble of Gallager with a block length of $n = 2004$ bits (see Example 4). The transmissions are assumed to take place over binary-input AWGN channels. In plot (a), lower bounds on the expected rates for memoryless hybrid-ARQ schemes with and without deadlines (see (37), and (35), respectively) are shown for $T = 0.002$ and 0.004 (and deadlines of $Q = 2$ and 4 transmissions). In plot (b), upper bounds on the error probability are provided for the considered schemes. For the case of $Q = 2$, the lower bounds on the expected rate and upper bounds on the decoding error probability are also provided in plots (a) and (b), respectively, assuming incremental-redundancy ARQ at the decoder (see (39)).

the applied protocol, the transmitter is informed regarding the one-bit feedback with some delay that is guaranteed (by the protocol) to be before the next time slot of the retransmission. As for the condition of noiseless feedback, loosening this condition results in an inevitable synchronization errors (see, e.g., a similar observation in [9]). Since the hybrid-ARQ schemes presented in this section require only one-bit feedback, even if these synchronization errors should be kept low in comparison with the block error performance, they are typically achievable with relatively low resources.

All feedback schemes in this section assumed a one-bit (ACK/NACK) feedback. It is interesting to compare these results to the potential gain achieved for systems where the available feedback supports more than binary signaling. The following example compares the performance of Forney's scheme (without deadlines) over the binary-input



(a) Lower bounds on the expected rates

(b) Upper bounds on the error probability

Fig. 8: Performance bounds of hybrid-ARQ schemes based on an expurgated, octal-alphabet and regular (8,16) LDPC code ensemble with a block length of $n = 1008$ symbols (see Example 6). The transmission is assumed to take place over an AWGN channel with 8-PSK modulation. In plot (a), lower bounds on the expected rates for memoryless hybrid-ARQ schemes with and without deadlines (see (37), and (35), respectively) are shown for $T = 0.01$ (and possible deadlines of $Q = 2$ and 4 transmissions). In plot (b) upper bounds on the error probability are provided for the considered schemes.

AWGN channel in Example 8 to the performance of a scheme due to Yamamoto and Itoh [35]:

Example 10 (**A comparison to the Yamamoto-Itoh scheme).** Consider the expurgated binary and regular LDPC code ensembles of Gallager, and the memoryless hybrid-ARQ scheme without feedback in Example 8. The lower bounds on the expected rates and upper bounds on the error probabilities of memoryless ARQ schemes without deadlines (Forney's scheme) are compared in Figs. 9(a) and. 9(b), respectively, to those of a Yamamoto-Itoh scheme [35] (based on the same code ensemble). The Yamamoto-Itoh scheme is based on the existence of an immediate and noiseless feedback that allows the receiver to send back to the transmitter its decoded message (we assume that ML decoding is performed at the receiver). Each cycle of transmission is divided into two stages: in the first (the message mode), the transmitter sends the coded information to the receiver. Then, based on the feedback, that is the decoded message, the transmitter sends a control signal informing the receiver if the decoding is correct or not (the control mode). In the latter case, if the decoding was unsuccessful, a retransmission of the message is applied in the next cycle. In this example, we denote by λ the fraction of the cycle that is provided for the message mode (i.e., a fraction $1-\lambda$ is provided for the control mode). For example, if $\lambda = 0.96$ and the ensemble block length is $n = 504$ bits, than additional $\left(\frac{1-\lambda}{\lambda}\right) n = 21$ control mode bits are appended to the coded information in each cycle of transmission. We assume a BPSK signaling in the control mode. We adapt our notation to the setting of the Yamamoto-Itoh scheme, and get that the probability for an erroneous decoding in the control mode $(P_{\text{e,cm}})$ is given by

$$P_{\text{e,cm}} = Q\left(\sqrt{\frac{2(1-\lambda)nE_{\text{s}}}{\lambda\, N_0}}\right)$$

where

$$Q(x) \triangleq \frac{1}{\sqrt{2\pi}} \int_x^\infty e^{-\frac{t^2}{2}} dt$$

denotes the complementary Gaussian cumulative distribution function. Let P_{e}^{ML} denote the decoding error probability of a single block under ML decoding (referring to the message mode). The probability of undetected decoding results in where there is an error in the ML decoding of the message, and the the decoding in the control mode has failed, so

$$P_{\text{ue}} = P_{\text{e}}^{\text{ML}} \cdot P_{\text{e,cm}}.$$

A retransmission of a message occurs when either the ML decoding of the message is correct but the decoding in the control mode is wrong or vice versa. Since these two parts are decoded separately and the channel is memoryless, than the probability of retransmission is given by

$$\begin{aligned} P_{\text{x}} &= \left(1 - P_{\text{e}}^{\text{ML}}\right) P_{\text{e,cm}} + P_{\text{e}}^{\text{ML}} \left(1 - P_{\text{e,cm}}\right) \\ &\leq P_{\text{e}}^{\text{ML}} + P_{\text{e,cm}}. \end{aligned}$$

Replacing P_{x} with its above upper bound gives a lower bound on the expected rate in (35), and an upper bound on the decoding error probability in (36). We rely on these bounds for studying performance bounds related to the Yamamoto-Itoh scheme, and these bounds are shown in Fig. 9 when this scheme is incorporated with an expurgated ensemble of binary and regular LDPC codes. The results show, as expected, that the additional feedback resources allows for a considerable improvement in error performance.

V. Upper Bounds under suboptimal decoding with erasures

In this section, upper bounds on decoding error probabilities are derived for the suboptimal decoding rule in (7).

Proposition 7. Consider the transmission of a block code $\mathcal{C}$ of block length n and M codewords, and let $p(\mathbf{y}|\mathbf{x})$ designate the transition probability of the channel where $\mathbf{x} \in \mathcal{C}$ is the transmitted codeword and $\mathbf{y} \in \mathcal{Y}^n$ is the received vector. Then, the conditional block error probability $P_{\text{e}|m}$, and the conditional undetected error probability $P_{\text{ue}|m}$, under the suboptimal decoding rule in (7) satisfy

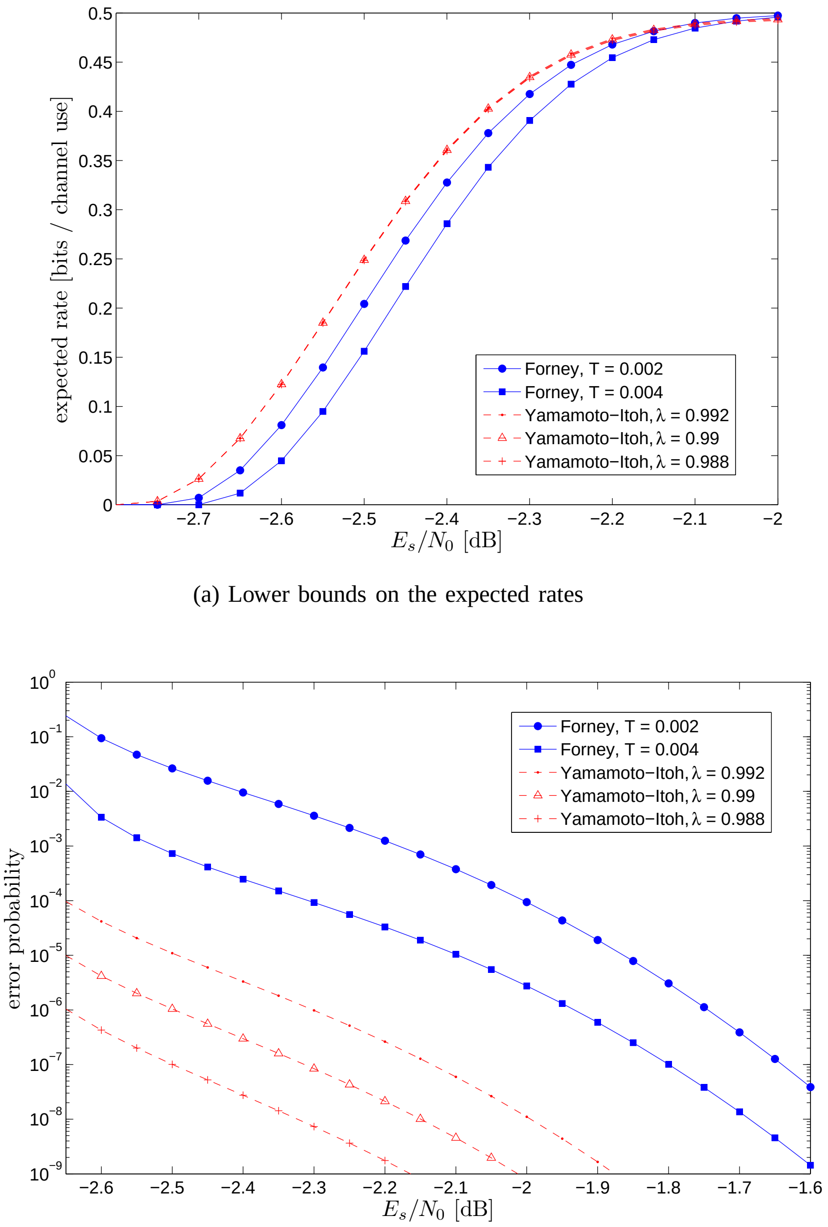
$$P_{\text{e}|m} \leq e^{nsT} D_{\text{B}}(m, G_n^m, s, \rho), \;\; 0 \leq s \leq \rho \leq 1 \tag{41}$$

$$P_{\text{ue}|m} \leq e^{-nsT} D_{\text{B}}(m, G_n^m, s, \rho), \;\; 0 \leq s \leq \rho \leq 1 \tag{42}$$

where $D_{\text{B}}(m, G_n^m, s, \rho)$ is defined in (11), and G_n^m is an arbitrary non-negative function over $\mathcal{Y}^n$ which possibly depends on the codeword $\mathbf{x}_m$, $1 \leq m \leq M$.

Proof: See Appendix G. ■

Remark 14. The upper bound on the block error probability in (41) coincides with the upper bound on the total error probability provided in (9) under the optimal generalized decoding rule. On the other hand, the upper bounds on the undetected error probabilities under the optimal and suboptimal decoding rules in (10) and (42), respectively, are different.



(a) Lower bounds on the expected rates

(b) Upper bounds on the error probability

Fig. 9: Lower bounds on the expected rate, and upper bounds on the decoding error probability, of some Yamamoto-Itoh schemes. The schemes make use of the expurgated, binary and regular (6,12) LDPC code ensemble of Gallager in Example 4, and the transmission is assumed to take place over a binary-input AWGN channel with BPSK signaling. The bounds of Forney's memoryless schemes in Example 8 are also provided for comparison.

The following corollary is a particularization of Proposition 7 for the ensemble of fully random block codes of length n and rate R whose transmission takes place over memoryless channels:

Corollary 4. Consider the transmission of block codes over a memoryless communication channel. Then, there exists a block code satisfying

$$P_{\mathrm{e}} \leq e^{-nE_1(R,T)}$$
$$P_{\mathrm{ue}} \leq e^{-nE_2^*(R,T)}$$

where $R \triangleq \frac{\ln M}{n}$ is the code rate (in nats per channel use), $E_1(R,T)$ is defined in (14),

$$E_2^*(R,T) \triangleq \max_{0 \leq s \leq \rho \leq 1,\ q_X} \Big(E_0(s, \rho, q_X) - \rho R + sT \Big)$$

E_0 is as defined in (15), and q_X is an arbitrary probability distribution over $\mathcal{X}$.

Proof: The proof follows the same arguments as the proof of Corollary 1. ■

The following bound is provided for the case of binary linear block codes whose transmission takes place over an MBIOS channel (the generalization of the bound to non-binary linear block codes, as provided in [18], is direct):

Corollary 5. Consider an (n,k) binary linear block code $\mathcal{C}$ whose transmission takes place over an MBIOS channel with a transition probability law p. Then the block error probability P_{e}, and the undetected error probability P_{ue}, under the generalized decoding rule in (7) satisfy

$$P_{\text{e}} \leq e^{-n\left(E(\rho,R,\mathcal{C}) - \frac{\rho T}{1+\rho}\right)}, \ 0 \leq \rho \leq 1 \tag{43}$$

$$P_{\text{ue}} \leq e^{-n\left(E(\rho,R,\mathcal{C}) + \frac{\rho T}{1+\rho}\right)}, \ 0 \leq \rho \leq 1 \tag{44}$$

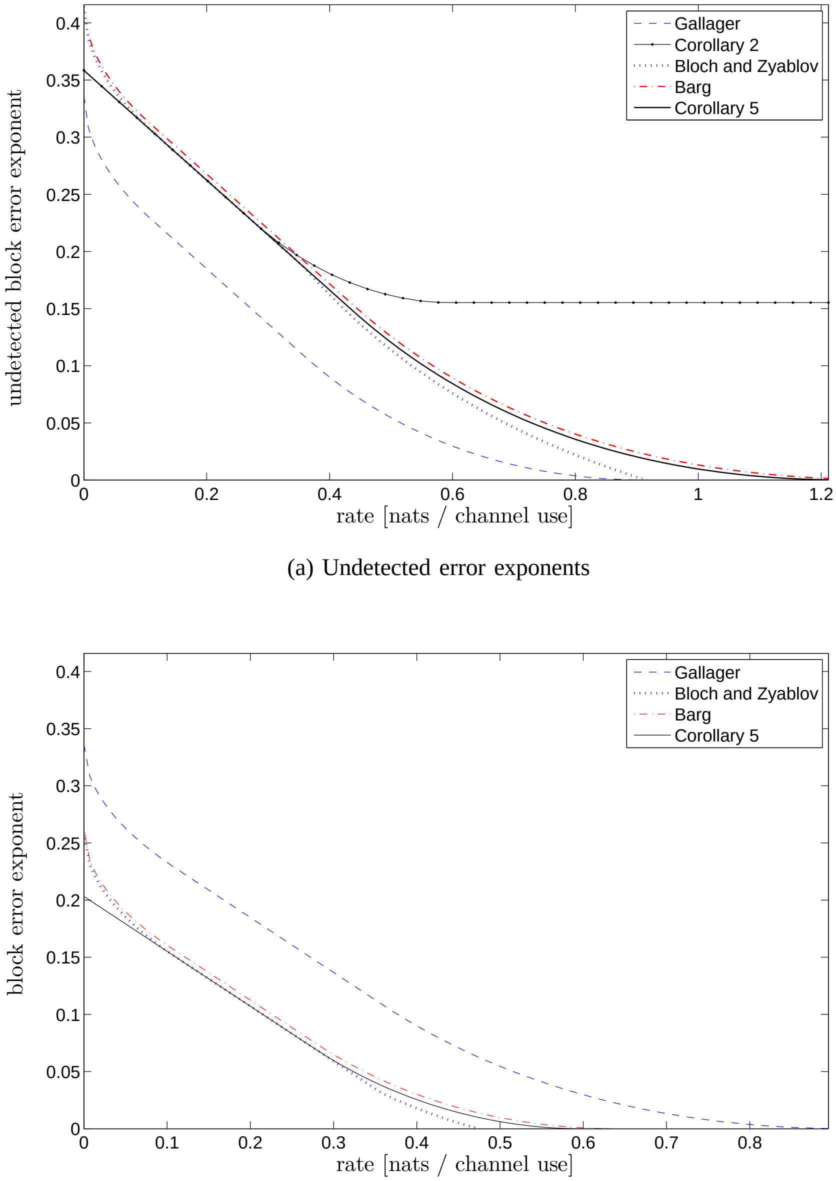
where R is the code rate (in nats per channel use), and $E\left(\rho, R, \mathcal{C}\right)$ is defined in (21).

Proof: The proof follows from Proposition 7, and its derivation is similar to the way where Corollary 2 is derived from Proposition 6. ■

Remark 15. As in Corollary 2, the bounds of Corollary 5 resemble to the SFB, and they may therefore be considered as a generalization of the SFB for the case at hand.

Remark 16. For all rates below some (finite) rate thresholds, the bounds in Corollary 5 on the decoding error for linear block codes under the suboptimal LR rule in Definition 3, coincide with those under the optimal decoding rule in Definition 2. To see this, observe first that the upper bounds in (19) and (43) are identical. It is left to consider the upper bounds in (20) and (44) on the undetected error probability. Note first that $E_0(\rho) - \rho R$ (E_0 is defined in (22)) is a concave function of $0 \leq \rho \leq 1$, and it is optimized for rates below $E_0'(1)$ at $\rho = 1$ (see, e.g., [32, p. 135]). Moreover, $\frac{\rho}{1+\rho}$ is a monotonic increasing function of $0 \leq \rho \leq 1$. This implies that if $\frac{T}{4} < E_0'(1)$, then at all rates below $E_0'(1) - \frac{\ln(\alpha(\mathcal{C}))}{n} - \frac{T}{4}$, the error exponents of the upper bounds in (20) and (44) are both maximized at $\rho = 1$, and they therefore coincide. A similar observation is provided in [17, p. 82] for the ensemble of fully random block codes. Specifically, it is observed in [17] that up to some rate threshold, the upper bounds under the suboptimal LR decoding rule for the ensemble of fully-random block codes coincide exponentially with those provided by Forney in [13].

Example 11 (Error exponents of fully random binary linear block codes). Fully random binary and linear (n,k) block codes are considered where, as mentioned in Example 3, $\alpha(\mathcal{C}) = 1$ (see (23)). For the particular case of transmission over a BSC, the error exponents for the considered ensemble are studied in [2] and [3]. The lower bounds on the block error exponents and the undetected error exponents from [2] and [3] are compared in Fig. 10(a), and 10(b), respectively, to the bounds provided in Corollary 5. The bounds are derived for a BSC with a crossover probability of $p = 0.07$ and a decoding parameter $\tau = 0.03$ (see (8) where these are the same parameters studied in [2, Fig. 1]). The error exponent provided by Gallager for the case of ML decoding is also provided for comparison, in addition to the undetected error exponent under the optimal generalized decoding rule. Apart from low rates, where the bounds in [2] and [3] outperform those provided in Corollary 5, the latter bounds on the error exponents lie in between the two previously reported bounds from [2] and [3] (see Fig. 10). Moreover, in the rate region beyond the critical rate, where the bound in [2] outperform the bound in [3], the derived bounds perform in close proximity to the tightest known bound. The superiority of the undetected error exponent under the optimal decoding rule is clearly pronounced. This comparison is further studied in Fig. 11 where the lower bounds on the undetected error exponents under the optimal and suboptimal generalized-decoding rules are provided for the same parameters as in Example 3 ($T = 0$, 0.025, 0.05, 0.1 and 0.15), assuming that transmission takes place over a BSC with a crossover probability of $p = 0.11$, and over binary-input AWGN channel with $E_{\text{s}}/N_0 = -2.8$ dB. For the case where $T = 0$, both considered exponents, for optimal and suboptimal generalized-decoding rules, coincide with each other and with the (non-expurgated) random coding error exponent of Gallager [15]. As observed in Remark 16, it is evident that for low to moderate code rates, the bounds under optimal and suboptimal generalized decoding rules coincide. However, as the coding rates approach the channel capacity, the lower bounds on the undetected block error exponents under the suboptimal generalized-decoding, are considerably loosened in comparison to the lower bound under the optimal generalized decoding.

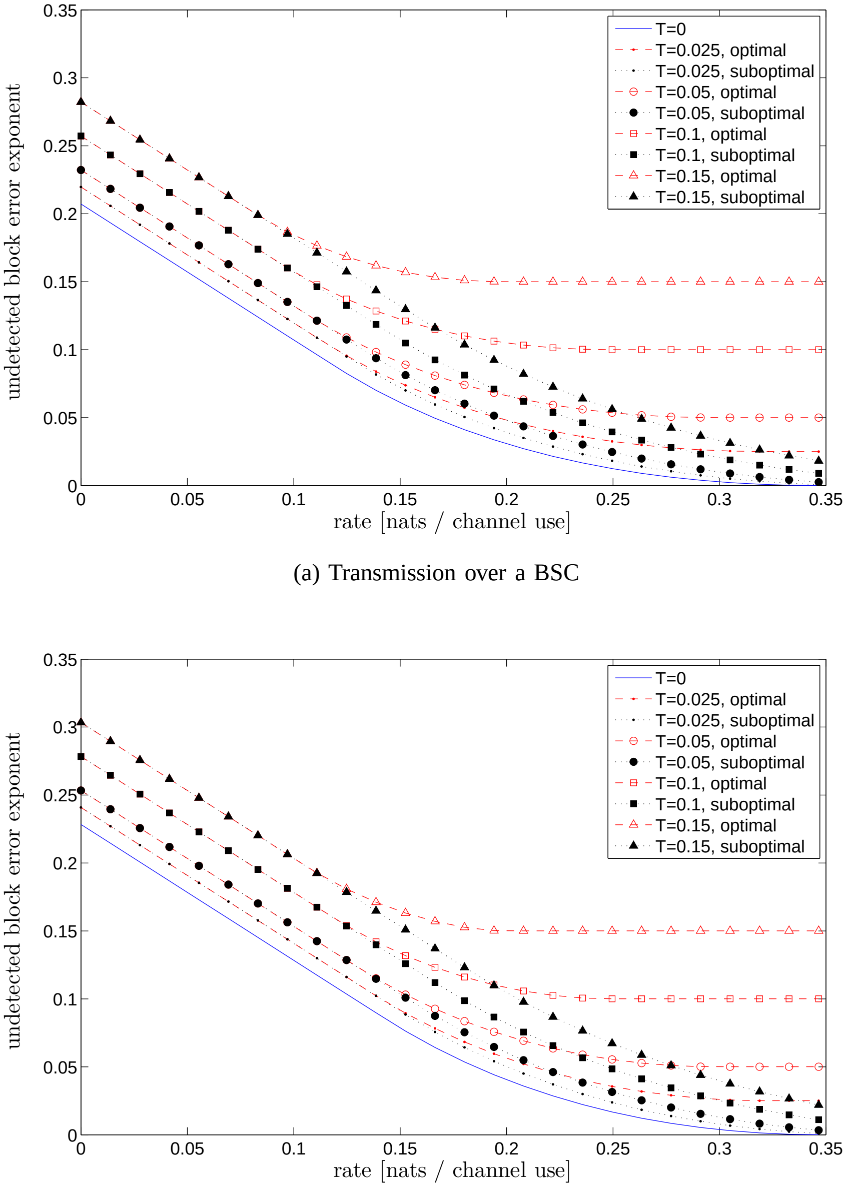


(a) Undetected error exponents

(b) Error exponents

Fig. 10: Lower bounds on the block error exponents of fully-random binary linear block codes whose transmission takes place over a BSC with a crossover probability of $p = 0.07$, under the suboptimal decoding rule in (8) with $\tau = 0.03$. The lower bounds on the undetected block error exponents in [2, Theorem 2], [3] (see also [2, Theorem 1]), and Corollary 5 (see (44)) are provided in plot (a), together with Gallager's random-coding error exponent under ML decoding [15], and the lower bound on the undetected error exponent in Corollary 2 (see (20)) under the optimal generalized decoding rule. The lower bounds on the error exponents in [2, Theorem 2], [3], and Corollary 5 (see (43)) are provided in plot (b) (the lower bound of Gallager for the random-coding error exponent under ML decoding is also provided for comparison).

Corollary 6. Under the assumptions and notation in Corollary 3, the block error probability P_{e} and the undetected



(a) Transmission over a BSC

(b) Transmission over a binary-input AWGN Channel

Fig. 11: Lower bounds on the undetected error exponents of fully-random binary linear block codes under the suboptimal generalized decoding rule in (7). The bounds based on Corollary 5, are provided in plots (a) and (b), assuming that the transmission takes place over a BSC with a crossover probability of $p = 0.11$, and a binary-input AWGN channel with $E_{\text{s}}/N_0 = -2.8$ dB, respectively. The lower bounds on the error exponents under the optimum generalized decoding rule in (5), studied in Example 3, are also provided for comparison.

error probability P_{ue} under the suboptimal decoding rule in (7), satisfy

$$P_{\text{e}} \leq e^{\frac{n\rho T}{1+\rho}} \cdot D_{\text{s}}(\rho, \mathcal{C}), \quad 0 \leq \rho \leq 1 \tag{45}$$

$$P_{\text{ue}} \leq e^{-\frac{n\rho T}{1+\rho}} \cdot D_{\text{s}}(\rho, \mathcal{C}), \quad 0 \leq \rho \leq 1 \tag{46}$$

where $D_{\text{s}}(\rho, \mathcal{C})$ is defined in (29).

Proof: Setting $s = \frac{\rho}{1+\rho}$, $G_n^m(\mathbf{y}) = \prod_{i=1}^n g(y_i)$ where g is as defined in (24), the proof follows from Proposition 7 in the same way as the proof in [18, Theorem 3]. ■

Consider the particular case of binary linear block codes whose transmission takes place over the binary-input

AWGN channel with BPSK modulation. The bound of Divsalar (see [8] and [24, Sec. 3.2.4]) provides a closed-form expression for an upper bound on the block error probability under ML decoding. The following proposition provides a similar bound under the LR decoding rule in Definition 3:

Proposition 8. Consider the transmission of a binary linear block code over the AWGN channel with BPSK modulation, then the error and undetected error probabilities under the LR decoding in (7) satisfy

$$P_{\text{e}} \leq \sum_{d=d_{\min}}^{n} \min\left\{ \exp\left(-nE_{\text{e}}\left(\frac{d}{n}, \frac{E_{\text{s}}}{N_0}\right)\right), |\mathcal{C}_d|\, Q\left(\sqrt{\frac{2E_{\text{s}}d}{N_0}} - \frac{nT}{2\sqrt{\frac{2dE_{\text{s}}}{N_0}}}\right)\right\} \tag{47}$$

$$P_{\text{ue}} \leq \sum_{d=d_{\min}}^{n} \min\left\{ \exp\left(-nE_{\text{ue}}\left(\frac{d}{n}, \frac{E_{\text{s}}}{N_0}\right)\right), |\mathcal{C}_d|\, Q\left(\sqrt{\frac{2E_{\text{s}}d}{N_0}} + \frac{nT}{2\sqrt{\frac{2dE_{\text{s}}}{N_0}}}\right)\right\} \tag{48}$$

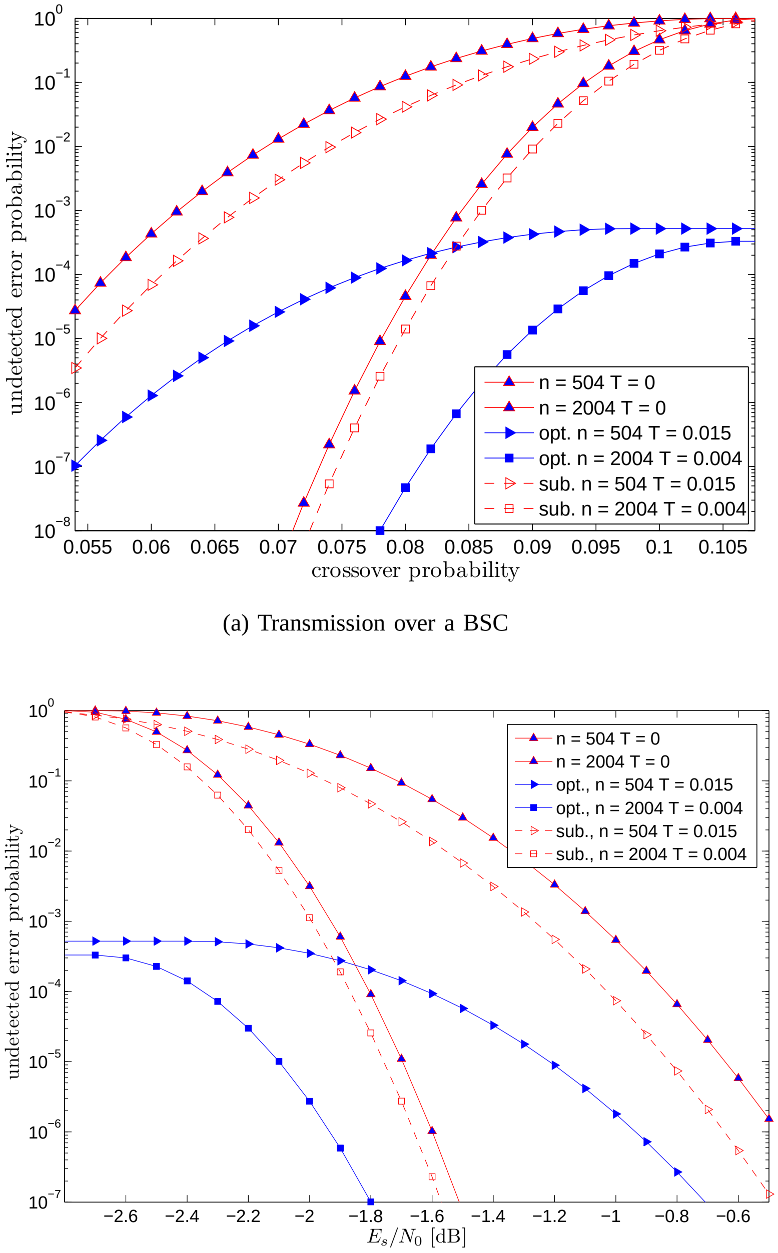
where $d_{\min}$ is the minimum Hamming distance of the code, n is the block length of the code, $|\mathcal{C}_i|$ is the number of codewords whose Hamming weight equals i, T is the decoding parameter in (7), E_{s} is the energy per transmitted (coded) symbol, $\frac{N_0}{2}$ is the two-sided power spectral density of the white Gaussian noise, and

$$\begin{aligned}
E_{\text{e}}\left(\delta, \frac{E_{\text{s}}}{N_0}\right) &\triangleq E_{\text{D}}\left(\delta, \frac{E_{\text{s}}}{N_0}\right) - \frac{T\xi}{2}, \\
E_{\text{ue}}\left(\delta, \frac{E_{\text{s}}}{N_0}\right) &\triangleq E_{\text{D}}\left(\delta, \frac{E_{\text{s}}}{N_0}\right) + \frac{T\xi}{2} \\
E_{\text{D}}\left(\delta, \frac{E_{\text{s}}}{N_0}\right) &\triangleq -r_n(\delta) + \frac{1}{2}\ln\left(\beta + (1-\beta)e^{2r_n(\delta)}\right) + \frac{\beta\delta}{1-(1-\beta)\delta}\frac{E_{\text{s}}}{N_0} \\
\beta &\triangleq \sqrt{\frac{E_{\text{s}}}{N_0}\frac{2(1-\delta)}{\delta(1-e^{-2r_n(\delta)})} + \left(\frac{1-\delta}{\delta}\right)^2\left(\left(1+\frac{E_{\text{s}}}{N_0}\right)^2 - 1\right)} - \frac{1-\delta}{\delta}\left(1+\frac{E_{\text{s}}}{N_0}\right)
\end{aligned}$$

$$\begin{aligned}
r_n(\delta) &\triangleq \frac{\ln|\mathcal{C}_d|}{n}, \quad \delta \triangleq \frac{d}{n} \\
\xi &\triangleq \frac{\beta}{\beta + (1-\beta)(1-\delta)}.
\end{aligned}$$

Proof: See Appendix H. ■

Example 12 (Error performance of expurgated binary and regular LDPC code ensembles under suboptimal generalized decoding with erasures). Consider an expurgation of the binary and regular LDPC code ensembles in Example 4 (with block lengths of 504 and 2004 bits). The upper bound in (46), on the undetected error probability under the generalized decoding rule with erasures in (7), is provided in Figs. 12(a) and 12(b), assuming that the transmission takes place over a BSC and a binary-input AWGN channel, respectively. The upper bounds under the optimal generalized decoding rule are also provided for a comparison, in addition to the upper bound under the generalized decoding rule with $T = 0$ (which coincides with the upper bound on the error probability under ML decoding). It is evident that the resulting bounds under the suboptimal generalized decoding rule are loosened in comparison to the bounds under the optimal generalized decoding rule. This result is expected from the previous example where the undetected error exponents are studied for fully-random linear block codes. In Fig 13, the upper bounds on the undetected error probability in Corollary 6 are compared with those provided in Proposition 8. The provided bounds are for the binary regular and expurgated LDPC code ensembles in Example 4 (with block lengths of 504 and 2004 bits), and for a similar ensemble with a block length of 10008 bits and $D_n = 800$. The parameter T in (7) is chosen, for this comparison, to be 0.0198, 0.0050, and $9.992 \cdot 10^{-4}$, respective to the considered block lengths. It is evident that the simple bound in (48) is loosened in comparison to the bound in (46), but only by a relatively small difference.



(a) Transmission over a BSC

(b) Transmission over a binary-input AWGN channel

Fig. 12: Upper bounds on the undetected error probabilities of some expurgated ensembles of binary and regular (6,12) LDPC codes under the optimal and sub-optimal generalized decoding rules in (5) and (7), respectively. The upper bound in Corollary 6 is shown in plots (a) and (b), assuming that the transmission takes place over a BSC and a binary-input AWGN channel, respectively. The upper bounds in Corollary 3, studied in Examples 4 and 5, are also provided for comparison.

VI. Upper bounds under fixed-size list decoding

In this section, upper bounds on the block error probability are derived for the fixed-size list decoding (see Definition 4). As mentioned in Section II, the block error event in this case corresponds to the possibility that the decoded list does not include the transmitted codeword.

Proposition 9. Consider the transmission of a block code $\mathcal{C}$ with M codewords of length n, and let $p(\mathbf{y}|\mathbf{x})$ designate the transition probability of the channel where $\mathbf{x} \in \mathcal{C}$ is the transmitted codeword and $\mathbf{y} \in \mathcal{Y}^n$ is the received vector. Consider the case where a fixed-size list decoder is used where the size of the list is denoted by L. Then, the

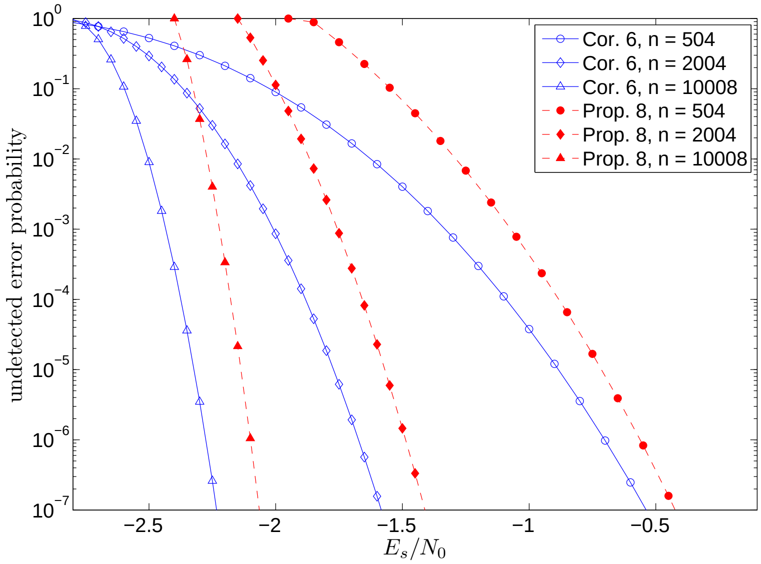


Fig. 13: A comparison between the upper bounds in (46) and (48), on the undetected error probability under the LR generalized decoding rule in (7). The comparison is provided for binary expurgated and regular (6,12) LDPC code ensembles of Gallager with block lengths of 504, 2004 and 10008 bits whose transmissions take place over binary-input AWGN channels with BPSK modulation.

conditional block error probability $P_{\mathrm{e}|m}$, given that the m-th message is transmitted satisfies

$$P_{\mathrm{e}|m} \le \left(\sum_{\mathbf{y}} G_n^m(\mathbf{y})\, p(\mathbf{y}|\mathbf{x}_m)\right)^{1-\rho} \left(\frac{1}{L}\sum_{m'\neq m}\sum_{\mathbf{y}} p(\mathbf{y}|\mathbf{x}_m)\, G_N^m(\mathbf{y})^{1-\frac{1}{\rho}} \left(\frac{p(\mathbf{y}|\mathbf{x}_{m'})}{p(\mathbf{y}|\mathbf{x}_m)}\right)^{\frac{s}{\rho}}\right)^{\rho}. \qquad (49)$$

where $0 \le s \le \rho \le 1$ are real-valued parameters, and G_n^m is an arbitrary non-negative function over $\mathcal{Y}^n$ which possibly depends on the codeword $\mathbf{x}_m$, for $1 \le m \le M$.

Proof: See Appendix I. ■

The following corollary is a particularization of Proposition 9 for the ensemble of fully-random block codes, with fixed block length and rate, whose transmission takes place over a memoryless channel:

Corollary 7. Consider the transmission of a block code $\mathcal{C}$ over a memoryless communication channel. Then, under the notation in Proposition 9, there exists a block code whose block error probability P_{e} under fixed-size list decoding satisfies

$$P_{\mathrm{e}} \le e^{-nE_{\mathrm{r}}\left(R-\frac{1}{n}\ln L\right)}$$

where $R \triangleq \frac{\ln M}{n}$ is the code rate (in nats per channel use),

$$E_{\mathrm{r}}(R) \triangleq \max_{0\le\rho\le1,\ q_X} \Big(E_0(\rho, q_X) - \rho R\Big) \qquad (50)$$

$$E_0(\rho, q_X) \triangleq -\ln\left(\sum_{y\in\mathcal{Y}}\left(\sum_{x\in\mathcal{X}} q_X(x)\, p(y|x)^{\frac{1}{1+\rho}}\right)^{1+\rho}\right)$$

and q_X is a probability distribution over the input alphabet $\mathcal{X}$.

Proof: Fix a probability distribution q_X over $\mathcal{X}$, and consider the ensemble of random block codes where each codeword is chosen independently according to $q_{\mathbf{X}}(\mathbf{x}) = \prod_{i=1}^{n} q_X(x_i)$. First, we apply the bound in (49) for a specific realization of a codebook, with $s = \frac{\rho}{1+\rho}$ and

$$G_n^m(\mathbf{y}) \triangleq \left(\sum_{\mathbf{x}} q_{\mathbf{X}}(\mathbf{x}) \left(\frac{p(\mathbf{y}|\mathbf{x})}{p(\mathbf{y}|\mathbf{x}_m)} \right)^{\frac{s}{\rho}} \right)^{\rho}.$$

The proof follows by a random coding argument, and by choosing the optimal probability distribution q_X. ■

Remark 17. The upper bound in Corollary 7 coincides exponentially with the sphere-packing lower bound in [26]. Because of its mathematical resemblance, it may be considered as a generalization of the well-known random-coding error-exponent of Gallager [15], for the case at hand.

The following bound is provided for the case of binary linear block codes whose transmission takes place over an MBIOS channel:

Corollary 8. Consider an (n, k) binary linear block code $\mathcal{C}$ whose transmission takes place over an MBIOS channel. Then, the block error probability P_{e} under fixed-size list-decoding, satisfies

$$P_{\text{e}} \le e^{-nE_{\text{r}}\left(R+\frac{1}{n}\ln\left(\frac{\alpha(\mathcal{C}}{L}\right)\right)} \tag{51}$$

where

$$E_{\text{r}}(R) \triangleq \max_{0\le\rho\le 1} \left(E_0(\rho) - \rho R\right)$$

and R is the code rate (in nats per channel use), L is the list size, and $E_0(\rho)$ and $\alpha(\mathcal{C})$ are defined in (22) and (23), respectively.

Proof: According to Proposition 4, it is necessary to analyze only the conditional error event assuming that the all-zero codeword is transmitted. Setting $G_n^0(\mathbf{y}) = \prod_{i=1}^n g(y_i)$ in (49), it follows that

$$\begin{aligned} P_{\text{e}} \le & \left(\sum_{y\in\mathcal{Y}} g(y) p(y|0) \right)^{n(1-\rho)} \\ & \left(\frac{1}{L} \sum_{i=1}^{n} |\mathcal{C}_i| \left(\sum_{y\in\mathcal{Y}} g(y)^{1-\frac{1}{\rho}} p(y|0) \right)^{n-i} \left(\sum_{y\in\mathcal{Y}} g(y)^{1-\frac{1}{\rho}} p(y|1)^{\lambda} p(y|0)^{1-\lambda} \right)^{i} \right)^{\rho} \end{aligned} \tag{52}$$

where $|\mathcal{C}_i|$ denotes the number of codewords whose Hamming distance is i, $1 \le i \le n$. The proof follows from (52) by setting $\lambda = \frac{1}{1+\rho}$ where g is as defined in (24) (see similar derivation in [24, Section 4.4.1]). ■

Remark 18. For the particular case of fully-random linear block codes, the bound in (51) coincides with the bound in Corollary 7 for fully-random block codes.

Remark 19. The bound in Corollary 8 resembles to the SFB [27], and therefore may be considered as a generalization of the SFB for the case at hand.

Remark 20. The bound in (52) ban be generalized to non-binary linear block codes using a similar derivation as in [18]. Note, however, that in [18], non-binary codes are studied under ML decoding and not list-decoding. Nevertheless, the similarity of the bound in (49) to the upper bounds derived in [18] allows to use the same arguments for the case at hand (see Appendix I).

Corollary 9. Under the assumptions and notation in Corollary 3, the block error probability probability P_{e} under fixed-size list-decoding where L denotes the size of the list, satisfies

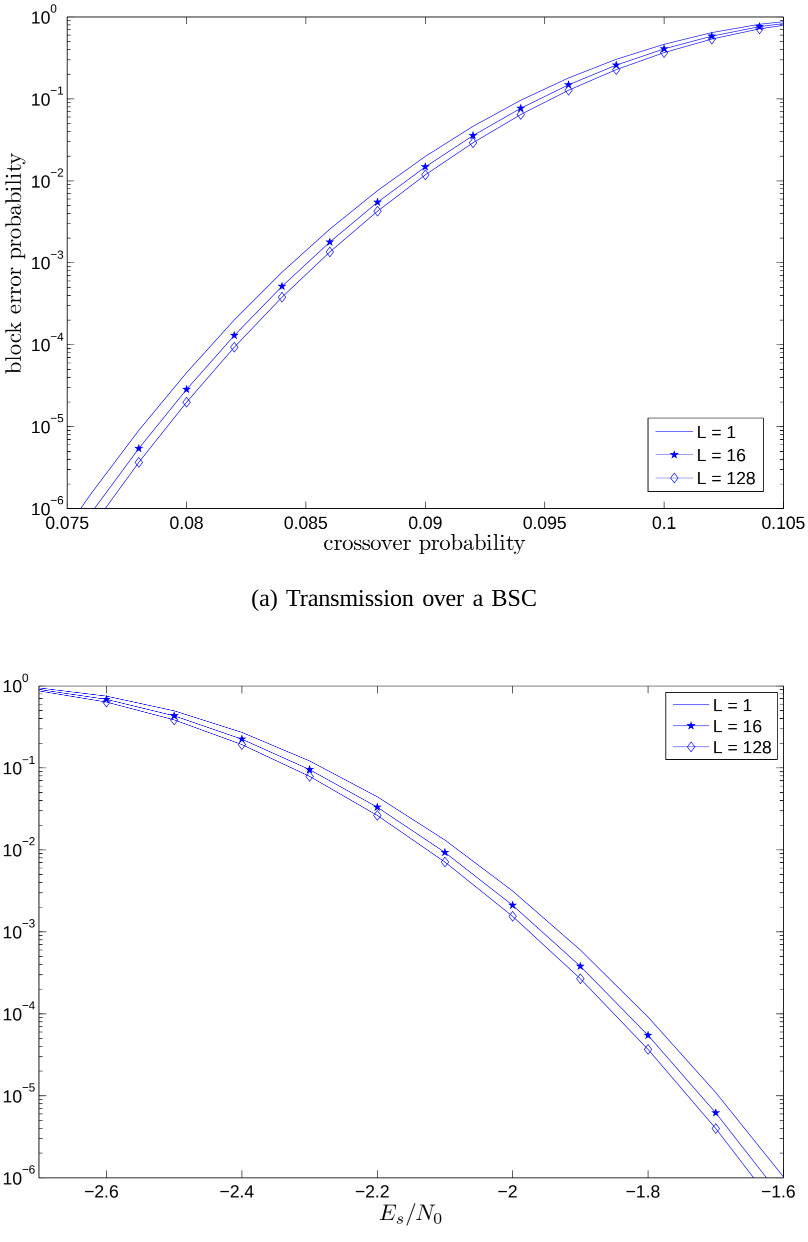
$$P_{\text{e}} \le A(\rho)^{n(1-\rho)} \left(\frac{1}{L} \sum_{1\le l\le n} P(l) \binom{n}{l} B(\rho)^{n-l} C(\rho)^{l} \right)^{\rho} \tag{53}$$

where $A(\rho)$, $B(\rho)$, and $C(\rho)$ are defined in (30)–(32).

Proof: Setting $s = \frac{\rho}{1+\rho}$ and $G_n^m(\mathbf{y}) = \prod_{i=1}^n g(y_i)$ where g is defined in (24), the proof follows from Proposition 9 in the same way as the proof in [18, Theorem 3]. ■

Remark 21. In the derivation of the bound in (51), a sum is upper bounded by a product of the maximal summand with the number of summands. This operation is avoided in the derivation of the bound in (53). Hence, the bound in Corollary 9 is tighter than the one in Corollary 8.

Remark 22. For the particular case of binary linear block codes, the symmetry condition in (26) is not mandatory and the bound in Corollary 9 follows by replacing the term $P(l)\binom{n}{l}$ with the distance spectrum of the considered code (ensemble).

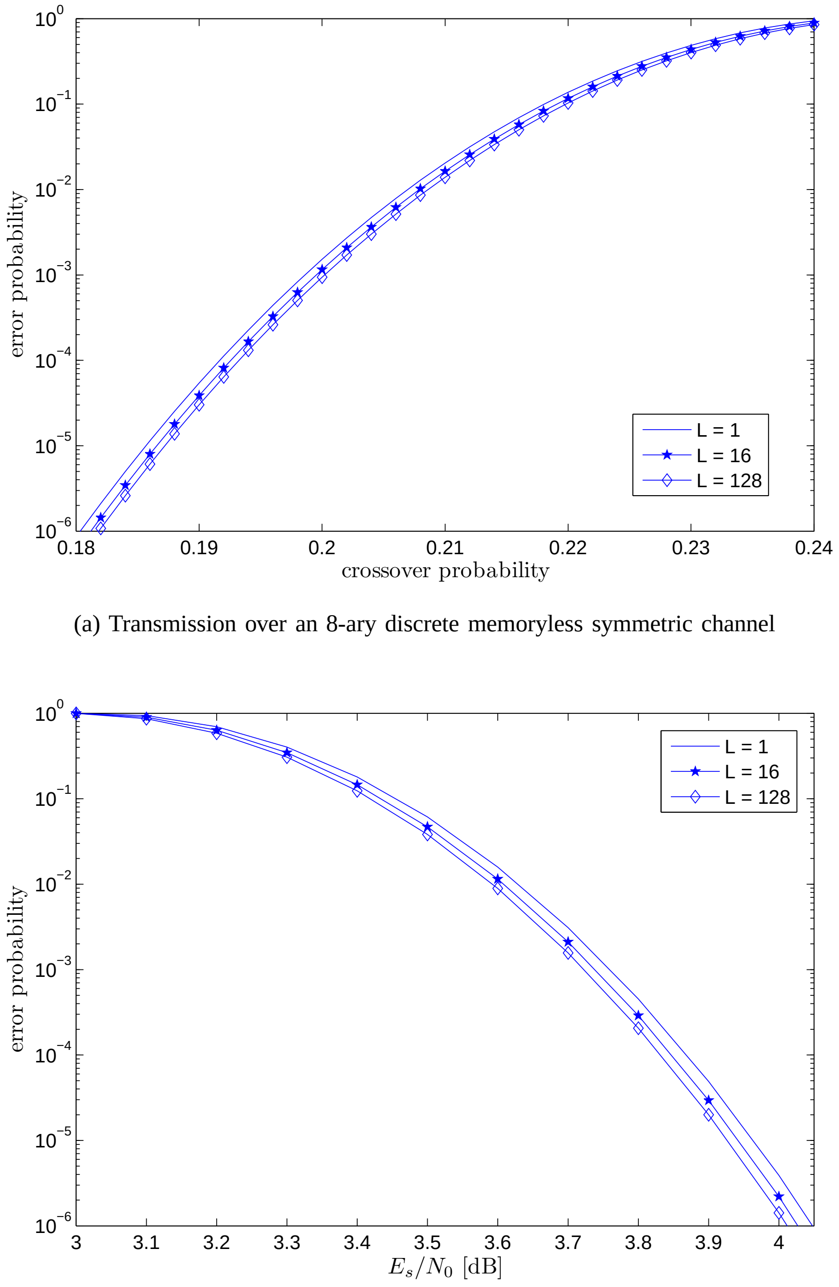


(a) Transmission over a BSC

(b) Transmission over a binary-input AWGN channel

Fig. 14: Upper bounds on the error probability for an expurgation of Gallager's ensemble of binary and regular (6,12) LDPC codes with a block length of 2004 bits (see Example 4). A list decoder is assumed where the size of the list is set to L. The upper bound in Corollary 9 is provided for some values of L. The bounds are shown in plots (a) and (b), respectively, for the case where the transmission takes place over a BSC and a binary-input AWGN channel.

Example 13 (Error performance of an expurgated ensemble of binary and regular LDPC codes under fixed-size list decoding). Consider the expurgation of Gallager's ensemble of binary and regular (6,12) LDPC codes with a block length of 2004 bits (see Example 4). Upper bounds on the block error probability under fixed-size list-decoding are shown in Figs. 14(a) and 14(b), assuming that the transmission takes place over a BSC and a binary-input AWGN channel, respectively. The upper bound in Corollary 9 is evaluated for list sizes of $L = 1$, 16, and 128 codewords. Note that the upper bound for $L = 1$ corresponds to ML decoding. The bounds on the error probability show some marginal improvement by increasing the considered list size from $L = 1$ to 128.



(a) Transmission over an 8-ary discrete memoryless symmetric channel

(b) Transmission over an AWGN channel with 8-PSK modulation

Fig. 15: Upper bounds on the error probability for an expurgation of Gallager's ensemble of regular (8,16) LDPC codes with octal alphabet and a block length of 1008 symbols (see Example 6). A list decoder is considered where the size of the list is set to L. The upper bound in Corollary 9 is provided in plots (a) and (b) for several values of L, assuming that the transmission takes place over an 8-ary discrete memoryless symmetric channel and an AWGN channel with 8-PSK modulation, respectively.

Example 14 (Error performance of an expurgated ensemble of non-binary and regular LDPC codes under fixed-size list decoding). Consider the expurgation of Gallager's ensemble of regular (8,16) LDPC codes with octal alphabet and a block length of 1008 symbols (see Example 6). Upper bounds on the block error probability under fixed-size list decoding are shown in Figs. 15(a) and 15(b), assuming that the transmission takes place over an 8-ary discrete memoryless symmetric channel and an AWGN channel with 8-PSK modulation, respectively. The bound in Corollary 9 is evaluated for list sizes of $L = 1$, 16, and 128 codewords. similarly to the case of binary code ensembles, only marginal improvement in the error performance is observed by increasing the value of L from 1 to 128.

VII. Summary and Conclusions

This paper considers the performance of several generalized decoding rules over memoryless symmetric channels. Three types of generalized decoding rules are considered:

1) The optimal generalized decoding rule in [13] with erasures and variable list sizes.
2) The suboptimal likelihood-ratio (LR) decoding rule with erasures (see [2] and [13]).
3) A fixed-size list decoding rule (see [12] and [34]) where the decoder outputs a list with includes the L most probable codewords (where the value of L is set a-priori).

The independence of the error performance on the transmitted codeword is proved in Propositions 2-4 for the considered decoding rules. Specifically, it is shown that the undetected error probability, block error probability (of both undetected errors and erasures), list decoding error probability, and the expected size of the decoded list are all independent of the transmitted codeword when the transmission takes place over a memoryless symmetric channel.

Upper bounds on the decoding error probability are provided. Moreover, upper bounds on the expected size of the decoded list are derived. The derivation of these bounds is based on a generalization of a bounding technique of Duman and Salehi (see, e.g., [8], [10], [11], [24]). The provided bounds are suitable for the analysis of structured and random codes (or code ensembles) over memoryless symmetric channels. Both binary and non-binary code ensembles are studied in this paper under generalized decoding rules. When binary codes are considered, the bounds are based on the distance spectra of the codes, and when non-binary ensembles are studied, the complete composition spectra are required under the symmetry assumption in (26). For the case of LR decoding of binary linear block codes, a derivation of a closed-form expression is provided via a similar derivation to [8] which applies to ML decoding.

Several applications and particularizations of the provided bounds are studied. First, the random coding error exponents in [13] are reproduced, in addition to some error exponents under the suboptimal LR decoding rule with erasures. These error exponents are derived by applying the new bounds to fully random block codes. Next, a derivation of the error exponents of fully random linear block codes under optimal and suboptimal (LR) generalized decoding is provided. The resulting error exponents under the suboptimal LR decoding rule are compared with a recent improvement in [2], where the ensemble of binary fully random linear block codes over binary symmetric channels (BSC) is studied. This comparison shows good proximity of the provided error exponents with the results in [2]. In addition, it is shown that the error exponents for the fully random linear block codes under the suboptimal LR decoding rule, coincide for low rates with the corresponding error exponents under the optimal decoding rule. This observation is similar to an observation in [17], where the ensemble of fully random block codes is considered. A Lower bound on the error exponent under fixed-size list-decoding is also derived as an application. The resulting bound coincides exponentially with the corresponding sphere-packing lower bound in [26].

Applications of the bounds for the performance analysis of structured code ensembles are further exemplified for some expurgated ensembles of (binary and non-binary) regular low-density parity-check (LDPC) codes. The error performance under some generalized decoding rules for these LDPC code ensembles is studied assuming that the transmission takes place over memoryless symmetric channels. Specifically, undetected error, and total block error (including undetected errors and erasures) probabilities are evaluated under optimal and suboptimal (LR) generalized decoding rules. In addition, the error performance for list decoding applications is studied for these ensembles. Both fixed-size and variable-size list decoding are considered, and an upper bound on the expected list-size is evaluated for the latter case.

An analysis of various hybrid automatic-repeat request (ARQ) schemes is also provided in this work. A noiseless and immediate one-bit feedback channel is assumed, where erasure-outputs at the decoder triggers, via this feedback channel, the retransmissions of messages. Hybrid-ARQ schemes with and without deadlines are considered, in addition to schemes with incremental redundancy. Upper bounds on the error probability and lower bounds on the expected overall rate are provided and exemplified in this paper.

APPENDIX A
PROOF OF PROPOSITION 2

The following proof holds for memoryless symmetric channels with discrete-output alphabets, and the generalization to continuous-output alphabets is direct. We state first the following technical lemma:

Lemma 1. let x_1, x_2, x_3 be arbitrary symbols in $\mathcal{X}$, and let p be a transition probability law of a memoryless symmetric channel. Then,

$$p\Big(\mathcal{T}\big(\mathcal{T}(y,x_1),x_2\big)|x_3\Big) = p\big(\mathcal{T}(y,x_1+x_2)|x_3\big)$$

where $\mathcal{T}$ is a mapping which satisfies the properties in Definition 1.

Proof: the reader is referred to [18, Appendix A]. ■

Assuming that all the codewords are sent with equal probability, the decision regions in (5) satisfy

$$\begin{aligned}\Lambda_m &\overset{(a)}{=} \left\{\mathbf{y}: \ \frac{p(\mathbf{y}|\mathbf{x}_m)}{\sum_{m'\neq m} p(\mathbf{y}|\mathbf{x}_{m'})} \geq e^{nT}\right\} \\ &\overset{(b)}{=} \left\{\mathbf{y}: \ \frac{\prod_{i=1}^n p(y_i|x_{m,i})}{\sum_{m'\neq m}\prod_{i=1}^n p(y_i|x_{m',i})} \geq e^{nT}\right\} \\ &\overset{(c)}{=} \left\{\mathbf{y}: \ \frac{\prod_{i=1}^n p(\mathcal{T}(y_i,-x_{m,i})|0)}{\sum_{m'\neq m}\prod_{i=1}^n p(\mathcal{T}(y_i,-x_{m',i})|0)} \geq e^{nT}\right\}\end{aligned} \tag{54}$$

where (a) follows from (5) and the equal a-priori message probability assumption, (b) holds since the channel is memoryless, and (c) follows from the symmetry of the channel (see (1)). Let $\mathbf{z} = (z_1,\ldots,z_n)$ be defined as

$$z_i \triangleq \mathcal{T}(y_i,-x_{m,i}), \quad 1 \leq i \leq n \tag{55}$$

where m is the index of the transmitted codeword. From Lemma 1, it follows that $\mathbf{y} \in \Lambda_m$ if and only if $\mathbf{z} \in \tilde{\Lambda}_m$ where

$$\tilde{\Lambda}_m \triangleq \left\{\mathbf{z}\in\mathcal{Y}^n: \ \frac{\prod_{i=1}^n p(z_i|0)}{\sum_{m'\neq m}\prod_{i=1}^n p(\mathcal{T}(z_i,x_{m,i}-x_{m',i})|0)} \geq e^{nT}\right\}, \quad 1 \leq m \leq q^k.$$

Using the linearity of the code, it follows that

$$\tilde{\Lambda}_m = \left\{\mathbf{z}\in\mathcal{Y}^n: \ \frac{\prod_{i=1}^n p(z_i|0)}{\sum_{l\neq 0}\prod_{i=1}^n p(\mathcal{T}(z_i,x_{l,i})|0)} \geq e^{nT}\right\}.$$

Since the set $\tilde{\Lambda}_m$ is independent of the index m, then

$$\tilde{\Lambda}_m = \tilde{\Lambda}_1 \text{ for all } 1 \leq m \leq q^k. \tag{56}$$

As a result, the conditional block error probability of the m-th message in (2) satisfies

$$\begin{aligned}P_{\mathrm{e}|m} &= \sum_{\mathbf{z}\in\tilde{\Lambda}_m^{\mathrm{c}}} p(\mathbf{z}|\mathbf{0}) \\ &\overset{(a)}{=} \sum_{\mathbf{z}\in\tilde{\Lambda}_1^{\mathrm{c}}} p(\mathbf{z}|\mathbf{0})\end{aligned}$$

where (a) follows from (56). This concludes the proof of the message independence property for the block error event.

We continue in proving the message independence property for the undetected error event (or the expected number of incorrect codewords when list decoding is considered). Assuming a memoryless symmetric channel, it follows from (1) and (4) that

$$\begin{aligned} P_{\text{ue}|m} &= \sum_{m' \neq m} \sum_{\mathbf{y} \in \Lambda_{m'}} p(\mathbf{y}|\mathbf{x}_m) \\ &= \sum_{m' \neq m} \sum_{\mathbf{y} \in \Lambda_{m'}} \prod_{i=1}^{n} p\big(\mathcal{T}(y_i, -x_{m,i})|0\big) \end{aligned} \tag{57}$$

where from (54)

$$\Lambda_{m'} = \left\{ \mathbf{y} : \; \frac{\prod_{i=1}^{n} p(\mathcal{T}(y_i, -x_{m',i})|0)}{\sum_{m'' \neq m'} \prod_{i=1}^{n} p(\mathcal{T}(y_i, -x_{m'',i})|0)} \geq e^{nT} \right\}.$$

Let $\mathbf{z}$ be a vector defined as in (55), then from Lemma 1

$$p\big(\mathcal{T}(y_i, -x_{m',i})|0\big) = p\big(\mathcal{T}(z_i, x_{m,i} - x_{m',i}|0\big), \quad i = 1, \ldots, n.$$

Hence, given that $\mathbf{x}_m$ is the transmitted codeword, then $\mathbf{y} \in \Lambda_{m'}$ for some $m' \neq m$ if and only if $\mathbf{z} \in \Gamma_{m,m'}$ where

$$\Gamma_{m,m'} \triangleq \left\{ \mathbf{z} \in \mathcal{Y}^n : \; \frac{\prod_{i=1}^{n} p(\mathcal{T}(z_i, x_{m,i} - x_{m',i})|0)}{\sum_{m'' \neq m'} \prod_{i=1}^{n} p(\mathcal{T}(z_i, x_{m,i} - x_{m'',i})|0)} \geq e^{nT} \right\}. \tag{58}$$

From (55), the conditional undetected error probability in (57) is rewritten in the form

$$P_{\text{ue}|m} = \sum_{m' \neq m} \sum_{\mathbf{z} \in \Gamma_{m,m'}} p(\mathbf{z}|\mathbf{0}). \tag{59}$$

Using the linearity of the code, then $x_{m,i} - x_{m'',i} = \big(x_{m,i} - x_{m',i}\big) + \big(x_{m',i} - x_{m'',i}\big) = x_{l_1,i} + x_{l_2,i}$ for some indices l_1 and l_2 which correspond to non-zero codewords. Let $\mathbf{x} \triangleq \mathbf{x}_{l_1}$ and $\tilde{\mathbf{x}} = \mathbf{x}_{l_2}$, then the conditional undetected error probability in (59) is expressed equivalently in the form

$$P_{\text{ue}|m} = \sum_{\substack{\mathbf{x} \in \mathcal{C} \\ \mathbf{x} \neq \mathbf{0}}} \sum_{\mathbf{z} \in \Gamma(\mathbf{x})} p(\mathbf{z}|\mathbf{0})$$

where, based on (58),

$$\Gamma(\mathbf{x}) \triangleq \left\{ \mathbf{z} \in \mathcal{Y}^n : \; \frac{\prod_{i=1}^{n} p(\mathcal{T}(z_i, x_i)|0)}{\sum_{\substack{\tilde{\mathbf{x}} \in \mathcal{C} \\ \tilde{\mathbf{x}} \neq \mathbf{0}}} \prod_{i=1}^{n} p(\mathcal{T}(z_i, x_i + \tilde{x}_i)|0)} \geq e^{nT} \right\}.$$

This proves the independence property for the undetected error event, and it concludes the proof of Proposition 2.

Appendix B
Proof of Proposition 3

Similarly to Appendix A, also the following proof considers memoryless symmetric channels with discrete-output alphabets, where the generalization to continuous output alphabets is direct. Let p be the transition probability function of the considered channel, $\mathcal{C}$ be an (n, k) linear block code over an alphabet whose cardinality is q, and $\mathcal{T}$ be a mapping as specified in Definition 1. It is assumed that all the codewords of $\mathcal{C}$ are sent with equal probability. For an arbitrary set $\Lambda \subseteq \mathcal{Y}^n$ and a codeword $\mathbf{x}_m \in \mathcal{C}$, let

$$\mathcal{Z}_m(\Lambda) \triangleq \{\mathbf{z} \in \mathcal{Y}^n : \; \mathcal{T}(z_i, x_{m,i}) \in \Lambda\}. \tag{60}$$

In addition, we use the notation $\Lambda^{\text{LR}}(\mathbf{x}_m)$ for the decision region Λ_m^{LR} in (7) of the codeword $\mathbf{x}_m$. Note that for the concerned decoding rule with $T > 0$, the decision regions are disjoint. The following technical lemma is introduced:

Lemma 2. Let $\mathcal{Z}_m$ be the mapping defined in (60), and Λ_m^{LR} be the decision region in (7). Then,

$$\mathcal{Z}_m\left(\Lambda_{m'}^{\text{LR}}\right) = \Lambda^{\text{LR}}(\mathbf{x}_{m'} - \mathbf{x}_m), \quad \forall \, m, m' \in \{1, \ldots, q^k\}. \tag{61}$$

Proof: Let us choose $\mathbf{z} \in \mathcal{Z}_m\left(\Lambda_{m'}^{\text{LR}}\right)$, and let $\mathbf{y} = (y_1, \ldots, y_n)$ be defined via the equality

$$y_i = \mathcal{T}(z_i, x_{m,i}), \quad i = 1, \ldots, n. \tag{62}$$

From (7) and (60)

$$\frac{p(\mathbf{y}|\mathbf{x}_{m'})}{p(\mathbf{y}|\mathbf{x}_{m_2'})} \geq e^{nT}$$

where $\mathbf{x}_{m'}$ and $\mathbf{x}_{m_2'}$ are the most probable codewords, in a descending order, for $\mathbf{y}$ as a received vector. Using the symmetry of the channel, it follows from (1) that

$$p(\mathbf{y}|\mathbf{x}_{m'}) = p(\mathbf{z}|\mathbf{x}_{m'} - \mathbf{x}_m).$$

As a result, $\mathbf{x}_{m'} - \mathbf{x}_m$ is the most probable codeword if $\mathbf{z}$ is the received vector (otherwise, if there exists a codeword $\mathbf{x} \neq \mathbf{x}_{m'} - \mathbf{x}_m$ which is more probable, then there exists a more probable codeword for $\mathbf{y}$ which is different from $\mathbf{x}_{m'}$). The same argument shows that $\mathbf{x}_{m_2'} - \mathbf{x}_m$ is the second most probable codeword for $\mathbf{z}$, and

$$\frac{p(\mathbf{z}|\mathbf{x}_{m'} - \mathbf{x}_m)}{p(\mathbf{z}|\mathbf{x}_{m_2'} - \mathbf{x}_m)} \geq e^{nT}.$$

This verifies that $\mathbf{z} \in \Lambda^{\text{LR}}(\mathbf{x}_{m'} - \mathbf{x}_m)$ which shows that $\mathcal{Z}_m\left(\Lambda_{m'}^{\text{LR}}\right) \subseteq \Lambda^{\text{LR}}(\mathbf{x}_{m'} - \mathbf{x}_m)$. To show the opposite inclusion, which then yields that these two sets are equal, let $\mathbf{z} \in \Lambda^{\text{LR}}(\mathbf{x}_{m'} - \mathbf{x}_m)$. This implies that the codeword $\mathbf{x}_{m'} - \mathbf{x}_m$ is the most probable codeword if $\mathbf{z}$ is the received vector, and

$$\frac{p(\mathbf{z}|\mathbf{x}_{m'} - \mathbf{x}_m)}{p(\mathbf{z}|\mathbf{x}_{m_2''})} \geq e^{nT}$$

where $\mathbf{x}_{m_2''}$ is the second most probable codeword for $\mathbf{z}$. Again, using the symmetry of the channel, for a vector $\mathbf{y}$ as in (62), it follows that $\mathbf{x}_{m'}$ is the most probable codeword for $\mathbf{y}$, $\mathbf{x}_{m_2''} + \mathbf{x}_m$ is the second most probable codeword for $\mathbf{y}$, and

$$\frac{p(\mathbf{y}|\mathbf{x}_{m'})}{p(\mathbf{y}|\mathbf{x}_{m_2''} + \mathbf{x}_m)} \geq e^{nT}.$$

As a result, $\mathbf{z} \in \mathcal{Z}_m\left(\Lambda_{m'}^{\text{LR}}\right)$, which yields that $\Lambda^{\text{LR}}(\mathbf{x}_{m'} - \mathbf{x}_m) \subseteq \mathcal{Z}_m\left(\Lambda_{m'}^{\text{LR}}\right)$. This concludes the proof of (61). ■

From (62), the conditional block error probability satisfies

$$\begin{aligned} P_{\text{e}|m} &= \sum_{\mathbf{y} \notin \Lambda_m^{\text{LR}}} p(\mathbf{y}|\mathbf{x}_m) \\ &\overset{\text{(a)}}{=} \sum_{\mathbf{z} \notin \mathcal{Z}_m(\Lambda_m^{\text{LR}})} p(\mathbf{z}|\mathbf{0}) \\ &\overset{\text{(b)}}{=} \sum_{\mathbf{z} \notin \Lambda^{\text{LR}}(\mathbf{0})} p(\mathbf{z}|\mathbf{0}) \end{aligned}$$

where (a) follows from (1) and (62), and (b) follows from (61). This proves the message independence property for the conditional block error probability. Using the same arguments, the message independence property is established for the conditional undetected error probability:

$$\begin{aligned} P_{\text{ue}|m} &= \sum_{m' \neq m} \sum_{\mathbf{y} \in \Lambda_{m'}^{\text{LR}}} p(\mathbf{y}|\mathbf{x}_m) \\ &= \sum_{m' \neq m} \sum_{\mathbf{z} \in \mathcal{Z}_m\left(\Lambda_{m'}^{\text{LR}}\right)} p(\mathbf{z}|\mathbf{0}) \\ &= \sum_{m' \neq m} \sum_{\mathbf{z} \in \Lambda^{\text{LR}}(\mathbf{x}_{m'} - \mathbf{x}_m)} p(\mathbf{z}|\mathbf{0}) \\ &= \sum_{\substack{\mathbf{x} \in \mathcal{C} \\ \mathbf{x} \neq \mathbf{0}}} \sum_{\mathbf{z} \in \Lambda^{\text{LR}}(\mathbf{x})} p(\mathbf{z}|\mathbf{0}) \end{aligned}$$

where the second equality follows from (62) and since the mapping $\mathcal{T}$ is bijective, the third equality follows from (61), and the last equality follows from the linearity of the code.

APPENDIX C
PROOF OF PROPOSITION 4

Considering ties as error events[2], the conditional block error probability for a list of size L satisfies

$$P_{\mathsf{e}|m} = \sum_{\mathbf{y} \in \Lambda_m^L} p(\mathbf{y}|\mathbf{x}_m) \tag{63}$$

where

$$\Lambda_m^L \triangleq \left\{ \mathbf{y} \in \mathcal{Y}^n : \; \exists \{m_i\}_{i=1}^L \text{ s.t. } m_i \neq m, \; p(\mathbf{y}|\mathbf{x}_{m_i}) \geq p(\mathbf{y}|\mathbf{x}_m) \; \forall \; 1 \leq i \leq L \right\} \tag{64}$$

is the complementary of the decision region of $\mathbf{x}_m \in \mathcal{C}$ under list decoding of fixed-size L (here $\{m_i\}_{i=1}^L$ is a sequence of distinct integers), i.e., if $\mathbf{y} \in \Lambda_m^L$ then the codeword $\mathbf{x}_m$ is not included in the list for a received vector $\mathbf{y}$. Using the change of variables in (62), it follows from (63) that for linear block codes whose transmission takes place over memoryless symmetric channels

$$P_{\mathsf{e}|m} = \sum_{\mathbf{z} \in \mathcal{Z}_m(\Lambda_m^L)} p(\mathbf{z}|\mathbf{0})$$

where $\mathcal{Z}_m\left(\Lambda_m^L\right)$ is as defined in (60). The following lemma concludes the proof of Proposition 4:

Lemma 3. Let $\mathcal{Z}_m$ be a mapping defined in (60), and Λ_m^L be the decoding region of $\mathbf{x}_m \in \mathcal{C}$ under list decoding with a fixed size L. Then,

$$\mathcal{Z}_m\left(\Lambda_m^L\right) = \Lambda_1^L$$

for all $1 \leq m \leq q^k$, where Λ_1^L is the complementary of the decision region of the all-zero codeword $\mathbf{x}_1 = \mathbf{0}$ under list decoding of size L.

Proof: Let us choose $\mathbf{z} \in \mathcal{Z}\left(\Lambda_m^L\right)$. From (60), there exists $\mathbf{y} \in \Lambda_m^L$ where

$$y_i = \mathcal{T}(z_i, x_{m,i}), \quad i = 1, \ldots, n \tag{65}$$

and $\mathcal{T}$ is a specified in Definition 1. From (64), there exists a list of L distinct codewords, $\{\mathbf{x}_{m_i}\}_{i=1}^L$, for which

$$p(\mathbf{y}|\mathbf{x}_{m_i}) > p(\mathbf{y}|\mathbf{x}_m), \quad i = 1, \ldots, L. \tag{66}$$

Using the symmetry of the channel, it follows that

$$p(\mathbf{z}|\mathbf{x}_{m_i} - \mathbf{x}_m) \geq p(\mathbf{z}|\mathbf{0}). \tag{67}$$

This assures that $\mathbf{z} \in \Lambda_1^L$, which shows that $\mathcal{Z}_m\left(\Lambda_m^L\right) \subseteq \Lambda_1^L$.

Next, in order to show the opposite inclusion, let $\mathbf{z} \in \Lambda_1^L$. Then, there exists a list of L non-zero codewords $\{\mathbf{x}_{m_i}\}_{i=1}^L$, $m_i \neq 1$, satisfying

$$p(\mathbf{z}|\mathbf{x}_{m_i}) \geq p(\mathbf{z}|\mathbf{0})$$

and therefore from the symmetry of the mapping $\mathcal{T}$ and the equality in (65), we get

$$p(\mathbf{y}|\mathbf{x}_{m_i} + \mathbf{x}_m) \geq p(\mathbf{y}|\mathbf{x}_m)$$

It assures that $\mathbf{z} \in \mathcal{Z}_m\left(\Lambda_m^L\right)$ which implies that $\Lambda_1^L \subseteq \mathcal{Z}_m\left(\Lambda_m^L\right)$. This two inclusions complete the proof of the lemma. ■

[2] Such a pessimistic assumption is reasonable, see also a similar assumption in [32, p. 59].

APPENDIX D
PROOF OF PROPOSITION 5

Let Λ_m be the generalized decision region as defined in (5). For $\mathbf{y} \notin \Lambda_m$, it follows that

$$1 = e^{nT} e^{-nT} \leq e^{nT} \left(\sum_{m' \neq m} \frac{p(\mathbf{y}|\mathbf{x}_{m'})}{p(\mathbf{y}|\mathbf{x}_m)} \right). \tag{68}$$

Let s and ρ satisfy $0 \leq s \leq \rho \leq 1$, and recall the following inequality (see [32, p.197]):

$$\sum_i a_i \leq \left(\sum_i a_i^{\lambda} \right)^{\frac{1}{\lambda}} \tag{69}$$

which holds if $a_i \geq 0$ and $0 < \lambda \leq 1$. Setting

$$a_i = \frac{p(\mathbf{y}|\mathbf{x}_i)}{p(\mathbf{y}|\mathbf{x}_m)}, \quad \lambda = \frac{s}{\rho}$$

it follows from (2), (68) and (69) that the conditional error probability of the m-th message satisfies

$$\begin{aligned} P_{\mathrm{e}|m} &\leq e^{nTs} \sum_{\mathbf{y} \in \Lambda_m^c} p(\mathbf{y}|\mathbf{x}_m) \left(\sum_{m' \neq m} \frac{p(\mathbf{y}|\mathbf{x}_{m'})}{p(\mathbf{y}|\mathbf{x}_m)} \right)^s \\ &\leq e^{nTs} \sum_{\mathbf{y} \in \Lambda_m^c} p(\mathbf{y}|\mathbf{x}_m) \left(\sum_{m' \neq m} \left(\frac{p(\mathbf{y}|\mathbf{x}_{m'})}{p(\mathbf{y}|\mathbf{x}_m)} \right)^{\frac{s}{\rho}} \right)^{\rho}. \end{aligned} \tag{70}$$

Let $\psi_n^m(\mathbf{y})$ designate an arbitrary probability tilting measure (which may depend on the transmitted codeword), then it follows that

$$\begin{aligned} P_{\mathrm{e}|m} &\leq e^{nTs} \sum_{\mathbf{y}} \psi_n^m(\mathbf{y}) \psi_n^m(\mathbf{y})^{-1} p(\mathbf{y}|\mathbf{x}_m) \left(\sum_{m' \neq m} \left(\frac{p(\mathbf{y}|\mathbf{x}_{m'})}{p(\mathbf{y}|\mathbf{x}_m)} \right)^{\frac{s}{\rho}} \right)^{\rho} \\ &\leq e^{nTs} \sum_{\mathbf{y}} \psi_n^m(\mathbf{y}) \left(\psi_n^m(\mathbf{y})^{-\frac{1}{\rho}} p(\mathbf{y}|\mathbf{x}_m)^{\frac{1}{\rho}} \sum_{m' \neq m} \left(\frac{p(\mathbf{y}|\mathbf{x}_{m'})}{p(\mathbf{y}|\mathbf{x}_m)} \right)^{\frac{s}{\rho}} \right)^{\rho}. \end{aligned}$$

Next, invoking Jensen's inequality gives

$$P_{\mathrm{e}|m} \leq e^{nTs} \left(\sum_{\mathbf{y}} \psi_n^m(\mathbf{y})^{1-\frac{1}{\rho}} p(\mathbf{y}|\mathbf{x}_m)^{\frac{1}{\rho}} \sum_{m' \neq m} \left(\frac{p(\mathbf{y}|\mathbf{x}_{m'})}{p(\mathbf{y}|\mathbf{x}_m)} \right)^{\frac{s}{\rho}} \right)^{\rho}.$$

This concludes the proof of (9) by setting

$$\psi_n^m(\mathbf{y}) = \frac{G_n^m(\mathbf{y}) p(\mathbf{y}|\mathbf{x}_m)}{\sum_{\mathbf{y}} G_n^m(\mathbf{y}) p(\mathbf{y}|\mathbf{x}_m)}$$

where $G_n^m(\mathbf{y})$ is an arbitrary non-negative function.

An undetected error event occurs if the received vector is included in the decision region of a codeword which differs from the transmitted codeword. Consequently, the average undetected error event satisfies

$$P_{\mathrm{ue}} = \frac{1}{M} \sum_{m=1}^{M} \sum_{\mathbf{y} \in \Lambda_m} \sum_{m' \neq m} p(\mathbf{y}|\mathbf{x}_{m'}). \tag{71}$$

Note that in the case where list decoding is considered (i.e., the decision regions are not disjoint), the LHS of (71) is no longer a probability. However, for the latter case this expression equals the number of incorrect codewords in the decoded list. It follows from (71) that for $0 \leq s \leq 1$, the undetected error probability satisfies

$$\begin{aligned} P_{\text{ue}} &= \frac{1}{M} \sum_{m=1}^{M} \sum_{\mathbf{y} \in \Lambda_m} p(\mathbf{y}|\mathbf{x}_m) \left(\frac{\sum_{m' \neq m} p(\mathbf{y}|\mathbf{x}_{m'})}{p(\mathbf{y}|\mathbf{x}_m)} \right)^{s} \left(\frac{\sum_{m' \neq m} p(\mathbf{y}|\mathbf{x}_{m'})}{p(\mathbf{y}|\mathbf{x}_m)} \right)^{1-s} \\ &\leq e^{nT(s-1)} \frac{1}{M} \sum_{m=1}^{M} \sum_{\mathbf{y}} p(\mathbf{y}|\mathbf{x}_m) \left(\sum_{m' \neq m} \frac{p(\mathbf{y}|\mathbf{x}_{m'})}{p(\mathbf{y}|\mathbf{x}_m)} \right)^{s} \end{aligned} \qquad (72)$$

where the last inequality holds since for $\mathbf{y} \in \Lambda_m$ and $0 \leq s \leq 1$

$$\left(\frac{p(\mathbf{y}|\mathbf{x}_m)}{\sum_{m' \neq m} p(\mathbf{y}|\mathbf{x}_{m'})} \right)^{1-s} \geq e^{nT(1-s)}.$$

The rest of the proof follows in a similar way to the derivation of (9) when comparing the bound in (70) with (72).

Appendix E
Proof of Corollary 1

Consider the ensemble of fully random block codes of length n symbols where the $M = e^{nR}$ codewords of a codebook are chosen independently at random according to the probability distribution $q_{\mathbf{X}}$ on $\mathcal{X}^n$.

Let $D_{\{\mathbf{x}_i\}_{i=1}^{M}}(m, G_n^m, s, \rho)$ denote the functional $D_{\text{B}}(m, G_n^m, s, \rho)$ in (11) where the dependence on a specific codebook $\{\mathbf{x}_i\}_{i=1}^{M}$ is expressed explicitly. Given a fixed codeword $\mathbf{x}_m$ for the m-th message, the expectation over the other $M-1$ codewords on the right-hand side of (9) gives that for $0 \leq s \leq \rho \leq 1$

$$\begin{aligned} &\sum_{\{\mathbf{x}_i\}_{i=1}^{M} \setminus \{\mathbf{x}_m\}} \left(\prod_{i \neq m} q_{\mathbf{X}}(\mathbf{x}_i) \right) D_{\{\mathbf{x}_i\}_{i=1}^{M}}(m, G_n^m, s, \rho) \\ &\overset{(a)}{\leq} \left(\sum_{\mathbf{y}} G_n^m(\mathbf{y}) p(\mathbf{y}|\mathbf{x}_m) \right)^{1-\rho} \\ &\quad \left(\sum_{m' \neq m} \sum_{\mathbf{x}_{m'}} q_{\mathbf{X}}(\mathbf{x}_{m'}) \sum_{\mathbf{y}} p(\mathbf{y}|\mathbf{x}_m) G_N^m(\mathbf{y})^{1-\frac{1}{\rho}} \left(\frac{p(\mathbf{y}|\mathbf{x}_{m'})}{p(\mathbf{y}|\mathbf{x}_m)} \right)^{\frac{s}{\rho}} \right)^{\rho} \\ &= (M-1)^{\rho} \left(\sum_{\mathbf{y}} G_n^m(\mathbf{y}) p(\mathbf{y}|\mathbf{x}_m) \right)^{1-\rho} \\ &\quad \left(\sum_{\mathbf{x}'} q_{\mathbf{X}}(\mathbf{x}') \sum_{\mathbf{y}} p(\mathbf{y}|\mathbf{x}_m) G_N^m(\mathbf{y})^{1-\frac{1}{\rho}} \left(\frac{p(\mathbf{y}|\mathbf{x}')}{p(\mathbf{y}|\mathbf{x}_m)} \right)^{\frac{s}{\rho}} \right)^{\rho} \end{aligned} \qquad (73)$$

where (a) follows from (11) and by invoking Jensen's inequality. Next, by substituting the non-negative function

$$G_n^m(\mathbf{y}) \triangleq \left(\sum_{\mathbf{x}} q_{\mathbf{X}}(\mathbf{x}) \left(\frac{p(\mathbf{y}|\mathbf{x})}{p(\mathbf{y}|\mathbf{x}_m)} \right)^{\frac{s}{\rho}} \right)^{\rho}$$

in (73), one obtains that for $0 \leq s \leq \rho \leq 1$ and $m = 1, \ldots, M$

$$\begin{aligned} &\sum_{\{\mathbf{x}_i\}_{i=1}^{M} \setminus \{\mathbf{x}_m\}} \left(\prod_{i \neq m} q_{\mathbf{X}}(\mathbf{x}_i) \right) D_{\{\mathbf{x}_i\}_{i=1}^{M}}(m, G_n^m, s, \rho) \\ &\leq (M-1)^{\rho} \sum_{\mathbf{y}} p(\mathbf{y}|\mathbf{x}_m) \left(\sum_{\mathbf{x}'} q_{\mathbf{X}}(\mathbf{x}') \left(\frac{p(\mathbf{y}|\mathbf{x}')}{p(\mathbf{y}|\mathbf{x}_m)} \right)^{\frac{s}{\rho}} \right)^{\rho}. \end{aligned}$$

By averaging $D_{\{\mathbf{x}_i\}_{i=1}^M}(m, G_n^m, s, \rho)$ over the M codewords, we get that for every index m $(1 \leq m \leq M)$

$$\begin{aligned}
&\sum_{\{\mathbf{x}_i\}_{i=1}^M} \left(\prod_{i=1}^M q_{\mathbf{X}}(\mathbf{x}_i)\right) \; D_{\{\mathbf{x}_i\}_{i=1}^M}(m, G_n^m, s, \rho) \\
&= \sum_{\mathbf{x}_m} q_{\mathbf{X}}(\mathbf{x}_m) \sum_{\{\mathbf{x}_i\}_{i=1}^M \backslash \{\mathbf{x}_m\}} \left(\prod_{i \neq m} q_{\mathbf{X}}(\mathbf{x}_i)\right) \; D_{\{\mathbf{x}_i\}_{i=1}^M}(m, G_n^m, s, \rho) \\
&\leq (M-1)^\rho \sum_{\mathbf{y}} \sum_{\mathbf{x}_m} q_{\mathbf{X}}(\mathbf{x}_m) p(\mathbf{y}|\mathbf{x}_m) \left(\sum_{\mathbf{x}'} q_{\mathbf{X}}(\mathbf{x}') \left(\frac{p(\mathbf{y}|\mathbf{x}')}{p(\mathbf{y}|\mathbf{x}_m)}\right)^{\frac{s}{\rho}}\right)^\rho \\
&= (M-1)^\rho \sum_{\mathbf{y}} \left\{ \left(\sum_{\mathbf{x}} q_{\mathbf{X}}(\mathbf{x})\, p(\mathbf{y}|\mathbf{x})^{1-s}\right) \left(\sum_{\mathbf{x}'} q_{\mathbf{X}}(\mathbf{x}') p(\mathbf{y}|\mathbf{x}')^{\frac{s}{\rho}}\right)^\rho \right\}.
\end{aligned} \tag{74}$$

Since the right-hand side of (74) does not depend on the index m, then this bound also applies to the expectation of the quantity $\frac{1}{M}\sum_{m=1}^M D_{\{\mathbf{x}_i\}_{i=1}^M}(m, G_n^m, s, \rho)$. Therefore, there exists a block code for which the value of this quantity is not larger than the average over the considered ensemble, i.e.,

$$\begin{aligned}
&\frac{1}{M} \sum_{m=1}^M \; D_{\{\mathbf{x}_i\}_{i=1}^M}(m, G_n^m, s, \rho) \\
&\leq (M-1)^\rho \sum_{\mathbf{y}} \left\{ \left(\sum_{\mathbf{x}} q_{\mathbf{X}}(\mathbf{x})\, p(\mathbf{y}|\mathbf{x})^{1-s}\right) \left(\sum_{\mathbf{x}'} q_{\mathbf{X}}(\mathbf{x}') p(\mathbf{y}|\mathbf{x}')^{\frac{s}{\rho}}\right)^\rho \right\}.
\end{aligned} \tag{75}$$

From (9), (10) and (75), it follows that the above block code satisfies simultaneously

$$\begin{aligned}
P_{\text{e}} &= \frac{1}{M} \sum_{m=1}^M P_{\text{e}|m} \\
&\leq e^{nsT} \cdot \frac{1}{M} \sum_{m=1}^M \; D_{\{\mathbf{x}_i\}_{i=1}^M}(m, G_n^m, s, \rho) \\
&\leq e^{nsT}(M-1)^\rho \sum_{\mathbf{y}} \left\{ \left(\sum_{\mathbf{x}} q_{\mathbf{X}}(\mathbf{x})\, p(\mathbf{y}|\mathbf{x})^{1-s}\right) \left(\sum_{\mathbf{x}'} q_{\mathbf{X}}(\mathbf{x}') p(\mathbf{y}|\mathbf{x}')^{\frac{s}{\rho}}\right)^\rho \right\} \\
&< e^{n(sT+\rho R)} \sum_{\mathbf{y}} \left\{ \left(\sum_{\mathbf{x}} q_{\mathbf{X}}(\mathbf{x})\, p(\mathbf{y}|\mathbf{x})^{1-s}\right) \left(\sum_{\mathbf{x}'} q_{\mathbf{X}}(\mathbf{x}') p(\mathbf{y}|\mathbf{x}')^{\frac{s}{\rho}}\right)^\rho \right\} \\
&= e^{-n\left(E_0(s,\rho,q_X) - \rho R - sT\right)}
\end{aligned}$$

and

$$\begin{aligned}
P_{\text{ue}} &< e^{n\left((s-1)T+\rho R\right)} \sum_{\mathbf{y}} \left\{ \left(\sum_{\mathbf{x}} q_{\mathbf{X}}(\mathbf{x})\, p(\mathbf{y}|\mathbf{x})^{1-s}\right) \left(\sum_{\mathbf{x}'} q_{\mathbf{X}}(\mathbf{x}') p(\mathbf{y}|\mathbf{x}')^{\frac{s}{\rho}}\right)^\rho \right\} \\
&= e^{-n\left(E_0(s,\rho,q_X) - \rho R - (s-1)T\right)}
\end{aligned}$$

where the last two equalities follow from (15), and since the input distribution and the channel are assumed to be memoryless, i.e.,

$$p(\mathbf{y}|\mathbf{x}) = \prod_{i=1}^n p(y_i|x_i), \quad q_{\mathbf{X}}(\mathbf{x}) = \prod_{i=1}^n q_X(x_i).$$

The proof of Corollary 1 is completed by optimizing the bounds over the parameters ρ and s (where $0 \leq s \leq \rho \leq 1$) and the input distribution q_X. This gives the exponents E_1 and E_2 in (14) for the upper bounds on P_{e} and P_{ue}, respectively.

APPENDIX F
PROOF OF PROPOSITION 6

The bounds in Proposition 6 are derived from Proposition 5 as follows: setting

$$p(\mathbf{y}|\mathbf{x}) = \prod_{i=1}^{n} p(y_i|x_i)$$

and

$$G_n^m(\mathbf{y}) = \prod_{i=1}^{n} g(y_i)$$

in (11), and relying on the useful rule for interchanging sum and product signs $\sum_{\mathbf{y}} \prod_{i=1}^{n} f(y_i) = \prod_{i=1}^{n} \sum_{y_i} f(y_i)$, one gets from (9) the RHS of (16) as an upper bound on $P_{\text{e}|0}$. Since the considered block code is linear and the communication channel is memoryless and symmetric, the bound in (16) follows from the message independence property in Proposition 2. The derivation of the bound in (17) relies on (10) where it is first proved that for a linear block code whose transmission takes place over a memoryless symmetric channel, the resulting expression for $D_{\text{B}}(m, G_n^m, s, \rho)$ is independent of m. To this end, let $\mathcal{T}$ be a mapping as defined in Definition 1, then for all $1 \leq i \leq n$

$$\begin{aligned}
&\sum_{m' \neq m} \sum_{y \in \mathcal{Y}} g(y)^{1-\frac{1}{\rho}} p(y|x_{m,i}) \left(\frac{p(y|x_{m',i})}{p(y|x_{m,i})} \right)^{\frac{s}{\rho}} \\
&= \sum_{m' \neq m} \sum_{y \in \mathcal{Y}} g(y)^{1-\frac{1}{\rho}} p(\mathcal{T}(y, -x_{m,i})|0) \left(\frac{p(\mathcal{T}(y, -x_{m,i})|x_{m',i} - x_{m,i})}{p(\mathcal{T}(y, -x_{m,i})|0)} \right)^{\frac{s}{\rho}} \\
&= \sum_{l \neq 0} \sum_{z \in \mathcal{Y}} g(y)^{1-\frac{1}{\rho}} p(z|0) \left(\frac{p(z|x_{l,i})}{p(z|0)} \right)^{\frac{s}{\rho}}.
\end{aligned}$$

As a result, it follows that for a memoryless and symmetric channel

$$\frac{1}{M} \sum_{m=1}^{M} D_{\text{B}}(m, G_n^m, s, \rho) = D(g, s, \rho) \tag{76}$$

where $D(g, s, \rho)$ is introduced in (18). The proof of the upper bound on P_{ue} as given in (17) is completed by substituting (76) in (17).

APPENDIX G
PROOF OF PROPOSITION 7

Let Λ_m^{LR} designate the decision region in (7), then for $\mathbf{y} \notin \Lambda_m^{\text{LR}}$

$$\frac{p(\mathbf{y}|\mathbf{x}_m)}{p(\mathbf{y}|\mathbf{x}_{m_2})} < e^{nT}$$

where $\mathbf{x}_{m_2}$ is the second most probable codeword. Hence, for $s \geq 0$, the conditional block error probability satisfies

$$P_{\text{e}|m} \leq e^{nsT} \sum_{\mathbf{y}} p(\mathbf{y}|\mathbf{x}_m) \sum_{m' \neq m} \left(\frac{p(\mathbf{y}|\mathbf{x}_{m'})}{p(\mathbf{y}|\mathbf{x}_m)} \right)^s. \tag{77}$$

The bound in (41) follows from (77), using the arguments following (70).

For $\mathbf{y} \in \Lambda_{m'}^{\text{LR}}$ where $m' \neq m$, it follows from (7) that

$$\frac{p(\mathbf{y}|\mathbf{x}_{m'})}{p(\mathbf{y}|\mathbf{x}_{m_2'})} \geq e^{nT}$$

where $\mathbf{x}_{m_2'}$ is the second most probable codeword given the received vector $\mathbf{y}$ at the channel output. As a result, the conditional undetected block error probability satisfies, for all $s \geq 0$, the following upper bound:

$$P_{\text{ue}|m} \leq e^{-nsT} \sum_{\mathbf{y}} p(\mathbf{y}|\mathbf{x}_m) \sum_{m' \neq m} \left(\frac{p(\mathbf{y}|\mathbf{x}_{m'})}{p(\mathbf{y}|\mathbf{x}_m)} \right)^s.$$

The rest of the proof of (42) is, again, similar to the derivation following (70).

APPENDIX H
PROOF OF PROPOSITION 8

The derivation of the bounds in Proposition 8 is primarily identical to the analysis in [8] and [24, Section 3.2.4], for which the reader is referred for a complete treatment of the analysis under ML decoding. We assume a BPSK modulation over AWGN channel with energy E_{s} per transmitted coded symbol, and a white Gaussian noise with two-sided power spectral density of $\frac{N_0}{2}$. Hence, the received vector $\mathbf{y}$ satisfies

$$\mathbf{y} = \gamma \mathbf{x} + \mathbf{n} \tag{78}$$

where $\gamma \triangleq \sqrt{\frac{2E_{\text{s}}}{N_0}}$, $\mathbf{x} \in \mathcal{C} \subseteq \{-1,+1\}^n$ is the transmitted codeword (with BPSK modulation), and $\mathbf{n}$ is a normal random vector with independent coordinates (all with zero mean and unit variance). Setting

$$E_{\text{e}}(d) \triangleq \left\{ \mathbf{y} \in \mathcal{Y}^n : \ \frac{\max_{\mathbf{x} \in \mathcal{C}_d \setminus \{\mathbf{x}_0\}} p(\mathbf{y}|\mathbf{x})}{p(\mathbf{y}|\mathbf{x}_0)} \cdot e^{nT} \geq 1 \right\}.$$

where $\mathcal{C}_d$ is the set of all codewords whose Hamming weight is d, and $\mathbf{x}_0$ is the all-zero codeword, it follows from (7) and the union bound that the conditional decoding error probability is upper bounded by

$$P_{\text{e}|0} \leq \sum_{d=d_{\min}}^{n} \Pr\left(E_{\text{e}}(d)\right) \tag{79}$$

where d_{min} denotes the minimal Hamming distance of $\mathcal{C}$. Consider the following inequality on the probability of an error event:

$$\Pr(E) \leq \Pr(E, \mathbf{y} \in \mathcal{R}) + \Pr(\mathbf{y} \notin \mathcal{R}) \tag{80}$$

where E denotes an error event, $\mathbf{y} \in \mathcal{Y}^n$ is the received vector, and $\mathcal{R} \subseteq \mathbf{Y}^n$. From (79) and (80), it follows that

$$P_{\text{e}|0} \leq \sum_{d=d_{\min}}^{n} \Big(\Pr\big(E_{\text{e}}(d), \mathbf{y} \in \mathcal{R}\big) + \Pr\big(\mathbf{y} \notin \mathcal{R}\big) \Big). \tag{81}$$

Using the union bound, we have

$$\begin{aligned} \Pr\big(E_{\text{e}}(d), \mathbf{y} \in \mathcal{R}\big) &\leq \sum_{\mathbf{x} \in \mathcal{C}_d} \Pr\left(\frac{p(\mathbf{y}|\mathbf{x})}{p(\mathbf{y}|\mathbf{x}_0)} \, e^{nT} \geq 1, \ \mathbf{y} \in \mathcal{R} \right) \\ &\overset{\text{(a)}}{=} \sum_{\mathbf{x} \in \mathcal{C}_d} \Pr\left(\langle \mathbf{y}, \mathbf{x} \rangle \geq \langle \mathbf{y}, \mathbf{x}_0 \rangle - \frac{nT}{\gamma}, \ \mathbf{y} \in \mathcal{R} \right) \end{aligned} \tag{82}$$

where equality (a) follows from (78), and $\langle \mathbf{x}, \mathbf{y} \rangle \triangleq \sum_{i=1}^{n} x_i y_i$ denotes the scalar multiplication of the vectors $\mathbf{x}$ and $\mathbf{y}$. Similarly to the derivation of bound in [8] (under ML decoding), we choose

$$\mathcal{R} \triangleq \Big\{ \mathbf{y} : \ \|\mathbf{y} - \eta \gamma \mathbf{x}_0\|^2 \leq nr^2 \Big\} \tag{83}$$

where η and r are arbitrary parameters which are subject to optimization. In addition, define

$$\begin{aligned} Z &\triangleq \langle \mathbf{y}, \mathbf{x} \rangle - \langle \mathbf{y}, \mathbf{x}_0 \rangle \\ W &\triangleq \|\mathbf{y} - \eta \gamma \mathbf{x}_0\|^2 - nr^2 \end{aligned}$$

then it follows from (82) and (83), using the Chernoff bound that

$$\Pr\big(E_{\text{e}}(d), \ \mathbf{y} \in \mathcal{R}\big) + \Pr\big(\mathbf{y} \notin \mathcal{R}\big) \leq e^{\frac{tnT}{\gamma}} \, |\mathcal{C}_d| \, \mathsf{E}\Big[e^{tZ+uW}\Big] + \mathsf{E}\Big[e^{sW}\Big] \tag{84}$$

for all $t \geq 0$, $u \leq 0$, and $s \geq 0$. Evaluating the expectations in (84) and setting $t = \frac{\gamma}{2}(1 - 2u\eta)$, we have similarly to [8] and [24, Section 3.2.4]:

$$\begin{aligned} \Pr\big(E_{\text{e}}(d), \ \mathbf{y} \in \mathcal{R}\big) + \Pr\big(\mathbf{y} \notin \mathcal{R}\big) \leq \; & e^{\frac{nT(1-ru\eta)}{2}} \, |\mathcal{C}_d| \, e^{-nur^2} \left(f_1\left(\gamma, u, \eta\right)\right)^{n-d} \left(f_2\left(\gamma, u, \eta\right)\right)^{d} \\ & + e^{-nsr^2} \left(f_1\left(\gamma, s, \eta\right)\right)^{n} \end{aligned} \tag{85}$$

where

$$f_1(\gamma,\alpha,\eta) \triangleq \frac{e^{\frac{(1-\eta)^2\gamma^2\alpha}{1-2\alpha}}}{\sqrt{1-2\alpha}}$$
$$f_2(\gamma,\alpha,\eta) \triangleq \frac{e^{-\frac{\gamma^2(1-2\alpha\eta^2)}{2}}}{\sqrt{1-2\alpha}}, \quad \alpha < \frac{1}{2}.$$

Optimizing the term e^{nr^2} on the right-hand side of (85), gives

$$\Pr\big(E_{\mathrm{e}}(d), \mathbf{y} \in \mathcal{R}\big) + \Pr\big(\mathbf{y} \notin \mathcal{R}\big) \le 2^{h_2\left(\frac{s}{s-u}\right)} A^{-\frac{u}{s-u}} B^{\frac{s}{s-u}}, \quad 0 < s < \frac{1}{2},\ u \le 0 \tag{86}$$

where

$$A \triangleq \left(f_1(\gamma,s,\eta)\right)^n$$
$$B \triangleq e^{\frac{nT(1-ru\eta)}{2}} |\mathcal{C}_d| \left(f_1(\gamma,u,\eta)\right)^{n-d} \left(f_2(\gamma,u,\eta)\right)^d$$

and h_2 designates the binary entropy function on base 2. Using the change of variables

$$\rho \triangleq \frac{s}{s-u}$$
$$\beta \triangleq \rho(1-2u)$$
$$\xi \triangleq \rho(1-2u\eta)$$

where $0 \le \rho \le 1, 0 \le \beta \le 1$, and $\xi \ge 0$, the bound in (86) transforms to

$$\Pr\big(E_{\mathrm{e}}(d),\ \mathbf{y} \in \mathcal{R}\big) + \Pr\big(\mathbf{y} \notin \mathcal{R}\big) \le 2^{h_2(\rho)} e^{-nE(E_{\mathrm{s}}/N_0, d/n, \beta, \rho, \xi) + \frac{nT\xi}{2}} \tag{87}$$

where

$$E(c,\delta,\beta,\rho,\xi) \triangleq -\rho r_n(\delta) - \frac{\rho}{2}\ln\left(\frac{\rho}{\beta}\right) - \frac{1-\rho}{2}\ln\left(\frac{1-\rho}{1-\beta}\right) + c\left(1-(1-\delta)\frac{\xi^2}{\beta} - \frac{(1-\xi)^2}{1-\beta}\right).$$

The parameters ρ, β and ξ are optimized in [8], [24] such that the error exponent $E(c,\delta,\beta,\rho,\xi)$ is maximized[3] (note that the bound for $T=0$ coincides with the bound which refers to ML decoding), setting the optimal parameters yields the first argument in (47). The second term inside the minimization on the right-hand side of (47) follows from a union bound on the error probability

$$P_{\mathrm{e}} \le \sum_{d=d_{\min}}^{n} \sum_{\mathbf{x}\in\mathcal{C}_d} \Pr\left(\frac{p(\mathbf{y}|\mathbf{x})}{p(\mathbf{y}|\mathbf{x}_0)} e^{nT} \ge 1\right)$$

where for every codeword $\mathbf{x} \in \mathcal{C}_d$

$$\Pr\left(\frac{p(\mathbf{y}|\mathbf{x})}{p(\mathbf{y}|\mathbf{x}_0)} e^{nT} \ge 1\right) = Q\left(\gamma\sqrt{d} - \frac{nT}{2\gamma\sqrt{d}}\right).$$

The derivation of the upper bound on the undetected error probability follows some similar arguments, and is therefore omitted.

[3] It is possible to obtain the optimized ρ and ξ when maximizing the entire exponent $E(c,\delta,\beta,\rho,\xi) + \frac{T\xi}{2}$. To this end, ξ needs to be shifted by $-\frac{T}{2}$ and the optimal ρ remains without change. The parameter β is required to be numerically optimized over $0 \le \beta \le 1$. Nevertheless, the resulting bound gives only a marginal gain over the bound which maximizes $E(c,\delta,\beta,\rho,\xi)$ without the addition of $\frac{T\xi}{2}$.

APPENDIX I
PROOF OF PROPOSITION 9

The main ingredient for proving the DS2 bound on the block error probability under ML decoding (and also the well known random-coding bound) is that for a received vector $\mathbf{y}$ which is not included in the decision region Λ_m as given in (3), the following inequality holds:

$$1 \leq \left(\sum_{m' \neq m} \left(\frac{p(\mathbf{y}|\mathbf{x}_{m'})}{p(\mathbf{y}|\mathbf{x}_m)} \right)^{\lambda} \right)^{\rho}, \quad \lambda, \rho \geq 0. \tag{88}$$

When an error event under fixed-size (L) list decoding is considered, there exists L distinct codewords, all different from the transmitted codeword, whose a-posterior probability is larger than the one of the transmitted codeword. Hence, the sum on the right-hand side of (88) is divided by L. Specifically for a received vector $\mathbf{y}$ that results in an error event, the following inequality is satisfied:

$$1 \leq \left(\frac{1}{L} \sum_{m' \neq m} \left(\frac{p(\mathbf{y}|\mathbf{x}_{m'})}{p(\mathbf{y}|\mathbf{x}_m)} \right)^{\lambda} \right)^{\rho}, \quad \lambda, \rho \geq 0 \tag{89}$$

Following the derivation of the DS2 bound in [24, p. 96] where the right-hand side of (88) is replaced with (89) leads to the derivation of the bound in Proposition 9.

REFERENCES

[1] C. Bai, B. Mielczarek, W. A. Krzymien, and I. J. Fair, “Improved analysis of list decoding and its application to convolutional codes and turbo codes,” *IEEE Trans. on Information Theory*, vol. 53, no. 2, pp. 615–627, February 2007.

[2] A. Barg, “Improved error bounds for the erasure/list scheme: The binary and spherical cases,” *IEEE Trans. on Information Theory*, vol. 50, no. 10, pp. 2503–2511, October 2004.

[3] E. L. Blokh and V. V. Zyablov, *Linear Concatenated Codes* (in Russian). Moscow, U.S.S.R.: Nauka, 1982.

[4] I. E. Bocharova, R. Johannesson, B. D. Kudryashov, and M. Loncar, “An improved bound on the list error probability and list distance properties,” *IEEE Trans. Information Theory*, vol. 54, no. 1, pp. 13–32, January 2008.

[5] D. Burshtein and G. Miller, “Asymptotic enumeration methods for analyzing LDPC codes,” *IEEE Trans. on Information Theory*, vol. 50, no. 6, pp. 1115–1131, June 2004.

[6] G. Caire, S. Shamai and S. Verdu, “Feedback and belief propagation,” *Proceedings of the 2006 Turbo Coding Symposium*, Munich, Germany, April 3–7, 2006.

[7] T. M. Cover and J. A. Thomas, *Elements of Information Thoery*, New. York: Wiley, 1991.

[8] D. Divsalar, “A simple tight bound on error probability of block codes with application to turbo codes,” the Telecommunications and Mission Operations (TMO) Progress Report 42–139, JPL, pp. 1–35, November 15, 1999. [Online]. Available: http://tmo.jpl.nasa.gov/progress_report/42-139/139L.pdf.

[9] S. C. Draper and A. Sahai, “Variable-length channel coding with noisy feedback,” *European Tran. on Teleccomunications*, vol. 19, no. 4, pp. 335–370, April 2008.

[10] T. M. Duman, *Turbo Codes and Turbo-Coded Modulation Systems: Analysis and Performance Bounds*, Ph.D. dissertation, Elect. Comput. Eng. Dep., Northeastern University, Boston, MA, USA, May 1998.

[11] T. M. Duman and M. Salehi, “New peformance bounds for turbo codes,” *IEEE Trans. on Communications*, vol. 46, no. 6, pp. 717–723, June 1998.

[12] P. Elias, “List decoding for noisy channels”, in *Proc. IRE WESCON Conf. Rec.*, vol. 2, pp. 94–104, 1957.

[13] G. D. Forney, “Exponential error bounds for erasure, list, and decision feedback schemes,” *IEEE Trans. Information Theory*, vol. 14, no. 2, pp. 206–220, March 1968.

[14] R. G. Gallager, *Low-density parity-check codes*, MA, USA:MIT Press, 1963.

[15] –, “A simple derivation of the coding theorem and some applications,” *IEEE Trans. on Information Theory*, vol. 11, pp. 3–18, January 1965.

[16] P. K. Gopala, Y. H. Nam, and H. El Gamal, “On the error exponents of ARQ channels with deadlines,” *IEEE Trans. on Information Theory*, vol. 53, no. 11, pp. 4265–4273, November 2007.

[17] T. Hashimoto, “Composite shceme LR + Th for decoding with erasures and its effective equivalence to Forney’s rule,” *IEEE Trans. on Information Theory*, vol 45, no. 1, pp. 78–93, January 1999.

[18] E. Hof, I. Sason, and S. Shamai (Shitz), “Performance bounds for non-binary linear block codes over memoryless symmetric channels,” *IEEE Trans. on Information Theory*, vol. 55, no. 3, pp. 977–996, March 2009.

[19] –, “Optimal generalized decoding of convolutional codes,” accepted to the *Tenth International Symposium on Communication Theory and Applications (ISCTA 09)*, Ambleside, UK, July 13-17, 2009.

[20] B. D. Kudryashov, “List decoding in a Gaussian channel,” *Prob. Inf. Transm.*, vol. 27, no. 3, pp. 30-38, July-September, 1991.

[21] N. Merhav, “Error exponents of erasure/list decoding revisted via moments of distance enumerators,” *IEEE Trans. on Information Theory*, vol. 54, no. 10, pp. 4439–4447, October 2008.

[22] A. Sahai, *Anytime information theory*, Phd. dissertation, Massachusetts Institute of Technology, 2001.

[23] A. Sahai and S. K. Mitter, “The necessity and sufficiency of anytime capacity for stabilization of a linear system over a noisy communication link. Part I: scalar systems,” *IEEE Trans. on Information Theory*, vol. 52, no. 8, pp. 3369-3395, August 2006.

[24] I. Sason and S. Shamai, *Performance Analysis of Linear Codes under Maximum-Likelihood Decoding: A Tutorial*, Foundations and Trends in Communications and Information Theory, vol. 3, no. 1–2, pp. 1–222, June 2006. [Online]. Available: http://www.ee.technion.ac.il/people/sason/monograph.html.

[25] S. Shamai and I. Sason, “Variations on the Gallager bounds, connections and applications,” *IEEE Trans. on Information Theory,* vol 48, no. 12, pp. 3029–3051, December 2002.

[26] C. Shannon, R. Gallager and E. Berlekamp, “Lower bounds to error probability for decoding on discrete memoryless channels,” *Information and Control*, vol. 10, Part 1: pp. 65–103, and Part 2: pp. 522–552, February/May 1967.

[27] N. Shulman, and M. Feder, “Random coding techniques for nonrandom codes,” *IEEE Trans. on Information Theory*, vol. 45, pp. 2101–2104, September 1999.

[28] E. Telatar and R. G. Gallager, “New exponential upper bounds to error and erasure probabilities,” *Proceedings 1994 IEEE International Symposium on Information Theory (ISIT 1994)*, p. 379, Trondheim, Norway, June 1994.

[29] E. Telatar, “Exponential bounds for list size moments and error probability,” *Proceedings 1998 IEEE Information Theory Workshop (ITW 1998)*, p. 60, Killarney, Ireland, June 1998.

[30] S. Tong, “Tangential-sphere bounds on the ensemble performance of ML decoded Gallager codes via their exact ensemble distance spectrum,” *Proceedings 2008 IEEE International Conference on Communications (ICC 2008)*, pp. 1150–1154, Beijing, China, May 19–23, 2008.

[31] A. J. Viterbi, “Error bounds for the white Gaussian and other very noisy memoryless channels with generalzied decision regions,” *IEEE Trans. Information Theory*, vol. 15, no. 2, pp. 279–287, March 1969.

[32] A. J. Viterbi and J. K. Omura, *Principle of Digital Communication and Coding*, 1979.

[33] G. Wiechman and I. Sason, “An improved sphere-packing bound for finite-length codes on symmetric memoryless channels,” *IEEE Trans. on Information Theory*, vol. 54, no. 5, pp. 1962–1990, May 2008.

[34] J. M. Wozencraft, “List decoding,” in Quaternary Progress Report, Cambridge, MA: MIT Research Lab. Elctron., vol. 48, pp. 90–95, 1958.

[35] H. Yamamoto and K. Itoh, “Asymptotic performance of a modified Schalkwijk-Barron scheme for channels with noiseless feedback,” *IEEE Trans. on Information Theory*, vol. 25, no. 6, pp. 729–733, November 1979.