

# Genetic Algorithm Aided Design of Near-Capacity Irregular Variable Length Codes

R. G. Maunder and L. Hanzo

Corresponding author: lh@ecs.soton.ac.uk

**Abstract**—In this paper we demonstrate that our ability to match the EXtrinsic Information Transfer (EXIT) function of an Irregular Variable Length Code (IrVLC) to that of a serially-concatenated inner code depends on the availability of a suite of component Variable Length Error Correction (VLEC) codebooks having a wide variety of inverted EXIT function shapes. We also show that the inverted EXIT function shape of a VLEC codebook depends on its coding rate and Error Correction Capability (ECC). This motivates the design of a Genetic Algorithm (GA) that searches the large VLEC parameter space to find codebooks having specific coding rates, ECCs and, hence, EXIT function shapes. The employment of this GA therefore facilitates the design of component VLEC codebook suites without the manual trial-and-error that is required when employing the state-of-the-art Heuristic Algorithm (HA) used as our bench marker, which cannot design component codebooks having specific EXIT function shapes.

## I. INTRODUCTION

It was shown in [1] that the area beneath the inverted EXtrinsic Information Transfer (EXIT) function [2] of a Variable Length Error Correction (VLEC) codebook  $C$  [3] is approximately equal to its coding rate  $R(C)$ . This coding rate depends on both the probabilities of occurrence of the various source symbol values and on the length of the corresponding VLEC codewords. More specifically, the  $K$ -ary source symbols have values of  $k \in [1, K]$  that occur with the probabilities  $\{P(k)\}_{k=1}^K$ , resulting in the entropy of  $E = -\sum_{k=1}^K P(k) \cdot \log_2[P(k)]$ . During VLEC encoding, the source symbols are mapped to binary codewords  $\{c_k\}_{k=1}^K$ , having lengths of  $\{L(c_k)\}_{k=1}^K$ . In order to ensure that each valid VLEC codeword sequence may be uniquely decoded, a lower bound equal to the source symbol entropy  $E$  is imposed upon the average codeword length  $L(C) = \sum_{k=1}^K P(k) \cdot L(c_k)$ . The coding rate  $R(C) = E/L(C)$  quantifies the redundancy within the VLEC codewords, which may be exploited for providing an error correcting capability during their decoding.

The Error Correction Capability (ECC) of a VLEC codebook  $C$  is typically characterized by its free distance lower bound  $\bar{d}_{\text{free}}(C)$  [3],

$$\bar{d}_{\text{free}}(C) = \min(d_{b_{\min}}(C), d_{d_{\min}}(C) + d_{c_{\min}}(C)), \quad (1)$$

where  $d_{b_{\min}}(C)$  was defined in [3] as the minimum block-distance between any pair of equal-length codewords in the

VLEC codebook  $C$ , while  $d_{d_{\min}}(C)$  and  $d_{c_{\min}}(C)$  were defined as the minimum divergence- and convergence-distances between any pair of unequal-length codewords, respectively.

Note that the ECCs of two VLEC codebooks cannot be distinguished using their free distance lower bounds  $\bar{d}_{\text{free}}(C)$ , if they happen to have the same integer value. For this reason, the Integer-Valued Free Distance (IV-FD) lower bound  $\bar{d}_{\text{free}}(C)$  is unsuitable for employment as the Objective Function (OF) or fitness function of a Genetic Algorithm (GA) [4] designed to generate VLEC codebooks having particular error correcting capabilities. This observation motivates the introduction of a novel Real-Valued Free Distance Metric (RV-FDM)  $D(C)$  in this paper in order to characterize the ECC of a VLEC codebook  $C$ . Since this RV-FDM is defined within the real-valued domain, it facilitates the unambiguous comparison of diverse VLEC codebooks' ECCs, even if they happen to have the same IV-FD lower bound.

In this paper, we will demonstrate that the number of inflections in the inverted EXIT function of a particular VLEC codebook  $C$  depends on its RV-FDM  $D(C)$ . This complements the proof of [5], which shows that if a VLEC codebook  $C$  has an IV-FD lower bound of  $\bar{d}_{\text{free}}(C) \geq 2$ , then it will have an inverted EXIT function that reaches the  $[E(C), E(C)]$  point of the EXIT chart and it will therefore support iterative decoding convergence to an infinitesimally low probability of error. Here  $E(C)$  is defined as the entropy of the VLEC-encoded bits, which is typically assumed to be unity and is given by  $E(C) = -\sum_{b=0}^1 \frac{L_b(C)}{L(C)} \cdot \log_2 \frac{L_b(C)}{L(C)}$ , where  $L_b(C) = \sum_{k=1}^K P(k) \cdot L_b(c_k)$  and  $L_b(c_k)$  is the number of bits in the VLEC codeword  $c_k$  that assume a value of  $b \in \{0, 1\}$ .

A Heuristic Algorithm (HA) was proposed for the construction of VLEC codes in [6], which was later refined in [7], [8]. These algorithms attempt to maximize the coding rate  $R(C)$  of a VLEC codebook  $C$  satisfying certain minimum integer-valued block-, divergence- and convergence-distances of  $d_{b_{\min}}(C)$ ,  $d_{d_{\min}}(C)$  and  $d_{c_{\min}}(C)$ , respectively. However, the nature of these HAs does not facilitate the direct control or prediction of the resultant VLEC codebook's coding rate  $R(C)$ , RV-FDM  $D(C)$  and, hence, inverted EXIT function shape. *By contrast, the novel contribution of this paper is that we propose a new GA [4], which is capable of generating VLEC codebooks having specific coding rates and RV-FDMs, together with desirable bit entropies.*

The rest of this paper is organized as follows. In Section II

we detail our novel RV-FDM, which allows the comparison of the ECC of two VLEC codebooks having equal IV-FD lower bounds. Our novel GA used for designing VLEC codebooks having both specific coding rates and RV-FDMs, as well as desirable bit entropies, is detailed in Section III. In Section IV, this GA is employed for designing a suite of component VLEC codebooks for use as Irregular Variable Length Coding (IrVLC) component codebooks [9]. Section V characterizes the suitability of these component VLEC codebooks for EXIT chart matching [10] and compares them with codebooks generated using the HA of [8]. Finally, we offer our conclusions in Section VI.

## II. THE FREE DISTANCE METRIC

In this section we introduce our RV-FDM  $D(C)$ , which allows the comparison of the ECC of two VLEC codebooks having equal IV-FD lower bounds  $\bar{d}_{\text{free}}(C)$ .

As described in Section I, the IV-FD lower bound  $\bar{d}_{\text{free}}(C)$  of a VLEC codebook  $C$  depends on the constituent codeword pairs that are ‘similar’, having the minimal block-, divergence- and convergence-distances of  $d_{b_{\min}}(C)$ ,  $d_{d_{\min}}(C)$  and  $d_{c_{\min}}(C)$ , respectively. The remaining so-called ‘dissimilar’ VLEC codeword pairs do not contribute to the IV-FD lower bound  $\bar{d}_{\text{free}}(C)$ , since they have block-, divergence- and convergence-distances that are higher than  $d_{b_{\min}}(C)$ ,  $d_{d_{\min}}(C)$  and  $d_{c_{\min}}(C)$ , respectively. Hence, we assume that the ECC of a VLEC codebook may be adequately characterized by modifying the IV-FD lower bound  $\bar{d}_{\text{free}}(C)$  to yield the RV-FDM according to

$$D(C) = \bar{d}_{\text{free}}(C) + F(C), \quad (2)$$

where  $F(C) \in [0, 1)$  quantifies the average similarity of the various VLEC codeword pairs, as we shall detail below. Note that  $\lfloor D(C) \rfloor = \bar{d}_{\text{free}}(C)$ , since we have  $F(C) \in [0, 1)$ . Hence, a VLEC codebook will have an inverted EXIT function that reaches the top right hand corner of the EXIT chart and will therefore support iterative decoding convergence to an infinitesimally low probability of error, if it has a RV-FDM of  $D(C) \geq 2$ , since this implies that it has an IV-FD lower bound of  $\bar{d}_{\text{free}}(C) \geq 2$ .

The average similarity of the various VLEC codeword pairs of a particular  $K$ -entry codebook  $C = \{\mathbf{c}_k\}_{k=1}^K$  may be quantified by

$$F(C) = \frac{\sum_{k^1=1}^{K-1} \sum_{k^2=k^1+1}^K P(k^1) \cdot P(k^2) \cdot F(\mathbf{c}_{k^1}, \mathbf{c}_{k^2})}{\sum_{k^1=1}^{K-1} \sum_{k^2=k^1+1}^K P(k^1) \cdot P(k^2)}, \quad (3)$$

where  $F(\mathbf{c}_{k^1}, \mathbf{c}_{k^2})$  quantifies the similarity of the specific pair of VLEC codewords  $\mathbf{c}_{k^1}$  and  $\mathbf{c}_{k^2}$ , as will be detailed below. Note that the calculation of  $F(C)$  in (3) takes into account the probabilities of occurrence  $P(k^1)$  and  $P(k^2)$  of the specific VLEC codewords  $\mathbf{c}_{k^1}$  and  $\mathbf{c}_{k^2}$ , respectively. These probabilities  $P(k^1)$  and  $P(k^2)$  must be considered, because the ECC of the VLEC codebook is related to the occurrence probabilities of the similar codewords, which contribute to the IV-FD lower bound  $\bar{d}_{\text{free}}(C)$ . For example, a stronger

error correcting capability can be expected, if only pairs of infrequently occurring VLEC codewords are similar. Note that in order to ensure that  $F(C) \in [0, 1)$ , VLEC codewords are not paired with themselves and a normalization factor of  $\sum_{k^1=1}^{K-1} \sum_{k^2=k^1+1}^K P(k^1) \cdot P(k^2) < 1$  is employed in the calculation of  $F(C)$  in (3).

As described above, the similarity of a specific pair of VLEC codewords  $\mathbf{c}_{k^1}$  and  $\mathbf{c}_{k^2}$  depends on the degree to which they contribute to the IV-FD lower bound. This may be determined according to Table I, in which  $d_b(\mathbf{c}_{k^1}, \mathbf{c}_{k^2})$  is the block-distance between a pair of equal-length VLEC codewords  $\mathbf{c}_{k^1}$  and  $\mathbf{c}_{k^2}$ , whilst  $d_d(\mathbf{c}_{k^1}, \mathbf{c}_{k^2})$  and  $d_c(\mathbf{c}_{k^1}, \mathbf{c}_{k^2})$  are the divergence- and convergence-distances between a specific pair of unequal-length codewords  $\mathbf{c}_{k^1}$  and  $\mathbf{c}_{k^2}$ , respectively [3].

TABLE I

THE SIMILARITY OF A SPECIFIC PAIR OF VLEC CODEWORDS  $\mathbf{c}_{k^1}$  AND  $\mathbf{c}_{k^2}$  DEPENDING ON THE DEGREE TO WHICH THEY CONTRIBUTE TO THE IV-FD LOWER BOUND.

Condition	$F(\mathbf{c}_{k^1}, \mathbf{c}_{k^2})$
if $L(\mathbf{c}_{k^1}) = L(\mathbf{c}_{k^2})$ and $d_{b_{\min}}(C) \leq d_{d_{\min}}(C) + d_{c_{\min}}(C)$ and $d_b(\mathbf{c}_{k^1}, \mathbf{c}_{k^2}) = d_{b_{\min}}(C)$	0
if $L(\mathbf{c}_{k^1}) \neq L(\mathbf{c}_{k^2})$ and $d_{d_{\min}}(C) + d_{c_{\min}}(C) \leq d_{b_{\min}}(C)$ and $d_d(\mathbf{c}_{k^1}, \mathbf{c}_{k^2}) = d_{d_{\min}}(C)$ and $d_c(\mathbf{c}_{k^1}, \mathbf{c}_{k^2}) = d_{c_{\min}}(C)$	0
if $L(\mathbf{c}_{k^1}) \neq L(\mathbf{c}_{k^2})$ and $d_{d_{\min}}(C) + d_{c_{\min}}(C) \leq d_{b_{\min}}(C)$ and $(d_d(\mathbf{c}_{k^1}, \mathbf{c}_{k^2}) = d_{d_{\min}}(C) \text{ xor } d_c(\mathbf{c}_{k^1}, \mathbf{c}_{k^2}) = d_{c_{\min}}(C))$	0.5
otherwise	1

## III. OVERVIEW OF THE PROPOSED GENETIC ALGORITHM

In this section we assume a basic familiarity with GAs [4] and we introduce a novel GA that facilitates the design of VLEC codebooks. During its operation, the GA continually generates candidate VLEC codebooks by mutating either the initial mother codebook  $C^0$  or a previously obtained candidate that has desirable properties, continuing until no further improvements can be obtained. The benefits of a candidate VLEC codebook  $C$  depend on its coding rate  $R(C)$ , RV-FDM  $D(C)$  and bit entropy  $E(C)$ . More specifically, we associate a high merit with candidate VLEC codebooks having a coding rate  $R(C)$  and a RV-FDM  $D(C)$  that are close to, but do not exceed, the user-specified limits  $R^{\text{lim}}$  and  $D^{\text{lim}}$  respectively, which are described in the GA parameter list of Table II. Furthermore, candidate VLEC codebooks having a near-unity bit entropy  $E(C)$  are regarded as having a high merit.

These aspects of a candidate VLEC codebook  $C$  are quantified using a composite quality metric  $M(C)$ , which yields a high value when a desirable coding rate  $R(C)$ , RV-FDM  $D(C)$  and bit entropy  $E(C)$  is exhibited. Our quality metric is defined here as

$$M(C) = \alpha^D \beta^D \frac{D(C) - D^{\text{best}}}{D^{\text{best}}} + \alpha^R \beta^R \frac{R(C) - R^{\text{best}}}{R^{\text{best}}} + \alpha^E \frac{E(C) - E^{\text{best}}}{E^{\text{best}}}, \quad (4)$$

where  $\alpha^D$ ,  $\alpha^R$ ,  $\alpha^E$ ,  $\beta^D$  and  $\beta^R$ , allow the relative importance of optimizing the RV-FDM, coding rate and bit entropy to

TABLE II

PARAMETERS THAT CONTROL THE OPERATION OF THE PROPOSED GA.

$C^0$	is the mother VLEC codebook.
$\alpha^D \geq 0, \alpha^R \geq 0$ and $\alpha^E \geq 0$	specify the relative importance of optimizing the RV-FDM, coding rate and bit entropy in (4), respectively.
$\beta^D \in \{-1, +1\}$ and $\beta^R \in \{-1, +1\}$	specify in (4) whether it is desirable to design candidate VLEC codebooks that have RV-FDMs and coding rates respectively that are higher than or lower than those of the mother codebook. A value of +1 specifies that it is desirable to increase the RV-FDM or coding rate, whilst a value of -1 indicates that a reduction is desirable.
$D^{\text{lim}}$ and $R^{\text{lim}}$	specify the maximum desirable RV-FDM when $\beta^D = +1$ and the maximum desirable coding rate when $\beta^R = +1$ , respectively. By contrast, they specify the minimum desirable RV-FDM when $\beta^D = -1$ and the minimum desirable coding rate when $\beta^R = -1$ .

be user-specified, as described in the GA parameter list of Table II. The values of  $D^{\text{best}}$ ,  $R^{\text{best}}$  and  $E^{\text{best}}$  are obtained by taking into consideration the RV-FDMs, coding rates and bit entropies of all candidate VLEC codebooks that have been considered during the GA so far. More specifically,  $D^{\text{best}}$  and  $R^{\text{best}}$  are the RV-FDM and coding rate that are closest to the specified limits  $D^{\text{lim}}$  and  $R^{\text{lim}}$ , respectively, while  $E^{\text{best}}$  is the highest bit entropy. These values are employed to normalize the RV-FDM, coding rate and bit entropy of the candidate VLEC codebook  $C$  and are constantly updated during the operation of the proposed GA. It is necessary to employ these constantly updated values, since it is often difficult to predict the values of the best RV-FDM, coding rate and bit entropy that will be found during the operation of the proposed GA in advance.

During the mutation of a particular candidate VLEC codebook  $C$ , the proposed GA employs four different types of operation, as follows:

- A randomly selected bit value of either 0 or 1 is *inserted* into a randomly selected position in a randomly selected codeword of the candidate VLEC codebook  $C$ .
- The value of a randomly selected bit in a randomly selected codeword of the candidate codebook  $C$  is *toggled* from 0 to 1 or from 1 to 0, as appropriate.
- A randomly selected bit in a randomly selected codeword of the candidate codebook  $C$  is *removed*.
- Two randomly selected codewords in the candidate codebook  $C$  are *swapped*.

During the mutation of a particular candidate VLEC codebook  $C$ ,  $M$  number of operations are performed, where the type of each operation is randomly selected from the above list. The number  $M$  of GA operations performed on the candidate VLEC codebook  $C$  is randomly selected, where the probability that  $M$  takes a particular value  $m \geq 1$  is given as  $P(M = m) = 2^{-m}$ . In this way, invoking a high number of mutations of the candidate VLEC codebook  $C$  is possible, but infrequent, since this is less likely to result in improved quality metrics.

However, this measure allows the GA to generate diverse codebooks in the interest of exploring the entire search-space sufficiently densely.

#### IV. DESIGN OF COMPONENT VLEC CODEBOOK SUITES

In this section we demonstrate the operation of the HA of [8] and the GA of Section III to design suites of component VLEC codebooks for use in IrVLCs [9]. We consider the IrVLC coding of  $K = 16$ -ary source symbol values, having the probabilities of occurrence that result from the Lloyd-Max quantization [11] [12] of independent Gaussian distributed source samples. These occurrence probabilities  $\{P(k)\}_{k=1}^K = \{0.008, 0.024, 0.043, 0.060, 0.076, 0.089, 0.097, 0.102, 0.102, 0.097, 0.089, 0.076, 0.060, 0.043, 0.024, 0.008\}$  are given by integrating the Gaussian Probability Density Function (PDF) between each pair of adjacent quantization decision boundaries [11]. These probabilities correspond to  $\{-\log_2(P(k))\}_{k=1}^K = \{6.93, 5.35, 4.55, 4.05, 3.72, 3.49, 3.36, 3.29, 3.29, 3.36, 3.49, 3.72, 4.05, 4.55, 5.35, 6.93\}$  bits of information per symbol, motivating the application of IrVLCs and giving a source entropy of  $E = 3.77$  bits/symbol.

As described in [9], an inverted IrVLC EXIT function is obtained as a weighted average of the inverted EXIT functions of the corresponding component VLEC codebooks. Here, EXIT chart matching [10] may be employed to shape the inverted IrVLC EXIT function so that an open EXIT chart tunnel may be created at near-capacity  $E_b/N_0$  values, potentially facilitating iterative decoding convergence to an infinitesimally low probability of error. However, the  $E_b/N_0$  value for which an open EXIT chart tunnel can be created is limited unless the component VLEC codebook suite is designed to have a wide variety of inverted EXIT function shapes. Furthermore, a suite like this can only be designed using the HA of [8] if a manually assisted trial-and-error procedure is employed, since it does not provide direct control over the resultant inverted EXIT function shapes, as described in Section I. By contrast, no trial-and-error intervention is required, when employing the GA of Section III, since it *does* provide direct control over the resultant inverted EXIT function shapes and is therefore more suitable for designing suites of component VLEC codebooks for use in IrVLCs. We note however, that some trial-and-error may be required on the one-off occasion when a user first familiarizes him or herself with the sensitivity of the GA to its parameters of Table II.

In order to ensure a fair comparison, trial-and-error was avoided when the HA of [8] was employed to design a suite of component VLEC codebooks. Instead, 11 component VLEC codebooks  $\{C^n\}_{n=1}^{11}$  having IV-FD lower bounds given by  $\bar{d}_{\text{free}}(C^n) = n + 1 \forall n \in [1 \dots 11]$  were generated. This was achieved in accordance with (1) by specifying a target minimum block-distance of  $d_{b_{\min}}(C^n) = \bar{d}_{\text{free}}(C^n)$ , a target minimum divergence-distance of  $d_{d_{\min}}(C^n) = \lceil \bar{d}_{\text{free}}(C^n)/2 \rceil$  and a target minimum convergence-distance of  $d_{c_{\min}}(C^n) = \lfloor \bar{d}_{\text{free}}(C^n)/2 \rfloor$  [8]. This approach is not labour-intensive, but results in a suite of component VLEC codebooks having only

the *limited* variety of inverted EXIT function shapes shown in Figure 1.

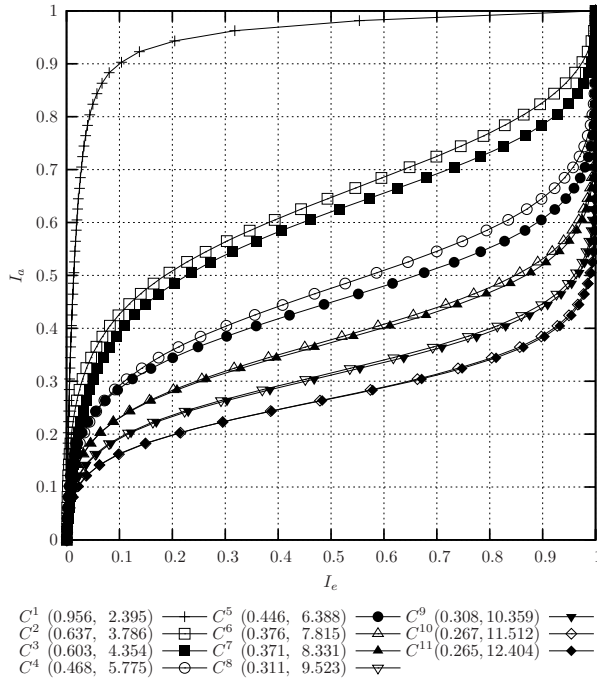


Fig. 1. Inverted EXIT functions of the IrVLC component codebook suite that was designed using the HA of [8]. Functions are labelled using the format  $C(R(C), D(C))$ .

Similarly, the GA of Section III was employed to generate a suite of 11 component VLEC codebooks  $\{C^n\}_{n=12}^{22}$  comprising two categories, namely  $\{C^n\}_{n=12}^{18}$  and  $\{C^n\}_{n=19}^{22}$ . During the generation of the first seven component VLEC codebooks  $\{C^n\}_{n=12}^{18}$ , the GA sought maximal RV-FDMs  $D(C)$  in order to provide the ‘S’-shaped inverted EXIT functions of Figure 2 that have up to two inflections. By contrast, during the generation of the remaining four component VLEC codebooks  $\{C^n\}_{n=19}^{22}$ , RV-FDMs  $D(C)$  as close to but no less than two were sought in order to provide the inverted EXIT functions of Figure 2 that contain no more than one inflection. Various coding rates  $R(C)$  were sought during the generation of the component VLEC codebooks  $\{C^n\}_{n=12}^{22}$  and, as a result the corresponding inverted EXIT functions of Figure 2 can be seen to exhibit more variety than those of Figure 1.

Note that the VLEC codebooks  $\{C^n\}_{n=1}^{11}$  designed using the HA of [8] were employed as mother codebooks during the design of the codebooks  $\{C^n\}_{n=12}^{22}$  using the GA of Section III. As a result, an assessment of the complexity associated with the proposed GA must include the complexity of performing the HA of [8] first. Hence, the complexity of the proposed GA is higher than that of the HA of [8], but was found to be reasonable, particularly since the GA is only invoked off-line, during the design of VLEC codebooks rather than during their use.

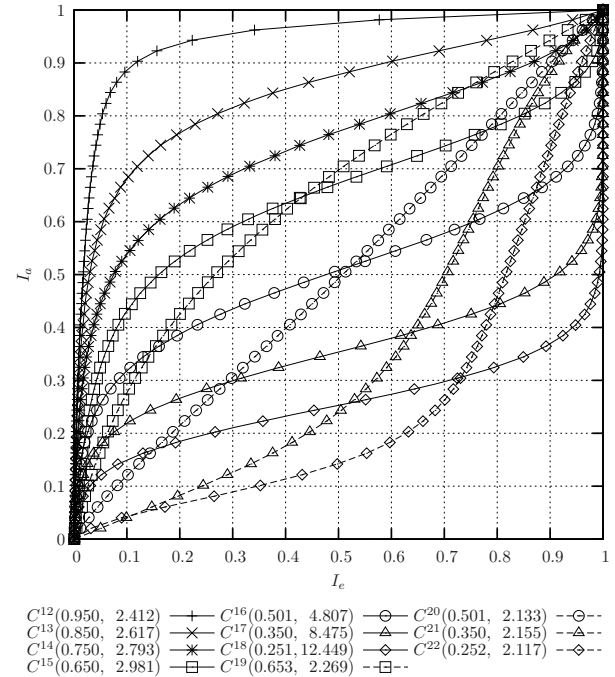


Fig. 2. Inverted EXIT functions of the IrVLC component codebook suite that was designed using the GA of Section III. Functions are labelled using the format  $C(R(C), D(C))$ .

## V. EXIT CHART MATCHING PERFORMANCE

Let us now consider the suitability of the suite of component VLEC codebooks  $\{C^n\}_{n=1}^{11}$  of Section IV that was designed using the HA of [8] and the suite of  $\{C^n\}_{n=12}^{22}$  designed using the GA of Section III for EXIT chart matching. We opted for matching the corresponding inverted IrVLC EXIT functions to the EXIT function of a memory-3 Unity Rate Code (URC) [13], which is employed to protect transmissions over a Binary Phase Shift Keying (BPSK) modulated uncorrelated Rayleigh fading channel having a range of  $E_b/N_0$  values. In each case, the EXIT chart matching algorithm of [10] was employed and the maximum IrVLC coding rate  $R$  for which an open EXIT chart tunnel could be created was recorded. These IrVLC coding rates  $R$ , together with the corresponding channel capacity [14], are plotted as a function of  $E_b/N_0$  in Figure 3.

The suite of component VLEC codebooks  $\{C^n\}_{n=12}^{22}$  of Section IV that was designed using the GA of Section III was found to be more suitable for use in EXIT chart matching than the set  $\{C^n\}_{n=1}^{11}$  designed using the HA of [8]. As shown in Figure 3, the suite  $\{C^n\}_{n=12}^{22}$  facilitates EXIT chart matching over a wider range of  $E_b/N_0$  values and facilitates the achievement of higher IrVLC coding rates  $R$ . Indeed, for an IrVLC coding rate of  $R = 0.55$ , the open EXIT chart tunnel of Figure 4 was created at 0.42 dB from the channel’s  $E_b/N_0$  capacity bound of 2.37 dB when employing the suite  $\{C^n\}_{n=12}^{22}$ , which compares favourably with the 1 dB discrepancy obtained, when employing the suite  $\{C^n\}_{n=1}^{11}$ . These findings may be explained by the higher degree of diversity

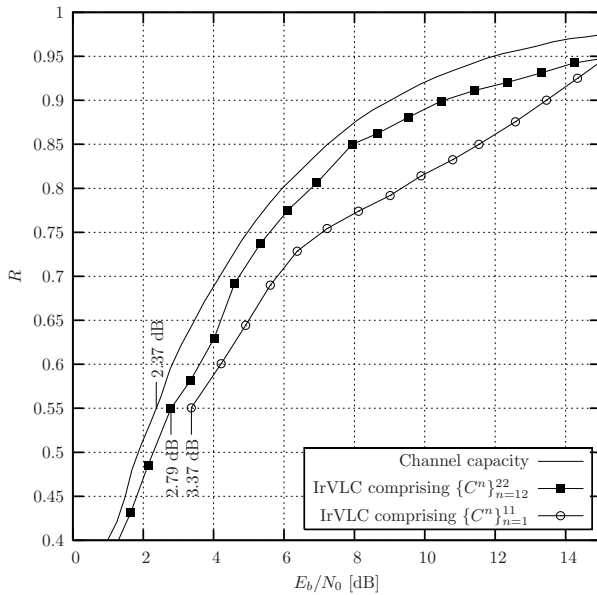


Fig. 3. The maximum coding rates for which open EXIT chart tunnels could be achieved for schemes employing a serial concatenation of IrVLCs based on the component VLEC codebook suites  $\{C^n\}_{n=1}^{11}$  and  $\{C^n\}_{n=12}^{22}$ , which are serially concatenated with a URC that protects transmission over a BPSK-modulated AWGN channel having a range of  $E_b/N_0$  values.

that is exhibited within the inverted component VLEC EXIT function shapes of Figure 2 than within those of Figure 1, as described in Section IV.

## VI. CONCLUSIONS

In this paper we have proposed the novel RV-FDM of (2), which allows the comparison of the ECC of two VLEC codebooks having equal IV-FD lower bounds. Furthermore, Figure 2 demonstrated that this RV-FDM affects the shape of the corresponding inverted VLEC EXIT function. This complements the property that the area below an inverted VLEC EXIT function equals the corresponding coding rate and the property that maintaining a RV-FDM of at least two guarantees an inverted EXIT function that reaches the top right hand corner of the EXIT chart. This motivates our novel GA-aided code design, generating VLEC codebooks having arbitrary coding rates, RV-FDMs and, hence, corresponding inverted EXIT function shapes. We demonstrated the ability of this novel GA to design a suite of component VLEC codebooks having a wide variety of inverted EXIT function shapes. This suite was compared with an equal-sized suite of codebooks that was designed using the HA of [8], which is incapable of designing component VLEC codebooks having arbitrary coding rates and RV-FDMs without imposing a significant degree of trial-and-error based manual intervention. Our GA was found to yield a suite of component VLEC codebooks that is better suited for IrVLC EXIT chart matching than that generated using the HA without employing trial-and-error.

## REFERENCES

[1] A. Ashikhmin, G. Kramer, and S. ten Brink, "Code rate and the area under extrinsic information transfer curves," in *Proceedings of the IEEE*

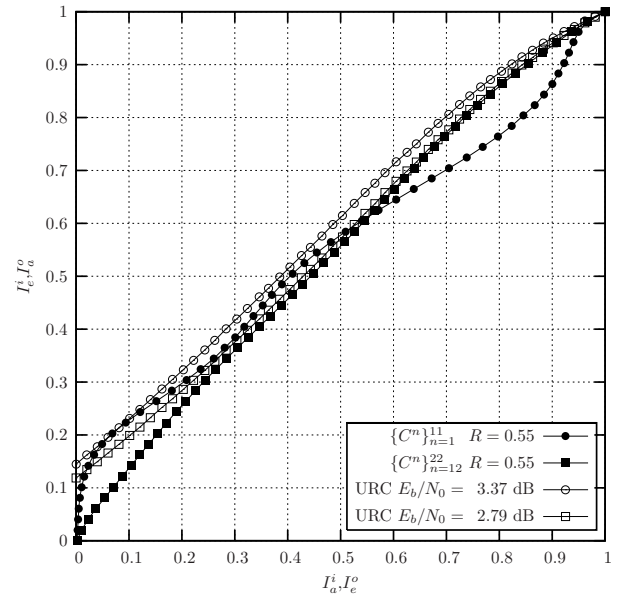


Fig. 4. Threshold EXIT chart tunnels obtained for schemes employing a serial concatenation of  $R = 0.55$ -rate IrVLCs based on the component VLEC codebook suites  $\{C^n\}_{n=1}^{11}$  and  $\{C^n\}_{n=12}^{22}$ , which are serially concatenated with a URC that protects transmission over a BPSK-modulated AWGN channel.

*International Symposium on Information Theory*, Lausanne, Switzerland, June 2002, p. 115.

- [2] S. ten Brink, "Convergence of iterative decoding," *Electronics Letters*, vol. 35, no. 10, pp. 806–808, May 1999.
- [3] V. Buttigieg and P. G. Farrell, "Variable-length error-correcting codes," *IEE Proceedings on Communications*, vol. 147, no. 4, pp. 211–215, August 2000.
- [4] D. E. Goldberg, *Genetic Algorithms in Search, Optimization and Machine Learning*. Addison-Wesley, 1989.
- [5] J. Klierer, N. Görtz, and A. Mertins, "Iterative source-channel decoding with Markov random field source models," *IEEE Transactions on Signal Processing*, vol. 54, no. 10, pp. 3688–3701, October 2006.
- [6] V. Buttigieg, "Variable-Length Error-Correcting Codes," Ph.D. dissertation, Department of Electronic Engineering, University of Manchester, Manchester, U.K., 1995.
- [7] C. Lamy and J. Paccout, "Optimised constructions for variable-length error correcting codes," in *Proceedings of the IEEE Information Theory Workshop*, Paris, France, March 2003, pp. 183–186.
- [8] J. Wang, L.-L. Yang, and L. Hanzo, "Iterative construction of reversible variable-length codes and variable-length error-correcting codes," *IEEE Communications Letters*, vol. 8, no. 11, pp. 671–673, November 2004.
- [9] R. G. Maunder, J. Wang, S. X. Ng, L.-L. Yang, and L. Hanzo, "On the performance and complexity of irregular variable length codes for near-capacity joint source and channel coding," to appear in *IEEE Transactions on Wireless Communications*.
- [10] M. Tüchler and J. Hagenauer, "EXIT charts of irregular codes," in *Proceedings of the Conference on Information Sciences and Systems*, Princeton, NJ, USA, March 2002, pp. 748–753.
- [11] S. Lloyd, "Least squares quantization in PCM," *IEEE Transactions on Information Theory*, vol. 28, no. 2, pp. 129–137, March 1982.
- [12] J. Max, "Quantizing for minimum distortion," *IRE Transactions on Information Theory*, vol. 6, no. 1, pp. 7–12, March 1960.
- [13] D. Divsalar, S. Dolinar, and F. Pollara, "Serial concatenated trellis coded modulation with rate-1 inner code," in *Proceedings of the IEEE Global Telecommunications Conference*, vol. 2, San Francisco, CA, USA, November 2000, pp. 777–782.
- [14] L. Hanzo, S. X. Ng, T. Keller, and W. Webb, *Quadrature Amplitude Modulation*. Chichester, UK: Wiley, 2004.