Performance Bounds and Distance Spectra of Variable Length Codes in Turbo/Concatenated Systems

Xavier Jaspar, Graduate Student Member, IEEE, and Luc Vandendorpe, Fellow, IEEE

Abstract—Variable length codes (VLCs), used in data compression, are very sensitive to error propagation in the presence of noisy channels. Addressing this problem with joint source-channel turbo techniques has been pursued in the literature and looks quite promising. But to date, most code-related conclusions are based on simulations. This paper states and proves several theoretical results about the robustness of prefix VLCs concatenated with linear error correcting codes (ECC), assuming a maximum likelihood decoder. Especially, an approximate and asymptotically tight distance spectrum of the concatenated code (VLC+ECC) is rigorously developed. Together with the union bound, it provides upper bounds on the symbol and frame/packet error rates.

I. INTRODUCTION

The potential of joint source-channel turbo (de)coding with a VLC as source code has been illustrated for the first time in [1], based on the serial concatenation of a VLC with a convolutional code (CC). This concatenation has then been improved in several directions, notably the VLC decoder [2], [3], the VLC itself [4]–[6] and the inner error correcting code [6]–[8] to cite a few. So far, however, most code-related conclusions have been based on simulations or on extrinsic information analysis, and a theoretical framework on the error correcting properties of VLCs in turbo/concatenated systems has been lacking.

The first attempt toward such a framework was in [9], where the author analyzed a transmission system consisting of a VLC stream alone (not framed, not concatenated) and provided expressions of the VLC stream spectrum and of the union bound for the Levenshtein symbol error rate. The second attempt was the extension of these results, in [5] and later independently in [6], to the serial concatenation of a VLC with an interleaver and a linear ECC. With a serial concatenation, the VLC stream is “framed” into an equivalent VLC block of \(N\) bits before entering the interleaver of length \(N\). Because of this, considering simply the VLC stream as in [9] is not sufficient. Instead, the framing must be taken into account to evaluate the spectrum of the VLC block (though not explicitly stated in [5]), just as is done in [10] with the “equivalent block code” of CCs in turbo codes. In [5], [6], expressions are given for the spectrum of the VLC block and for the spectrum of the concatenated code (VLC, interleaver, ECC) in order to obtain performance bounds. Unfortunately, in these contributions, the results are given without development and without proof, as simplistic combinations/extensions of previous results from [9], [10]. Important details are missing.

In this paper, we carry these contributions one step further by developing more accurate results as well as new results, and by proving them rigorously. We introduce the important concept of bounded VLC spectrum. Next, for VLCs with bounded spectrum, we develop an approximate expression of the distance spectrum of VLC blocks that is asymptotically tight with the correct spectrum as the interleaver length increases. At last, an expression of the distance spectrum of the global code (VLC + interleaver + ECC) is given. Together with the union bound, this spectrum provides eventually upper bounds on the bit, Levenshtein symbol and frame error rates. Because the union bound classically diverges at low signal to noise ratios (SNRs) on the channel, these upper bounds are untight at these SNRs. But, at moderate to large SNRs, they are tight with the simulation results and can thus be used for performance prediction.

The remainder of this paper is structured as follows. A few background results, the generic communication system and several assumptions are provided in Sections II–III. Spectra and bounds for concatenated VLC blocks are developed in Section IV. Related work in this field is discussed at the end of Section IV. Simulation results are reported in Section V and extensions of this work are discussed in Section VI.

II. BACKGROUND AND NOTATIONS

Random variables are written with capital letters and realizations with small letters. \(P(z)\) is the abbreviation of the probability \(P(Z = z).\) \(E\{\}\) is the expectation. The sequence \((Z_m, Z_{m+1}, \ldots, Z_n)\) is written \(Z_{m:n}\), or \(\hat{Z}\), when \(m, n\) can be omitted. \(\hat{Z}\) is the decision taken on \(Z\) at the receiver. \(\mathcal{A}\) is the alphabet of the source symbols. \(d_H(., .)\), \(d_S(., .)\) and \(d_{S_L}(., .)\) denote the Hamming, symbol and Levenshtein [11] symbol distances, \(l(.)\) and \(l_S(.)\) the bit length and symbol length, \(l\) and \(\sigma^2\) the VLC length average and variance. \(l_{gcd}\) is the greatest common divisor of the VLC lengths, \(gcd\{l(s) : s \in \mathcal{A}\}\), and \(l_{gcd}^* = gcd\{l(s) : s \in \mathcal{A}, P(s) > 0\}\). Besides, \(l_{\min} = \min\{l(s) : s \in \mathcal{A}\}\), \(l_{\max} = \max\{l(s) : s \in \mathcal{A}\}\), \(l_{\min}^* = \min\{l(s) : s \in \mathcal{A}, P(s) > 0\}\), \(l_{\max}^* = \max\{l(s) : s \in \mathcal{A}, P(s) > 0\}\). \(I(a)\) equals 1 iff \(a\) is true. \(|\mathcal{A}|\) is the cardinality of the set \(\mathcal{A}\). \(\mathbb{Z}\) is the set of integers and \(\mathbb{N}_{\geq 0} = \{n \in \mathbb{Z} : n \geq 0\}\), \(\mathbb{N}_{> 0} = \{n \in \mathbb{Z} : n > 0\}\).
A. Prefix VLCs, source of symbols and Markov model

Let $\mathcal{A} = \{0, 1\}$. A finite sequence $w = b_1 b_2 \ldots b_m$ of code letters $b_j \in \mathcal{A}$ is a word over $\mathcal{A}$, of length $l(w) = m$. Let $\mathcal{B}$ be the empty word, with $l(\mathcal{B}) = 0$ and $Rw = w = wR$ for any word $w$. Let $\mathcal{B}^+ \triangleq \cup_{i \geq 1} \mathcal{B}^i$. Let $\text{prefix}(w) = \{ p \in \mathcal{B}^+ : \forall v \in \mathcal{B}^+, w = pv \}$ be the set of all prefixes of $w$. A code is a set $\mathcal{V} \subset \mathcal{B}^+$ and its elements are called codewords. If no codeword is the prefix of another codeword, then $\mathcal{V}$ is a prefix VLC.

If, besides, no codeword is the suffix of another codeword, then the VLC is said reversible. In the following, only prefix VLCs are considered and we omit “prefix” for convenience. For any code $\mathcal{V}$, let $\text{prefix}(\mathcal{V}) \triangleq \cup_{w \in \mathcal{V}} \text{prefix}(w)$. Given a source of discrete, independent (memoriless) and identically distributed symbols over the alphabet $\mathcal{A}$, with $|\mathcal{A}| = |\mathcal{V}| \geq 2$, and characterized by the probability distribution $P(S)$, let $\text{vlc}(.)$ be a bijective mapping that associates to each symbol $s \in \mathcal{A}$ a unique codeword $w = \text{vlc}(s) \in \mathcal{V}$. The inverse mapping is denoted $\text{vlc}^{-1}(.)$. In the following, this mapping is always assumed implicitly: we use concepts associated with codewords equivalently with symbols and vice versa, e.g., $l(s) = l(\text{vlc}(s))$, $P(w) = P(\text{vlc}^{-1}(w))$. At last, let $\mathcal{V}^+ \triangleq \cup_{i \geq 1} \mathcal{V}^i$ and $\mathcal{A}^+ \triangleq \cup_{i \geq 1} \mathcal{A}^i$.

The source/VLC can be modeled as a Markov binary source with state space $\mathcal{X} = \{ \mathcal{B} \} \cup \text{prefix}(\mathcal{V})$, state $X_n \in \mathcal{X}$ and transition probability $P(U_n|X_n).$ The $n$th bit $U_n$ is associated with the transition from $X_{n-1}$ to $X_n$. The probability $P(U_n|X_{n-1})$ can be deduced from $P(S)$, $\text{vlc}(.)$ and $\mathcal{V}$. This model corresponds to the Balakirsky trellis [12], in Fig. 1, and $R$ is then the so-called root state.

At last, we introduce the notation $s_{m:n} \equiv u_{i:j} \equiv x_{i-1:j}$ to indicate that $u_{i:j} = \text{vlc}(s_{m:n}), l(s_{m:n}) = i - 1, l(s_{1:n}) = j, x_{i-1} = R$ and $P(x_n|x_{n-1}, u_n) = 1$ for $i \leq n \leq j$, that is, to indicate that $s_{m:n}, u_{i:j}$ and $x_{i-1:j}$ refer to the same realization of the source/VLC/Markov model.

B. Distance spectrum

Given a block code $C$, let $A_{w,h}^C$ be the average, over all codewords $c$, of the number of codewords $c$ of weight $w$. Upper (union) bounds on the frame and bit error rates (FER, BER) of frame-ML (maximum likelihood) decoding of a linear block code $C$ with $N$ information bits, are [10]

$$\text{FER} \leq \sum_{w,h} A_{w,h}^C P_h, \quad \text{BER} \leq \sum_{w,h} \frac{w!}{N} A_{w,h}^C P_h,$$

where $P_h$ is the pairwise error probability which is the conditional probability that a given codeword $c$, at distance $d_H(c, \hat{c}) = h$ from the transmitted codeword $c$, has a larger likelihood metric than $c$. For the additive white Gaussian noise (AWGN) channel with the BPSK modulation, $P_h = \frac{1}{2} \text{erfc} \left( \sqrt{h R_c E_b/N_0} \right)$, where $R_c$ is the global code rate, $N_0/2$ is the double-sided noise spectral density and $E_b$ the energy per bit of entropy. For the binary symmetric channel (BSC) with error probability $p$, $P_h \leq \left(4(1-p)p^{h/2} \right)$.

For the evaluation of the bounds (1), the distance spectrum is required. The abstract concept of uniform interleaver is introduced and methods are developed in [10] to compute efficiently the distance spectrum of concatenated linear codes. To summarize, consider a serially concatenated code $C$ whose outer and inner linear component codes are respectively $C_o$ and $C_i$. The interleaver length is $N$. Then,

$$A_{w,h}^C = \sum_{l=0}^{N} \frac{A_{w,l}^C A_{l,h}^C}{\binom{N}{l}}.$$

III. TRANSMISSION SYSTEM AND ASSUMPTIONS

A. Generic system encoder

Consider the source/VLC of Section II-A producing semi-infinite symbol stream $S_{1:}\infty$ and bit stream or VLC stream $U_{1:}\infty = \text{vlc}(S_{1:}\infty)$. In Fig. 2(a), the VLC stream is directly sent across the memoryless channel, without framing, without interleaver and without ECC. In Fig. 2(b), the VLC stream is framed into finite-length VLC blocks (frames or packets) of $N$ bits before the interleaving by $\Pi$ of length $N$. This framing is done according to a certain framing rule. We now
introduce the framing rule $F_b$ and focus on it—see also the rule $F_s$ in [13], more pedagogical but less practical, based on a constant number of symbols. Assume the VLC stream has been processed up to a certain symbol, and the symbol and bit indices have been shifted such that the next symbol to process is $S_1$ and the next bit is $U_1$. Then, given a fixed $N$ constant for all VLC blocks, the framing rule $F_b$ forms the next VLC block with $S_{1:N_s}$ ($N_s$ is variable) subject to

$$N_b = l(S_{1:N_s}) \leq N, \quad N_b + l(S_{N+1}) > N, \quad (3)$$

where $S_{N+1}$ becomes the first symbol of the next VLC block. If $N_b < N$, $N - N_b$ zeros are appended to $U_1$ to get $U_1$.

At the receiver, the frame-MAP (maximum a posteriori) detection is considered,

$$\hat{s} = \arg\max_{s} \{P(y, s)\}, \quad (4)$$

where $y$ is the received signal. In practice, for the system in Fig. 2(b), the frame-MAP detection is prohibitively complex because of the product code source/VLC + ECC. In the simulation results, we will use instead a frame-MAP joint source-channel iterative decoder, as a good approximation. It can be deduced [14] from the application of the sum-product algorithm (SPA) on the factor graph of the complete transmission scheme, followed by the Viterbi algorithm on the Balakirsky trellis.

B. Link with the proposed performance bounds

The proposed performance bounds in this paper are based on the ML (maximum likelihood) detection

$$\hat{s} = \arg\max_{s} \{P(y, s)\}, \quad (5)$$

which does not take into account the a priori probabilities of the source sequences compared to the MAP detection (4).

C. Assumptions

Here is the framework of assumptions used thereafter. Generalizations are discussed in Section VI.

**Assumption 1** (source/VLC): Let us consider a stationary memoryless source and a prefix VLC, as in Section II-A, such that $\exists s', s'' \in A$, $s' \neq s''$, $P(s') > 0$, $P(s'') > 0$. \hfill $\square$

**Assumption 2** (global system): In addition to Assumption 1, we consider in Fig. 2(b) finite-length frames $U_{1:N_b}$ and $\text{VLC}(S_{1:N_s})$, a uniform interleaver [10] of length $N$, a linear ECC and a frame-ML decoder (5). We assume furthermore that the decoder knows (and uses) the values of $N_b$, $N$ and not the value of $N_s$. Thus $N_b < N$, the decoder knows besides that the bits $U_{N_b+1:N}$ are zeros. \hfill $\square$

IV. PERFORMANCE BOUNDS

A. Definition of error event

The synchronization right before the bit $U_1$ between a transmitted stream $S_{1:∞} \equiv U_{1:∞} \equiv X_{0:∞}$ and a decoded stream $S_{1:∞} \equiv U_{1:∞} \equiv X_{0:∞}$ is characterized by the equality $X_{i-1} = \hat{x}_{i-1}$. In particular, the equality $X_{i-1} = \hat{x}_{i-1} = R$ implies that both streams have a codeword starting with $U_i$. Therefore, if there is no bit error after $U_i$, i.e., $U_{1:∞} = U_{1:∞}$, then all subsequent codewords are decoded correctly (up to a possible time shift). This is the idea underlying the concept of error event, see Fig. 1: An error event is generated by some errors and, in this paper, we consider that the error event starts with the beginning of its first erroneous codeword, say in $X_{i-1} = \hat{x}_{i-1} = R$ for some $i$, and ends as soon as $X_j = \hat{x}_j = R$ with $j$ the smallest integer $j > i$. Between $X_{i-1}$ and $X_j$, the equality $x_l = \hat{x}_l = R$ is not satisfied, so $i - 1$ and $j$ are the only elements of the set $\{l : i - 1 \leq l \leq j, x_l = \hat{x}_l = R\}$.

**Definition 3**: Under Assumption 2, the pair of sub-sequence $s_{m:n} = u_{i:j}$ and $s'_{m':n'} = \hat{u}_{i:j}$ form an error event $e = (u_{i:j}, \hat{u}_{i:j})$—by convention, we consider that this $e$ starts with $u_i$ or $s_{m:n}$ or $s'_{m':n'}$ if $u_{i:j} \neq \hat{u}_{i:j}$ and $i - 1, j$ if $e = (u_{i:j}, \hat{u}_{i:j})$ and $l = i - 1, j$ if $e = (u_{i:j}, \hat{u}_{i:j})$. By extension of previous notations, let $L(e) = I_{s_{m:n}}, L(e) = I_{s'_{m':n'}}$, $l(e) = I_{s_{m:n}}(e) = j - i + 1, d_{L}(e) = d_{L}(u_{i:j}, \hat{u}_{i:j})$ and $d_{H}(e) = d_{H}(u_{i:j}, \hat{u}_{i:j})$. \hfill $\square$

**Definition 4**: For $u_-, \hat{u}_- \in \mathcal{Y}^+$ with $l(u_-, \hat{u}_-) = l(\hat{u}_-, u_-)$, let $e_{\hat{u}_-, u_-} = (\hat{u}_-, u_-)$ the sequence of all error events formed by sub-sequences of $u_-$ and $\hat{u}_-$, that is, let it be the sequence $e_{1} = (u_1, \hat{u}_1), e_2 = (u_2, \hat{u}_2), \ldots, e_n = (u_n, \hat{u}_n)$ for some $n$ where $u_1, \hat{u}_1, u_2, \hat{u}_2, \ldots, \hat{u}_n, u_n \in \mathcal{Y}^+$ subject to $u_i = v_{0,i}^0, v_{1,i}^0, v_{2,i}^0, \ldots, v_{n,i}^0$ and $\hat{u}_i = v_{0,i}^0, v_{1,i}^0, v_{2,i}^0, \ldots, v_{n,i}^0$ with $v_0, v_1, \ldots, v_0 \in \mathcal{Y}^+$ union $\{R\}$. For some $\hat{s} = u_-, \hat{s} = u_-$, let $e_{\hat{s},(\hat{s}_-)'} = e_{\hat{s},(\hat{s}_-)'}$. \hfill $\square$

**Remark 5**: The probability is one that codewords in $U_{1:∞}$ (but not necessarily in $U_{1:∞}$ since the ML decoder may decode codewords of zero probability), and thus error events, start at bit positions $l \equiv 1 \mod (s_{\text{gcd}})$ by definition of $s_{\text{gcd}}$. Note, if $s_{\text{gcd}} = 1$, then $l \equiv 1 \equiv 0 \mod (s_{\text{gcd}})$ for all $l$. \hfill $\square$

**Example 6**: Consider the source alphabet $\mathcal{A} = \{a, b\}$, with $P(S = a) = p \in (0, 1)$ and $P(S = b) = 1 - p$, and the VLC $\{0, 11\}$ with VLC($a$) = 0 and VLC($b$) = 11. Consider the
following \( u_s, \tilde{u}_s \); and bit errors in positions 1, 4, 8 and 14:

\[
\begin{align*}
\begin{array}{c}
\begin{array}{ccccccccccc}
\begin{array}{cccccccccc}
0 & 1 & 1 & 1 & 1 & 0 & 0 & 0 & 0 & 1 & 1 & 1 & 0 \\
\hline
\text{a} & \text{b} & \text{a} & \text{a} & \text{a} & \text{b} & \text{b} & \text{b} & \text{a} \end{array}
\end{array}
\end{array}
\end{align*}
\]

\[
\begin{align*}
\begin{array}{c}
\begin{array}{ccccccccccc}
\begin{array}{cccccccccc}
1 & 1 & 1 & 1 & 0 & 0 & 0 & 0 & 1 & 1 & 1 & 1 \\
\hline
\text{b} & \text{b} & \text{b} & \text{a} & \text{a} & \text{a} & \text{b} & \text{b} & \text{a} \end{array}
\end{array}
\end{array}
\end{align*}
\]

There are two error events, \((aba, bb)\) starting with \(u_1\) and \((bbba, abb)\) starting with \(u_8\). There is no symbol error in between if we tolerate a symbol time shift. \(\square\)

### B. VLC stream spectrum

For a given realization \(u_{1:t} \equiv s_{1:m}\), let \(P(u_{1:t}) = P(s_{1:m})\) be the probability of observing \(s_{1:m}\) among all sequences of the same symbol length, i.e., \(P(u_{1:t}) = P(s_{1:m}) = \prod_{k=1}^{m} P(S = s_k)\).

**Definition 7:** The distance spectrum of a VLC stream, or **VLC stream spectrum**, is under Assumption 2

\[
A^{\text{vlc}}_{s_L,h,l,s} \triangleq \sum_{s_1, l \in Z^+ \text{ s.t. } (s_1, l) = l_k} P(s_{1:l}) A^{\text{vlc}}_{s_L,h,s_1:l_k},
\]

where \(A^{\text{vlc}}_{s_L,h,s} \triangleq \left\{|\tilde{s} \in \mathbb{Z}^+: d_{sL}(\tilde{s}, s) = h\} \right\}\) is the conditional spectrum given \(s\). Unnecessary subscripts may be marginalized for convenience, e.g.,\(A^{\text{vlc}}_{s_L,h} = \sum_{l_k} A^{\text{vlc}}_{s_L,h,l,s} \).\(\square\)

Given some fixed position \(k \in \mathbb{N}_{>0}\) in the transmitted stream \(S_{1:2^h}\), the spectrum \(A^{\text{vlc}}_{s_L,h,l,s}\) can be interpreted as the number of possible error events \(e\) starting with \(S_k\) with \(d_{sL}(e) = s_L, d_H(e) = h, l_S(e) = l_s, l_H(e) = l_h\), averaged over all possible \(S_{1:2^h}\). In particular, the VLC stream spectrum counts single error events and not combinations of several error events, unlike the VLC block spectrum in Section IV-F. Single error events are indeed sufficient to predict the performance of a VLC stream alone across a memoryless channel (Fig. 2(a)): A Levenshtein error at a given symbol \(s_k\) is triggered only by single error events involving \(S_k\) and is independent of error events appearing elsewhere in the stream. By contrast, for the concatenated system in Fig. 2(b), an error event involving \(S_k\) might be influenced by error events at other positions because of the (memory introduced by the) interleaver and the ECC.

**Remark 8 (language abuse):** The VLC stream spectrum is more properly the spectrum of the triplet formed by the VLC, the mapping \(\text{vlc}(\cdot)\) and the source. It depends notably on the source statistics. \(\square\)

**Definition 9:** The free distance of a VLC is given by \(d^{\text{vlc}}_f = \min\{h \in \mathbb{N}_{>0}: A^{\text{vlc}}_h \neq 0\}\).

In terms of spectrum, the sensitivity of VLCs to residual bit errors translates into the existence of arbitrarily large error events, i.e., large \(l_h, l_s\) with \(A^{\text{vlc}}_{s_L,h,l,s} > 0\) for small \(h\). Fortunately, generally in practice, \(A^{\text{vlc}}_{s_L,h,l,s} \to 0\) exponentially as \(l_s \to \infty\). To characterize this, we introduce the concept of bounded spectrum.

**Definition 10:** \(A^{\text{vlc}}_{s_L,h,l,s} \) is bounded if \(\exists c < 1, \forall h, 3l_h^c (\text{depending on } h), \forall l_h \geq l_h^c, A^{\text{vlc}}_{s_L,h,l,s} \leq c^h\). \(\square\)

The concept of bounded spectrum is important because it implies that long error events can be neglected in the proposed upper bounds; it is an assumption of Th. 14. A VLC has a bounded spectrum when [15] it has a synchronizing sequence \(u_0^c \) with \(P(u_0^c) > 0\); roughly speaking, when such a \(u_0^c\) exists, there is an increasing probability that \(u_0^c\) is received as more bits are received — the probability of not receiving \(u_0^c\) is exponentially decreasing —, forcing the decoder to synchronize and thus making error events shorter.

**Example 11:** Coming back to Example 6, we can calculate \(A^{\text{vlc}}_{s_L,h,l,s}\) analytically. By carefully enumerating all \(s\) involved in (6), we get: If \(s_k = h = 2\), then \(A^{\text{vlc}}_{s_L,h,l,s}\) is equal to \(p^2(1-p)^{l_r-2}\) when \(l_h = 2l_s - 2 \geq 2\), to \(2p(1-p)^{l_r-1}\) when \(l_h = 2l_s - 1 \geq 2\), to \((1-p)^{l_r}\) when \(l_h = 2l_s \geq 2\). Otherwise, it is equal to 0. The spectrum is bounded and, by summing over all \(l_s, l_h\), we get \(A^{\text{vlc}}_{s_L=2,h=2} = 1 + \frac{2}{p} - p\). \(\square\)

### C. Bounds for non-concatenated VLC streams

Based on the VLC stream spectrum, upper bounds can be developed for non-concatenated VLC streams (Fig. 2(a)). The upper bound on the Levenshtein [11] symbol error rate,

\[
\text{SER}_{L} \leq \sum_{h, s_L=1} P_h s_L A^{\text{vlc}}_{s_L,h},
\]

appeared originally in [9], where \(P_h\) is defined in Section II-B. Other bounds have been proved in [13, Th. 5.18–5.22] and an extension to the (non-Levenshtein) symbol error rate is considered in Section VI-A.

**Example 12:** Coming back to Example 11, \(\text{SER}_{L} \leq 2 P_h = (1 + \frac{1}{p} - p)\). \(\square\)

### D. Numerical evaluation of the VLC stream spectrum

There is no straightforward simplification to evaluate numerically \(A^{\text{vlc}}_{s_L,h,l,s}\). The expression in Def. 7 must be evaluated as is, i.e., by computing the conditional spectrum for all possible \(s\). —for a linear code, we need only the weight distribution, i.e., the conditional spectrum for the all-zero codeword. In practice, with an impact on the tightness of the performance bounds, we can limit the complexity by computing \(A^{\text{vlc}}_{s_L,h,l,s}\) up to some \(h\), if we are interested in large SNR analysis, and up to some \(l_s, l_h\), if \(c\) is small in Def. 10 — \(c\) is smaller if the VLC has [15] a larger probability to resynchronize. We refer to [9], where two algorithms are proposed to evaluate the quantity \(\sum_{s_L} s_L A^{\text{vlc}}_{s_L,h}\) and can be extended easily to \(A^{\text{vlc}}_{s_L,h,l,s}\).

### E. Framing rule, definition and properties

In Fig. 2(b), the VLC stream \(S_{1:2^h} \equiv U_{1:2^h}\) is framed into a semi-infinite sequence of VLC blocks \((S_{1:N}^1 S_{2:N}^1 S_{3:N}^1 \cdots U_{1:2^N}^1 U_{1:2^{N+1}}^1 \cdots )\) according to a certain framing rule \(F\). The successive \(S_{1:N}^m \equiv U_{1:N}^m, m \geq 1\), constitute a stochastic process which we refer to as \(B_{\text{vlc}}^m\) or \(B_{\text{vlc}}\) when \(m\) can be omitted — with transition probability \(P(B_{\text{vlc}}^m|B_{\text{vlc}}^{m-1}, F)\) and stationary probability \(P(B_{\text{vlc}}|F)\).

Let’s focus on the framing rule \(F_6\) (Section III-A). By (3), the values of \(N_b\) of non-zero probability satisfy

\[
N_b \equiv 0 \pmod{2^{l_g(c)}}, \quad N - l^{\text{max}}_b < N_b \leq N.
\]

(8)
The maximum value of $N_b$ of non-zero probability is thus $N_b^{\text{max}} \triangleq \frac{t_{\text{gcd}}^*}{t_{\text{gcd}}} [N/l_{\text{gcd}}^*]$. By definition of $F_b$, $B_{\text{vlc}}^m$ depends on the value of $N_b$ of $B_{\text{vlc}}^{m-1}$, which we denote $N_b^{(m-1)}$. Straightforwardly, under Assumption 2,
\[
P(B_{\text{vlc}} = s|n_b^{(1)}, F_b) = \frac{\prod_{k=1}^{l_b(s)} P(S = s_k)}{P(l(S) > N - n_b^{(1)})} \]
and, for $n_b^{(1)}$ subject to (8),
\[
P(n_b^{(1)} | F_b) = \frac{\prod_{l} P(l(S) > N - n_b^{(1)})}{\prod_{l} (N_{b}^{\text{max}} + \frac{t_{\text{gcd}}}{2} - \frac{l}{2})} \leq O(\theta{n_b^{(1)}}),
\]
where $0 \leq \theta < 1$. See Lemma 20 for the value of $\theta$.

By last, by (8), the admissible values of $N_b$ are concentrated [13] around $N_b \approx N_b^{\text{max}} + \frac{t_{\text{gcd}}}{2} - \frac{l}{2}$ for $N > l_{\text{max}}$, and the most probable values of $N_b$ are concentrated around $N_b/N_{b}^{\text{max}}$ for large $N$ with $N_s/N_b \rightarrow |N|/2$ as $N \rightarrow +\infty$.

### F. VLC block spectrum with framing rule $F_b$

**Definition 13:** Under Assumption 2, the distance spectrum of a VLC block $B_{\text{vlc}}$ with framing rule $F$ is
\[
A_{s_h,l,s,h}^F \triangleq \sum_s P(B_{\text{vlc}} = s|F) A_{s_h,l,s,h}^F,
\]
where the conditional spectrum $A_{s_h,l,s,h}^F$ is given by
\[
A_{s_h,l,s,h}^F = \left\{ \hat{u} : d_H(s, \hat{s}) = d_S(l, \hat{l}) = s, l \right\}.
\]

In other words, $A_{s_h,l,s,h}^F$ counts the codewords $\hat{u}$ of the VLC block with $d_S(l, \hat{u}) = s$ and $d_H(u, \hat{u}) = h$, averaged over all possible $u$. The VLC block spectrum is thus not limited to single error events since the pair $(u, \hat{u})$ may contain several error events, unlike the VLC stream spectrum.

Instead of evaluating (10) exactly, which is intractable, a transformation is now given to estimate the VLC block spectrum approximately from the VLC stream spectrum, similarly as in [10] for CCs.

**Theorem 14:** Given a VLC stream with bounded spectrum $A_{s_h,l,s,h}^{\text{vlc}}$, and the framing rule $F_b$, the VLC block spectrum $A_{s_h,l,s,h}^{F_b}$ can be approximated for small $h$ and asymptotically large $N$ under Assumption 2 by
\[
A_{s_h,l,s,h}^{F_b,\text{app}} \triangleq \sum_{l_b \leq N_b^{\text{max}}} \left( \frac{N_b^{\text{max}}}{t_{\text{gcd}}^*} - \frac{l}{t_{\text{gcd}}^*} + n \right) \left( \frac{t_{\text{gcd}}}{l} \right)^n T_{s_h,l,s,h,l,b,n}^{\text{vlc}}
\]
where $T_{s_h,l,s,h,l,b,n}^{\text{vlc}}$ is the coefficient of the polynomial
\[
T_{s_h,l,s,h,l,b,n}^{\text{vlc}}(S_L, H, L_s, L_b, \Omega) = \sum_{s_h,l,s,h,l,b,n} \left( A_{s_h,l,s,h,l,b}^{\text{vlc}}(S_L, H, L_s, L_b) \right)^n,
\]
with $A_{s_h,l,s,h,l,b}^{\text{vlc}}(S_L, H, L_s, L_b) \triangleq \sum_{l_b_l,h} A_{s_h,l,s,h,l,b}^{\text{vlc}} S_{s_h}^{L_s} H^h L_{s}^b L_{b}^h \Omega^h$. See proof in Appendix B.

Since $A_{s_h,l,s,h}^{F_b,\text{app}}$ is asymptotically tight with the real spectrum $A_{s_h,l,s,h}^{F_b}$ for small $h$ and large $N$, it can be used in Section IV-G to get tight performance predictions at high SNRs for concatenated systems with long interleavers. But, strictly speaking, we cannot guarantee that these predictions are “upper bounds”, because of the approximations made in the proof to obtain $A_{s_h,l,s,h}^{F_b,\text{app}}$. To ensure (strict) upper bounds, we introduce the concept of upper spectrum.

**Definition 15:** We say $A_{s_h,l,s,h}^{F_b,\text{app}}$ is an upper spectrum of the VLC block and we write $A_{s_h,l,s,h}^{F_b} \leq A_{s_h,l,s,h}^{F_b,\text{app}}$ if $\sum_{s_h \geq s} A_{s_h,l,s,h}^{F_b,\text{app}}$ is an upper bound on the performance for any $h, s \in \mathbb{N}_{\geq 0}$.

By induction, it follows that $\sum_{s_h \geq s} A_{s_h,l,s,h}^{F_b,\text{app}}$ is upper bound on the BER, SER$_{\text{app}}$, and FER.

**Theorem 16:** Under Assumption 2, given a VLC stream with bounded spectrum $A_{s_h,l,s,h}^{F_b,\text{app}}$, $A_{s_h,l,s,h}^{F_b,\text{app}} = \sum_{l} \left( \frac{N_b^{\text{max}}}{t_{\text{gcd}}^*} - \frac{l}{t_{\text{gcd}}^*} + n \right) \left( \frac{t_{\text{gcd}}}{l} \right)^n T_{s,h,l,b,n}^{\text{vlc}}$ is an upper spectrum for the framing rule $F_b$. See proof in [13, Section 5.C.3.4].

The differences between (12) and (14) are due to small differences in the corresponding proofs. To prove (12) in Appendix B, some probabilities are replaced by their asymptotic values as a first approximation (e.g., $P(\hat{u})$ by $t_{\text{gcd}}^*/l$ from Lemma 20). To prove (14), they are simply upper bounded by 1. Note, for fixed length codes (FLCs), $A_{s_h,l,s,h}^{F_b,\text{app}}$ is equal to $A_{s_h,l,s,h}^{F_b,\text{app}}$ and is almost exact; the only approximation is due to Lemma 21.

### G. VLC block concatenated with a linear code

Given the spectrum $A_{s_h,l,s,h}^{F_b}$ of a VLC block for a given framing rule $F$, we can use (2) to calculate the spectrum $A_{s_h,w,h}^{\text{gc}}$ of the global code (GC) in Fig. 2(b), i.e., the spectrum of the VLC block serially concatenated with the linear ECC. Indeed, by linearity of the ECC, it is self-evident that (2) holds with the conditional spectrum, i.e.,
\[
A_{s_h,w,h}^{\text{gc}}(w) = \sum_{s_s} P(s|F) A_{s_s,w,s}^{\text{gc}}(w).
\]

The summation of (2) is not used here so as to keep the dependence w.r.t. the number $w$ of VLC bit errors. Then, by (10) and by linearity of (15) in $A_{s_h,w,h}^{\text{gc}}$, $A_{s_h,w,h}^{\text{gc}} = \sum_{s} P(s|F) A_{s_s,w,s}^{\text{gc}} = \sum_{s} A_{s_s,w,s}^{\text{gc}} \frac{\text{A}_{\text{ecc}}}{w}$. By linearity again, this spectrum can be used in the union bounds (1) to get performance bounds on the frame-ML decoding of the global code, as in the next theorem.
Theorem 17 (Concatenated system): The frame-ML decoding performance of the system in Fig. 2(b) is upper bounded under Assumption 2 by the union bounds

\[ \text{FER} \leq \sum_{h} P_h \sum_{s,l,w} A_{s,l,w,h}^{\text{rc}} \]  
\[ \text{SER}_L \leq \sum_{h} P_h \sum_{s,l,w} N_{\text{min}}^{-1} A_{s,l,w,h}^{\text{rc}} \]  
\[ \text{BER} \leq \sum_{h} P_h \sum_{s,l,w} \frac{w}{N_{\text{min}}} A_{s,l,w,h}^{\text{rc}} \]  

where $N_{\text{min}}$ and $N_{\text{min}}^{-1}$ are the minimum admissible values of $N_s$ and $N_b$ of non-zero probability. \hfill \blacksquare

Corollary 18: The upper bounds (18)-(19) become approximate performance estimations if we replace $A_{s,l,h}^{\text{rc,app}}$ instead of the exact $A_{s,l,h}^{\text{rc}}$. To ensure strict upper bounds (but untight!), the upper spectrum $A_{s,l,h}^{\text{rc,app}}$ from (14) must be used instead.

H. Numerical evaluation of spectra and bounds for concatenated VLCs

1) Tight predictions at large SNRs with $A_{s,l,h}^{\text{rc,app}}$: Step 1, the stream spectrum $A_{s,l,h,l,b}^{\text{vlc}}$ (up to some $h$ and $b$) in Section IV-D. Step 2, the intermediate spectrum $T_{s,l,h,b,l,n}^{\text{vlc}}$ (up to some $n$), whose polynomial expression (13) can be implemented with multi-dimensional convolutions or FFTs. Step 3, the approximate VLC block spectrum $A_{s,l,h}^{\text{rc,app}}$ in (12). Step 4, the global concatenated code $A_{s,l,w,h}^{\text{rc}}$ in (16) with $A_{s,l,h}^{\text{rc}} \approx A_{s,l,h}^{\text{rc,app}}$. Step 5, performance estimations in Th. 17 with the slight modification suggested in Corollary 18.

2) Predictions with single event approximation: The computation-intensive steps are typically Steps 1 and 2. Step 1 is unavoidable. Step 2 can be avoided if (i) one is interested only in large SNR analysis, i.e., in the error floor region, and if (ii) the error floor region is dominated by errors of the concatenated code corresponding to single error events of the VLC. Condition (ii) depends on the ECC: for (parallel turbo codes based on) recursive CCs, it holds if $d_{\text{min}}^{\text{le}} \geq 2$; for low density parity check codes, it depends on the so-called trapping set. If (i) and (ii) are satisfied, then only single error events ($n = 1$) matter in Th. 14. If besides the interleaver is large compared to the “boundness” of the VLC spectrum (Def. 10), only the terms for small $l_s, l_b$ are significant and (12) becomes

\[ A_{s,l,h}^{\text{rc,app, single-event}} \approx \frac{N_{\text{max}}}{l} A_{s,l,h}^{\text{vlc}}. \]  

3) Strict upper bounds: Upper bounds can be obtained with Section IV-H1 if (i) no truncation is made at Steps 1 and 2, (ii) $A_{s,l,h}^{\text{rc,app}}$ is used instead of $A_{s,l,h}^{\text{rc,app}}$ at Step 3 and (iii) Corollary 18 is not used at Step 5. These bounds are however useful only for theoretical purposes, because item (i) is not tractable in practice for moderate interleaver sizes with today’s computers and knowledge.

I. Related work

To the best of our knowledge, the distance spectrum of VLC blocks concatenated with linear ECCs (Fig. 2(b)) has been first introduced in [5], and later independently in [6]. The authors in [5] consider a different framing rule, say $F_b$, which given a fixed $N_s$ forms the next VLC block with $S_{1:N_s}$, and sets $N = N_b = l(S_1:N_s)$ — compared to $F_h$, it requires to constantly adapt the interleaver and the ECC in Fig. 2(b) because $N$ depends on $S_1:N_s$. Due to the randomness of $N$, it is suggested in [5, eq. (5)] to develop the global code spectrum conditionally to $N$ and then to average it over $N$, i.e., with our notations, $A_{s,l,w,h}^{\text{rc,app}} = \sum_{n=0}^{N_{\text{max}}} P(n_b|F_h) A_{s,l,w,h}^{\text{rc,app}}(n_b|F_h) A_{s,l,w,h}^{\text{rc,app}}(n_b|F_h)$, where $N$ has been replaced by $N_b$ since $N = N_b$, $A_{s,l,w,h}^{\text{rc,app}}$ is the ECC spectrum subject to $n_b$ input bits, and $A_{s,l,w,h}^{\text{rc,app}}$ is the conditional VLC block spectrum given $N_b = n_b$. Unfortunately, this computation is difficult because, to the best of our knowledge, there is no simple approximate transformation such as (12) to calculate $A_{s,l,h}^{\text{rc,app}}$. The authors in [5] suggest as a first approximation to consider only the average interleaver length $N_{\text{av}}$ but leave unclear the calculation of $A_{s,l,h|n_b}^{\text{rc,app}}$, which is required. They present only the final result [5, eq. (9)], with our notations and extended with the Levenshtein symbol distance $s_l$.

\[ A_{s,l,h}^{\text{rc,app}} \leq \sum_{n=1}^{N} T_{s,l,h,n}. \]  

To illustrate their bound with simulation results, they use the VLC stream spectra computed in [16]. These spectra are however too marginalized (in the sense of Def. 7) and do not tight with the simulation results but, hopefully, go in the direction of an “upper” bound (which is desirable).

With the developments of this section, we can complete the results of [5] for the framing rule $F_b$. By the central limit theorem, the most probable values of $N_b = l(S_1:N_s)$ are highly concentrated around $N_s^\text{av}$ for large $N_s$. Thus $A_{s,l,w,h}^{\text{rc,app}} \approx A_{s,l,h}^{\text{rc,app}} A_{s,l,h}^{\text{rc,app}}(N_b|w)/N_b$, where $A_{s,l,h}^{\text{rc,app}}$ can be calculated approximately by Th. 14 with $N_b^{\text{max}} \approx N_b^\text{av}$ — or almost exactly by [13, Th. 5.26], the only approximation being due to Lemma 21.

V. Simulation results

The tightness of the proposed bounds at moderate and large SNRs is illustrated for the SER in Fig. 3, using the steps from Section IV-H1. The system considered consists in the source generators and the VLCs listed in Table I, framed according to $F_b$, serially concatenated with a uniform interleaver of size $N = 384$ and with a recursive systematic punctured CC encoder with generators $(07,05)_8$, where $07_8$ is the feedback. The parity bits of the CC are punctured to get a channel code rate of 2/3. The
At last, the tightness of the proposed bounds for the BER and the FER is also illustrated in Fig. 5 with another source, the “English alphabet” (see [4]), and stronger codes: (i) a Huffman VLC serially concatenated with an interleaver of size 4000 and a rate-1/2 parallel turbo code, (ii) a RVLC $d_{\text{f}}^{\text{vlc}} = 2$ with the same serial concatenation, (iii) an al-prompt VLC [9, Table A.1] with $d_{\text{f}}^{\text{vlc}} = 3$ serially concatenated with an interleaver of size 1000 and a rate-2/3 recursive CC. The bounds are tight at high SNRs. This time, however, notice how the bounds—more properly, the truncated estimations of the bounds—are below the simulations results at low SNRs. This is typical of stronger serially concatenated (other examples in [10]): These systems exhibit at low SNRs error rates dominated by error events with large Hamming distance $h$ that are discarded by truncating in practice the estimations of the spectra to some $h$ and $l_b$, thereby making the estimations of the bounds lower at low SNRs. Without truncation, the proposed bounds would have diverged upwards at low SNRs (because of the union bound), just as in Fig. 4.

VI. EXTENSION AND FURTHER WORK

A. The symbol error rate (SER)

With FLCs, the SER is straightforwardly tightly upper bounded by (7) and (18). But with VLCs that are not FLCs, the SER deserves some attention. It has a meaning only when some synchronization mechanism is in place, for example when the symbol counter is resynchronized after each VLC block. In that example, the SER is always (loosely) upper bounded by the FER.

Let us focus on VLC blocks and define the SER as

$$\mathbb{E}\left\{d_{\text{f}}(S_1;L_{\text{min}},\hat{S}_1;L_{\text{min}}) + I_{\text{S}}(\hat{S}_1) - I_{\text{S}}(S_1)\right\}$$

with $L_{\text{min}} = \min\{I_{\text{S}}(\hat{S}_1), I_{\text{S}}(S_1)\} - I_{\text{S}}(\hat{S}_1)$ and $l_b(\hat{S}_1)$ may be different by Assumption 2. Next, recall that in Section IV-A, the decoder was considered synchronized in $S_k$ if $X_n = X_n$ with $n = l(S_1;L_{\text{min}})$, i.e., if $l(S_1;k-1) = l(S_1;k-1)$ for some $k'$. For the SER, the symbol clocks $k$ and $k'$ must be synchronized and thus equal. But integrating this into Def. 3 leads to too many arbitrarily long error events.

Instead, we can loosely upper bound the SER of a non-concatenated VLC block with an extension of the VLC stream spectrum from Def. 7. Let $A_{s,k,s,d,e,h,t,b}^{\text{vlc}} \triangleq \sum_{s_1,i,\bar{e} \in \mathcal{E}} P_{s_1}(s_1) P_{i,d,e,h.t,b}^{\text{vlc}}$, with

$$A_{s,k,s,d,e,h,t,b}^{\text{vlc}} \triangleq \left\{ (s, \hat{s}, d_{s}, d_{e}, h) : d_{s}(s, \hat{s}) = d_{s}(s, \hat{s}) = d_{e}(s, \hat{s}) = d_{e}(s, \hat{s}) = d_{e}(s, \hat{s}) \right\}$$

and $d_{e}(s, \hat{s}) = l_s(\hat{s}) - l_s(s)$. When $d \neq 0$, the decoder is desynchronized and we obtain a loose upper bound by considering that the decoder produces symbol errors until the end of the frame (at most $N_{\text{max}}^s$), thus

$$\text{SER} \leq \sum_{h,s} P_h(s) A_{s,d,e=0,h}^{\text{vlc}} + N_{\text{max}}^s \sum_{d=0} A_{s,d,e,h}^{\text{vlc}}$$

by extension of (7). This bound can be tightened by replacing $N_{\text{max}}^s$ with an average value. At last, this reasoning can be extended to upper bound the SER of VLC blocks concatenated with linear codes, with a global code spectrum $A_{s,c,d,e,w,h}^{\text{vlc}}$ obtained by extensions of Th. 14, (16) and Th. 17.

### TABLE I

<table>
<thead>
<tr>
<th>Prob.</th>
<th>VLC</th>
<th>RVLC $d_f=1$</th>
<th>RVLC $d_f=2$</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.33</td>
<td>11</td>
<td>00</td>
<td>10</td>
</tr>
<tr>
<td>0.30</td>
<td>10</td>
<td>01</td>
<td>01</td>
</tr>
<tr>
<td>0.18</td>
<td>00</td>
<td>10</td>
<td>000</td>
</tr>
<tr>
<td>0.10</td>
<td>011</td>
<td>11</td>
<td>111</td>
</tr>
<tr>
<td>0.09</td>
<td>010</td>
<td>11011</td>
<td>11000</td>
</tr>
</tbody>
</table>

**Fig. 3.** Simulation results (solid lines) and union bounds (dashed lines) on the SER, for different VLCs listed in Table I and concatenated with a punctured recursive systematic CC encoder.
B. Performance bounds for $N_s$ known at the decoder

Knowing the number of symbols $N_s$ at the decoder generally increases the resilience at the expense of a higher decoding complexity. Compared to Section VI-A, the decoder knows the value of $N_s$, and thus $d_L(s; s) = 0$, that is, the error events in the set $\mathcal{E}(s; s)$ always satisfy $\sum_{c \in \mathcal{E}(s; s)} d_L(c) = 0$. Upper bounds on the BER, the SER, and the FER then follow from the extension of Th. 17 suggested at the end of Section VI-A if we restrict the summations to $d = 0$, i.e., to $A_{sL, d}^{\text{ VLC}}$. A loose upper bound on the SER is possible as well by noticing that all codeword pairs enumerated by $A_{sL, d, e, w, h}^{\text{ VLC}}$ with $e > 0$ form at least one error event with $d \neq 0$. Since $d \neq 0$, this error event desynchronizes the decoder and at most all symbols in the VLC block are erroneously decoded. Therefore, as an extension of (18),

$$
\text{SER} \leq \sum_{s,h} P_h \left( \frac{s}{N_s^{\min}} A_{s,d=0,c=0,h}^{\text{ VLC}} + \sum_{e>0} A_{s,d=0,c,e,h}^{\text{ VLC}} \right).
$$

C. Extension to the MAP decoder

With the frame-MAP decoder, the pairwise error probability is the probability that a given $\hat{s}$ has a larger a posteriori metric than the transmitted $s$. It depends thus on the a priori distance between $s$ and $\hat{s}$. We define the a priori distance between $s$ and $\hat{s}$ as $d_A(s, \hat{s}) \triangleq \log(P(s)) - \log(P(\hat{s}))$.

Then, the MAP pairwise error probability for the AWGN channel can be proved to be

$$
P_{h,a} = \frac{1}{2} \text{erfc} \left( \sqrt{h R_c E_b / N_0} + \frac{a}{4\sqrt{h R_c E_b / N_0}} \right) \quad (24)
$$

where $h = d_H(c, \hat{c})$ is the Hamming distance between the codewords associated with $s$ and $\hat{s}$, and $a = d_A(s, \hat{s})$. For the BSC with error probability $p$,

$$
P_{h,a} = \sum_{j=\max(1+|j^*|,0)}^h \binom{h}{j} p^j (1-p)^{h-j} + \frac{1}{2} \mathbb{1}\{ j^* \in \mathbb{Z}, 0 \leq j^* \leq h \} \binom{h}{j^*} p^{j^*} (1-p)^{h-j^*} \quad (25)
$$

where $j^* = (a / (1-p) + h) / 2$. See [13] for details.

Extending the bounds to the MAP decoder for VLC streams is thus as simple as adding the a priori distance $a = d_A(s, \hat{s})$ to the VLC spectrum from Def. 7, which leads to $A_{sL,b,a,h}^{\text{ VLC}}$. The free distance becomes $d_{L,\text{map}}^{\text{ VLC}} = \min \{ h \in \mathbb{N}_0 : \exists a \in \mathbb{R}, A_{h,a}^{\text{ VLC}} \neq 0 \}$ and the upper bound (7) becomes

$$
\text{SER}_L \leq \sum_{h \geq 1} \sum_{a \geq 0} \sum_{s \geq 1} s_L A_{sL,h,a}^{\text{ VLC}}. \quad (26)
$$
where \(A_{s_l,h,a}^{\text{vle}} = \sum_{d_l,h,a} A_{s_l,h,a}^{\text{vle}}\). In practice, we need some simplifications, though. The a priori distance is a real number and thus the triplet \((s_l, h, a)\) takes too many different values for practical computer storage of \(A_{s_l,h,a}^{\text{vle}}\). This can be solved easily by quantizing the a priori distance and limiting its range of values, for example with \(q(d_A(s, \hat{s}))\) for some function \(q(\cdot)\). Note, the upper bounds need to remain upper bounds despite \(q(\cdot)\). We can show that any odd function, \(q(d) = -q(-d)\), fulfills this requirement.

At last, for concatenated VLC blocks, the extension is very similar. We simply have to incorporate the a priori distance in the spectra and in the bounds, with obvious extensions of (13), (12) and (16). Note the extension of (13) is possible because \(d_A(s, \hat{s}) = \sum_{e \in E(s, \hat{s})} d_A(e)\).

D. Extension to sources with memory

Extending the results to sources with memory is straightforward in theory but can be difficult in practice. The evaluation of the spectrum has indeed a computational complexity that increases with the state space of the Markov model. Nonetheless, the complexity decreases when the source memory is associated with stronger error correcting capabilities because we can then limit the evaluation to shorter error events.

To summarize the extension, note the root state probability in the proofs must be extended to the probability to finish a given codeword at a given position, for all codewords. Also, an error event ends when both the transmitted and decoded sequences finish the same codeword at the same bit position. At last, considering sources with memory does not make sense without the MAP extension from Section VI-C since the ML bounds do not really take the source memory into account.

VII. Conclusions

Several theoretical results on the resilience of prefix VLCs concatenated with linear ECCs have been stated and proved, assuming an optimal maximum likelihood (ML) decoder. In a sense, this work pursues previous contributions [5, 6] one step further, by developing more accurate results as well as new results, and by proving them rigorously. We focused mainly on the ML decoder and on memoryless sources. Extensions to sources with memory and to the MAP decoder have been discussed in Section VI but further work in this direction would be interesting. Though not discussed, another interesting extension would be to replace the union bound with the Tangential Sphere Bound [17] to make the bounds tighter at lower SNRs. In addition, all theoretical results have been developed in the same framework of assumptions for the sake of clarity (Section III-C) but many results, taken apart, are certainly extensible to less restrictive assumptions.

Appendix

A. Root state probability, \(P(\rho_i)\)

The probability to start a codeword at a given bit position, or equivalently to be in the root state in the Balakirsky trellis, is an important quantity in the sequel.

**Definition 19:** Given \(S_{1:}\infty \equiv U_{1:}\infty \equiv X_{0:}\infty\), let \(\rho_i\) be the event that \(U_{1:i-1} \in \mathcal{V}^+, i.e., X_{i-1} = R\).

By definition, \(P(\rho_1) = 1\) and \(P(\rho_i) = P(\rho_i|\rho_{i-1})\). Besides, by stationarity and periodicity of \(X_n\), we have \(P(\rho_i|\rho_{i-1}) = P(\rho_{i-j+1}|\rho_{i-j})\) for \(j < i\) and \(j \equiv 1 \pmod{L_{\text{gcd}}}\).

**Lemma 20:** Under Assumption 1, the root state probability

\[
P(\rho_i) = \begin{cases} 
\frac{l_{\text{gcd}}^*}{l_{\text{gcd}}} + O(\delta^i), & \text{if } i \equiv 1 \pmod{L_{\text{gcd}}}, \\
0, & \text{otherwise},
\end{cases}
\]

where \(\theta = |\lambda_2|^{-1}/L_{\text{gcd}} < 1\) and \(\lambda_2\) is the root of second highest absolute value of the polynomial \(c(\lambda) = L_{\text{max}}/L_{\text{gcd}}\left(1 - \sum_{s \in \mathcal{E}} P(s)\lambda^{-l(s)}/L_{\text{gcd}}\right)\). If \(l(s) = l_{\text{max}}^* = l_{\text{gcd}}^*\) for all \(s \in \mathcal{E}\), let \(\lambda_2 = 1\).

**Sketch of Proof:** For a complete proof, see [13, Lemma 5.59]. For \(i \neq 1 \pmod{L_{\text{gcd}}}\), \(P(\rho_i) = 0\) follows from Remark 5. For \(i \equiv 1 \pmod{L_{\text{gcd}}}\), two main equations characterize \(P(\rho_i)\). Firstly, the event \(\rho_i\) means a symbol \(s\) finished with the bit \(u_{i-1}\), thus started with \(u_{i-1-l(s)}\), i.e., \(\rho_{i-1-l(s)}\). By summing over all \(s\),

\[
P(\rho_i) = \sum_{s \in \mathcal{E}} P(\rho_i|s) P(s) = \sum_{s \in \mathcal{E}} P(\rho_{i-1-l(s)}) P(s).
\]

Secondly, the opposite event of \(\rho_i\), of probability \(1 - P(\rho_i)\), means a symbol \(s\) started in any position \(i - j \leq l(s)\) and \(i - j - 1 \equiv 1 \pmod{L_{\text{gcd}}} \). Therefore,

\[
1 - P(\rho_i) = \sum_{s \in \mathcal{E}} \left(\frac{l(s)/L_{\text{gcd}}-1}{\sum_{k=1}^{l_{\text{max}}^*} P(\rho_{i-1-l_{\text{gcd}}^*})}\right) P(s).
\]

By assuming \(P(\rho_i) \to P\) for some \(P\) as \(i\) becomes large, we deduce \(P = l_{\text{gcd}}^*/l_{\text{gcd}}\) from (29). The convergence and its rate can be proved by analyzing the recurrence (28).

B. Proof of Th. 14

Here is a short introduction to the lemmas 21–23 used in the proof of Th. 14. The spectrum (10) can be written equivalently as

\[
A_{s_l,h}^{F_b} = \sum_{n_b \in U_{1:n_b}} \sum_{u_{1:n_b}} P(B_{vle} = u_{1:n_b}|F_b) A_{s_l,h|u_{1:n_b}}^{F_b},
\]

where the second summation enumerates all possible \(u_{1:n_b}\) and the conditional spectrum \(A_{s_l,h|u_{1:n_b}}^{F_b}\) involves all possible \(u_{1:n_b}\) given \(u_{1:n_b}\). Instead of enumerating all pairs \((u, \bar{u})\) directly, let us arrange the joint enumeration of \(u\) and \(\bar{u}\) such that it is sorted by the number \(n\) of error events \((e_1, e_2, \ldots, e_n) = \bar{e}(\bar{u}, u)\) by the bit lengths \(l_{1:n}\) of these events, by their starting bit positions \(j_{1:n}\), by their Levenshtein symbol distances \(s_{1:1:n}\) and by their Hamming distances \(h_{1:n}\). Then, Lemma 21 expresses \(d_S(\bar{u}, u)\) approximately as a function of \(d_S(e_m)\). Lemma 22 provides results related to the probability of having a VLC block \(u_{1:n_b}\) with codewords starting at the given positions \(j_{1:n}\). At last, Lemma 23 states that the intermediate spectrum \(T_{h_{1:n},b}^{\text{vle}}\) is bounded, enabling us to neglect long combinations of error events.

**Lemma 21:** Given \(s\) and \(\hat{s}\) with \(l(s) = l(\hat{s})\), \(d_S(s, \hat{s}) \leq \sum_{e \in E(s, \hat{s})} d_S(e)\). This bound is rather tight in practice. It is in most cases an equality but still a strict inequality with some combinations of error events.

□
Proof: Follows from the definition of the Levenshtein symbol distance $d_{S_n}(x)$. ■

Lemma 22: Given the set $D_{n,N}$ and the variable $K_{n,N}$,

$$D_{n,N} \triangleq \left\{ (d_1, \ldots, d_{n+1}) \in \mathbb{N}_{\geq 0}^{n+1} : \sum_{m=1}^{n+1} d_m = N \right\}, \quad (31)$$

$$K_{n,N} \triangleq \frac{1}{|D_{n,N}|} \sum_{d \in D_{n,N}} \prod_{m=1}^{n+1} P(\rho_{1+d_m}^{L_m\text{gcd}}), \quad (32)$$

where $|D_{n,N}| = \binom{N+n}{n}$, we have for fixed $n$

$$K_{n,N} \leq 1, \quad (33)$$

$$K_{n,N} = \left( \frac{\ell_{\text{gcd}}}{\ell} \right)^{n+1} \pm \mathcal{O}\left( \frac{n^2 \log^2(N)}{N} \right), \quad (34)$$

The same result holds if we replace some $P(\rho_{1+d_m}^{L_m\text{gcd}})$ with $P(\rho_{1+d_m}^{L_m\text{gcd}}|F_k)$ in (32).

It is certainly possible to tighten (34) but this is beyond scope since we focus on large values of $N$.

Proof: Given the definition of $D_{n,N}$, its cardinality is clearly equal to the number of ways we can arrange $n$ red balls among $N$ black balls, which is $\binom{N+n}{n}$. And (33) follows trivially from $P(\rho_{1+d_m}^{L_m\text{gcd}}) \leq 1$.

To prove (34), let us define a subset of $D_{n,N}$ whose elements $d_{1:n+1}$ are larger than $\Delta = [\log^2(N)]$:

$$D_{n,N}^\Delta \triangleq \{ d_{1:n+1} \in D_{n,N} : d_m \geq \Delta \text{ for } 1 \leq m \leq n+1 \}. \quad (35)$$

Then,

$$K_{n,N} = \frac{1}{|D_{n,N}|} \left( \sum_{d \in D_{n,N}} \prod_{m=1}^{n+1} P(\rho_{1+d_m}^{L_m\text{gcd}}) + \sum_{d \in D_{n,N}\setminus D_{n,N}^\Delta} \prod_{m=1}^{n+1} P(\rho_{1+d_m}^{L_m\text{gcd}}) \right). \quad (36)$$

In the first summation, the elements $d_i$ are all larger than $\Delta$, thus $P(\rho_{1+d_m}^{L_m\text{gcd}}) = \ell_{\text{gcd}}^{L_m}/\ell \pm \mathcal{O}(\ell_2^{\Delta})$ by Lemma 20. In the second summation, let us upper bound all factors $P(\rho_{1+d_m}^{L_m\text{gcd}})$ by 1.

$$K_{n,N} = \frac{|D_{n,N}^\Delta|}{|D_{n,N}|} \left( \frac{\ell_{\text{gcd}}}{\ell} \pm \mathcal{O}(\ell_2^{\Delta}) \right)^{n+1} + \mathcal{O}(1 - |D_{n,N}|/|D_{n,N}|). \quad (37)$$

Straighforwardly, $|D_{n,N}^\Delta| = \binom{N-(n+1)\Delta+n}{n}$. Besides, for fixed $n$ and $\Delta = [\log^2(N)]$,

$$1 \geq \frac{|D_{n,N}|}{|D_{n,N}^\Delta|} = \frac{N!}{(N-(n+1)\Delta+n)!} \quad \text{and} \quad \frac{(N-(n+1)\Delta+n)!}{(N-n)!} \geq \frac{(N+n)!}{(N-(n+1)\Delta)!} \geq \frac{(N+n)!}{(N-n)!} \geq \frac{(N+n)!}{(N+n)!} = \left(1 - \frac{n+(n+1)\Delta}{N+n} \right)^n \geq \mathcal{O}\left( \frac{n^2\Delta}{N+n} \right).$$

Substituting this into (37) proves (34). The proof holds also if some of the factors $P(\rho_{1+d_m}^{L_m\text{gcd}}|F_k)$ in (32) are replaced by $P(\rho_{1+d_m}^{L_m\text{gcd}}|F_k)$, by extension of Lemma 20 with 9). ■

Lemma 23: If the VLC spectrum $A_{L_m}^{\text{vlc}}$ is bounded (Def. 10), then the spectrum $T_{h_{1:m},h}^{\text{vlc}}$ of (34) is also bounded, i.e., $\exists e < 1, \forall h_m \geq \ell_{\text{gcd}}^{L_m}$ and $\sum_{m=1}^{n} A_{h_m}^{\text{vlc}} (h_{1:m},h) = O(e^\Delta)$. ■

Proof: Let us consider a fixed value of $h$. By definition,

$$T_{h_{1:m},h}^{\text{vlc}} = \sum_{n_h} \sum_{l_{1:m}} \sum_{l_m=1}^{\ell_{\text{gcd}}^{L_m}} \sum_{m=1}^{n} A_{h_m}^{\text{vlc}} (h_{1:m},h) \prod_{l_m=1}^{\ell_{\text{gcd}}^{L_m}}. \quad (38)$$

Since $A_{h_m}^{\text{vlc}} (h_{1:m},h) = \ell_{\text{gcd}}^{L_m}$ is bounded, $\exists e < 1, \forall h_m \geq \ell_{\text{gcd}}^{L_m}$ and $\sum_{m=1}^{n} A_{h_m}^{\text{vlc}} (h_{1:m},h) = O(e^\Delta)$. Then, for $h_m \leq h$, we have $A_{h_m}^{\text{vlc}} (h_{1:m},h) \leq K_h e^\Delta$ and $\prod_{l_m=1}^{\ell_{\text{gcd}}^{L_m}} \leq K_h e^\Delta$. The rest of the proof follows since $h = \sum_{m=1}^{n} h_m \geq n$ and the number of terms in (38) is polynomial in $h, l_h$.

Proof of Th. 14: If we restrict the enumeration in (30) to realizations $u_{1:n}$ of non-zero probability (the only ones that matter in (30)), the quantities $L_{1:m}, L_{1:n}, L_{1:n}, L_{1:n}$ are constrained as follows. For $L_{1:n}$, let $\mathcal{L}_{n,n}$ denote the set of possible lengths, $\mathcal{L}_{n,n} = \{ l_{1:n} : \forall m, \ l_{m} \geq \ell_{\text{gcd}}^{L_{m}} \}$. For $L_{1:n}$, let $\mathcal{J}_{1:n,n}$ be the set of possible positions. $\mathcal{J}_{1:n,n} = \{ l_{1:n} : \forall m, \ l_{m} \geq \ell_{\text{gcd}}^{L_{m}} \}$. For $L_{1:n}$, let $\mathcal{H}_{h} = \{ h_{1:n} : \sum_{m=1}^{n} h_m = h \}$ with $h = d_{\text{H}}(u_{1:n}, \bar{u}_{1:n})$. Unfortunately, we know from Lemma 21 that the link between $L_{1:n,n}$ and $L_{1:n,n} = d_{\text{L}}(u_{1:n}, \bar{u}_{1:n})$ is not unequivocal due to the inequality $\sum_{m=1}^{n} \sum_{l_m=1}^{\ell_{\text{gcd}}^{L_{m}}} \geq L_{1:n,n}$. Nevertheless, we make the approximation in the following that it is an equality because (i) this inequality is an equality in most cases; (ii) due to the sense of this inequality, replacing it with an equality will result eventually in a spectrum that counts more Levenshtein symbol errors than the correct spectrum $\sum_{m=1}^{n} \sum_{l_m=1}^{\ell_{\text{gcd}}^{L_{m}}}$ which is compatible with upper performance bounds. So we consider $S_{\text{L}} \triangleq \{ s_{1:n} \in L_{1:n,n} : \sum_{m=1}^{n} \sum_{l_m=1}^{\ell_{\text{gcd}}^{L_{m}}} \geq L_{1:n,n} \}$ in the following. By rearranging (30) as mentioned, the spectrum can be written approximately (the approximation is due to $s_{1:n} \in S_{\text{L}}$) as

$$A_{s_{1:n},h}^{\text{Fc}} \approx \sum_{n} \sum_{u_{1:n} \in L_{1:n,n}} \sum_{h_{1:n}} \sum_{L_{1:n,n}} \sum_{s_{1:n} \in S_{\text{L}}} P(u_{1:n}|s_{1:n},h_{1:n},L_{1:n,n}) \prod_{l_m=1}^{\ell_{\text{gcd}}^{L_{m}}} \quad (39)$$

where $A_{s_{1:n},h_{1:n},s_{1:n},h_{1:n},L_{1:n,n}}^{\text{Fc}} = \sum_{s_{1:n},h_{1:n},s_{1:n},h_{1:n},L_{1:n,n}} P(u_{1:n}|s_{1:n},h_{1:n},L_{1:n,n})$. $P(u_{1:n} | s_{1:n}, h_{1:n}, L_{1:n,n})$ counts the number of $u_{1:n}$ that form $n$ error events $e_1, e_2, \ldots, e_n$ with $u_{1:n}$ subject to the following constraints: $e_m$ starts with $u_{j_m}$, $l_m = \ell_{(e_m)}$, $h_m = d_{\text{H}}(e_m)$ and $s_{1:n} = d_{\text{L}}(e_m)$. Since $\bar{u}_{1:n}$ differs from $u_{1:n}$ only in these error events, the number of possible $\bar{u}_{1:n}$ is actually equal to the number of possible events $e_1, e_2, \ldots, e_n$ subject to the above constraints. But, subject to these constraints, these events are independent of each other and thus their number is equal to the number of $e_1$ times the number of $e_2$ times the number of $e_n$. At last, the number
of $c_m$ is precisely the number of sub-sequences $u_{y_{m}}$ that form an error event with $u_{y_{m}}$ subject to the above $s_{L,m}$ and $h_{m}$, with $u_{y_{m}} = u_{y_{m_{1}}:j_{m_{1}}+m_{1}-1}$, which is precisely given by the conditional spectrum $A_{s_{L,m},h_{m}} = A_{s_{L,m},h_{m}}|u_{y_{m}}$ from Def. 7. Consequently, by taking all these considerations into account, $A_{s_{L},h|u_{1:m},s_{L},h_{m}} = \prod_{m=1}^{n} A_{s_{L,m},h_{m}} = \prod_{m=1}^{n} |J_{m}|$. At last, note $p_j$ in (39) (see Def. 19) accounts for the fact that to have error events starting in positions $j$, the transmitted $u_{1:m}$ must necessarily have codewords starting at those positions. With $u_{y_{m}} = (u_{y_{m_{1}}}, u_{y_{m_{2}}}, \ldots, u_{y_{m_{n}}})$, by reorganizing the summations, expression (39) can be written as

$$A_{s_{L},h} = \sum_{n_b} \sum_{n \geq 1} \sum_{l \in \mathbb{L}} \sum_{j \in J_{l,m}} \sum_{h \in H_{l}} \sum_{s_{L} \in S_{L}} P(u_{y_{m}} | F_{b}) (\mathbb{P}(\rho_j, u_{y_{m}}, F_{b})) \times \left( \sum_{m=1}^{n} A_{s_{L,m},h} = A_{s_{L,m},h} | u_{y_{m}} \right) \times \sum_{u_{1:m} \in \mathbb{V}^{+}} P(B_{v_{c}} = u_{1:m} | \rho_j, n_b, u_{y_{m}}, F_{b}) .$$

We have immediately the following simplifications:

$$P(n_b | \rho_j, u_{y_{m}}, F_{b}) = P(n_b | \rho_j, u_{y_{m}, F_{b}}) = P(n_b | \rho_j, n_b, F_{b}),$$

$$P(n_b | \rho_j, u_{y_{m}, F_{b}}) = P(n_b | \rho_j, n_b, F_{b}) = P(n_b | n_b+1, F_{b}) = P(n_b+1 | \rho_j, n_b, F_{b}) ,$$

$$P(\rho_j, u_{y_{m}}, F_{b}) = P(\rho_j, F_{b}) \times P(u_{y_{m}} | \rho_j, F_{b})$$

$$\times \prod_{m=1}^{n_b} P(\rho_{j_{m_{1}}+m_{1}-1}, u_{y_{m_{1}}-1}, F_{b}) = P(u_{y_{m}}, \rho_j, F_{b}).$$

By (9), $P(\rho_j, u_{y_{m}}, F_{b}) = P(\rho_j, u_{y_{m}}, F_{b}) = P(\rho_{j_{m_{1}}+m_{1}-1}, F_{b})$, similarly for $P(n_b+1 | \rho_j, n_b, F_{b})$, and, when $j > 1$, $P(u_{y_{m}}, \rho_j, F_{b}) = P(u_{y_{m}}, \rho_j, F_{b})$, which enables the simplification

$$P(n_b | \rho_j, u_{y_{m}}, F_{b}) = P(n_b | \rho_j, u_{y_{m}} | F_{b}).$$

$$\sum_{u_{1:m}} P(u_{1:m} | m = 1) \sum_{u_{1:m}} P(u_{1:m} | m = 1) \times P(\rho_j | F_{b})$$

$$\times P(\rho_{j_{m_{1}}+m_{1}-1}, F_{b}) = P(\rho_{j_{m_{1}}+m_{1}-1}, F_{b}).$$

For $2 \leq m \leq n$, $j > 1$ by definition of $J_{l,n_b}$ and thus (41) holds. For $m = 1$, however, it holds only when $j > 1$. Nevertheless, we can neglect this because the number of $j \in J_{l,n_b}$ subject to $j > 1$ is negligible, as explained hereafter. Eventually,

$$A_{s_{L},h} = \prod_{m=1}^{n} A_{s_{L,m},h} = A_{s_{L,m},h} | u_{y_{m}} = \prod_{m=1}^{n} A_{s_{L,m},h} | u_{y_{m}} = \prod_{m=1}^{n} P(\rho_j | F_{b})$$

$$\times P(\rho_{j_{m_{1}}, F_{b}}) \sum_{n_b} \sum_{l \in \mathbb{L}} \sum_{j \in J_{l,n_b}} P(\rho_j | F_{b}) .$$

Let us simplify the summation over $j \in J_{l,n_b}$, by defining the new variables $d_{l,m+1}$ as

$$d_{l,m+1} = \left\{ \begin{array}{ll} \frac{(j_l - 1)}{l_{t_{gcd}}} & \text{if } m = 1, \\
\left( \frac{(j_l - 1 - m + l_{t_{gcd}})}{l_{t_{gcd}}} & \text{if } 2 \leq m \leq n, \end{array} \right) \right.$$}

Given the set $J_{l,n_b}$ of possible $j$, the set of possible $d_{l,m+1}$ is exactly $D_{n,(n_b-l_{sum})/(l_{t_{gcd}}}$ in Lemma 22, with $l_{sum} = \sum_{m=1}^{n} l_{m}$. Therefore, by invoking Lemma 22,

$$A_{s_{L},h} = \prod_{m=1}^{n} A_{s_{L,m},h} | u_{y_{m}} = \prod_{m=1}^{n} P(\rho_{j_{m_{1}}, F_{b}}) \sum_{n_b} \sum_{l \in \mathbb{L}} \sum_{j \in J_{l,n_b}} P(\rho_j | F_{b})$$

$$\times P(\rho_{j_{m_{1}}, F_{b}}) \sum_{n_b} \sum_{l \in \mathbb{L}} \sum_{j \in J_{l,n_b}} P(\rho_j | F_{b}) .$$

where $T_{s_{L},l_{sum},n_b}$ is defined in Th. 14.

To simplify this last expression further, we can capitalize on the largeness of $N$. For $N$ large, the possible values of $n_b$ are relatively close to $N_{b}^{max} \approx N$ by (8). Therefore, if $N$ is sufficiently large, $T_{s_{L},l_{sum},n_b}$ becomes negligible as $l_{sum} \rightarrow n_b$, since $T_{s_{L},l_{sum},n_b}$ decreases exponentially with $l_{sum}$ by Lemma 23. This enables two simplifications: (i) the summation over $l_{sum} \leq n_b$ can be replaced asymptotically by a summation over $l_{sum} \leq N_{b}^{max}$, (ii) the term $O(n^2 \log^2(n_b - l_{sum})/(n_b - l_{sum}))$, which becomes non-negligible when $l_{sum} \rightarrow n_b$ (but which is nevertheless always upper bounded by 1 by (33)) is largely compensated by the exponential decrease of $T_{s_{L},l_{sum},n_b}$ as $l_{sum} \rightarrow n_b$. Also, note that for $h$ small (assumption) $n_b$ is small by $H_{t_{gcd}}$ and thus $(n_b-l_{sum})/(l_{t_{gcd}}+n)$ is relatively close to $(N_{b}^{max}-l_{sum})/l_{t_{gcd}}$ for $N$, $n_b$ large. At last, recall that in order to legitimate (41) for $j_l = 1$, we used the fact that the number of $j \in J_{l,n_b}$ subject to $j_1 = 1$ is negligible. This is indeed the case. Let $J_{l,n_b}^{1,1} = \{ j \in J_{l,n_b} : j_l = 1 \}$. By extension of Lemma 22, $|J_{l,n_b}^{1,1}| = |D_{n-1,(n_b-l_{sum})/(l_{t_{gcd}}}$, and thus $|J_{l,n_b}^{1,1}|/|J_{l,n_b}| = O(n/(n_b-l_{sum}))$, which can be neglected for the same reasons as the term $O(n^2 \log^2(n_b - l_{sum})/(n_b - l_{sum}))$ in (45). By combining all these approximations,

$$A_{s_{L},h} = \prod_{m=1}^{n} P(\rho_{j_{m_{1}}, F_{b}}) \sum_{n_b} \sum_{l \in \mathbb{L}} \sum_{j \in J_{l,n_b}} P(\rho_j | F_{b}) .$$


\[ \left( \frac{n_{\max}}{l_{\gcd}} - \frac{l_{\gcd}}{n} + n \right) \left( \frac{t^*_{\gcd}}{t} \right)^{n+1} T_{\text{vcl}}^{\text{dec}} s_{l}, h_{l}, b_{l}=l_{\text{exam}}, n^* \]

In this last expression, only the factor \( P(n_b | \rho_{n_b+1}, F_b) \) depends on \( n_b \). It is the probability to have a symbol \( S \) that does not fit at the end of the current frame, i.e., whose length is strictly larger than \( N - n_b \). \( P(l(S) > N - n_b) \). Taking the constraint (8) into account,

\[ \sum_{n_b} P(n_b | \rho_{n_b+1}, F_b) = \sum_{l=0}^{l_{\max}/l_{\gcd} - 1} P(l(S)/l_{\gcd}^* > l) = \tilde{l}/l_{\gcd}^* . \]

This result, together with (46), concludes the proof. \( \square \)

**ACKNOWLEDGMENT**

We greatly thank the reviewers for the significant improvements they suggested for this paper.

**References**


