ON THE LOW-FREQUENCY SUPPRESSION PERFORMANCE OF DC-FREE RUNLENGTH-LIMITED MODULATION CODES

Volker Braun[1] and A.J.E.M. Janssen[2]
[1]Institute for Experimental Mathematics
Ellernstr. 29, 45326 Essen, Germany
[2]Philips Research Laboratories
Prof. Holstlaan 4, 5656 AAA Eindhoven
The Netherlands

Abstract— Basic trade-offs between the rate of combined DC-free runlength-limited (DCRLL) modulation codes and the amount of suppression of low-frequency components are presented. The main results are obtained by means of a numerical study of dependencies between statistical properties of ideal, 'maxentropic' DCRLL sequences. The numerical results are mathematically founded by proving the observed behavior of the Shannon capacity of the DCRLL constraints for asymptotically large values of digital sum variation. Presented characteristics of maxentropic DCRLL sequences comply with the corresponding properties of maxentropic pure DC-free sequences, as previously considered by Justesen and Immink. Knowledge of the maxentropic bounds enables us to evaluate the performances of implemented DCRLL codes with respect to their low-frequency suppression capability. Among the considered codes are the EFM code as applied in the Compact Disc system, and the EFMPlus code which has been adopted as the coding format of the Digital Versatile Disc system.

I. INTRODUCTION

IN THE Compact Disc system (CD), a combined DC-free runlength-limited (DCRLL) modulation code is used to transform the digital user (audio) bit stream into a sequence of binary channel symbols which is suitable for both storage on the disc and retrieval from the disc. The *runlength* is known as the number of consecutive like symbols occurring in a sequence. Runlength-limited sequences are characterized by two parameters, $T_{min} = (d + 1)$ and $T_{max} = (k + 1)$, denoting the minimum and the maximum runlengths occurring in the sequence. The parameter d controls the highest transition frequency while the parameter k ensures adequate frequency of transitions for synchronization of the read clock [2]. Suppression of the low-frequency content of the runlength-limited modulation sequence is employed in the CD system primarily to circumvent or reduce interaction between the data written on the disc and the servo systems that follow the track. According to Immink [8], the servo system is the Achilles' heel of the player as error correction is totally useless if track or clock loss occurs. Therefore, efficient suppression of the low-frequency components is the crucial criterion for designing DCRLL codes. The main goal is to achieve sufficient low-frequency suppression capability at minimum cost in *code rate*, i.e., the ratio between user bit rate and channel bit rate. In this paper, we determine the basic trade-offs between the rate of combined

Contributed Paper

DC-free runlength-limited codes and the amount of suppression of low-frequency components. Knowledge of these trade-offs enables us to evaluate the performances of several selected implemented DCRLL codes with respect to their low-frequency suppression capability. Among the considered codes are the EFM code as used in the CD system, and the EFMPlus code which, supported by this research study, has been adopted as the coding format of the Digital Versatile Disc system (DVD). As indicated above, DCRLL codes are very suitable for application in optical storage systems. The wide use of optical storage systems in applications like audio, multimedia or software distribution consequently gives an impression of the great engineering relevance of combined DC-free runlength-limited codes. However, DCRLL codes are certainly not confined to optical recording practice.

The following sections relate to the study of ideal, 'maxentropic' DCRLL sequences. In the next section, we will give an intuitive description of maxentropic DCRLL sequences, and will indicate how these sequences are modeled. A detailed description of the mathematical background relevant for this research study is far beyond the scope of this paper. As sources for further reading, providing insight into both the underlying mathematical background and the relation of this study to previous work, we recommend the book by Immink [1] and the article by Kerpez et al. [4]. An overview of modulation and coding for recording systems has been presented by, for example, Siegel and Wolf [3].

II. PRELIMINARIES

We assume that binary user information with a bit rate of $f_b = 1/T_b$ is translated into a coded channel sequence having the channel bit rate $1/T_c$, where $R = T_c/T_b$ denotes the rate of the code. We define the *Running Digital Sum* (RDS) of the encoded sequence $\{x_i\} = \{\ldots, x_{-1}, x_0, \ldots, x_i, \ldots\}$, $x_i \in \{-1, 1\}$, as

$$z_i = \sum_{j=-\infty}^{i} x_j = z_{i-1} + x_i.$$

Pierobon [6] showed that a sequence is DC-free if, and only if, the RDS assumes a finite number of values. This number is called the *Digital Sum Variation* (DSV), denoted by N. DCRLL sequences combine restrictions, or 'constraints', imposed on both the runlengths and the RDS. So, we characterize DCRLL sequences by three integer parameters (d, k, N), where by definition $0 \leq d < k \leq N - 2$. In

the absence of a specific k constraint, we treat the DCRLL constraints in the form (d, ∞, N), although a finite DSV implies a finite k constraint. By using this interpretation, we achieve that all our results obtained for ideal DCRLL sequences comply with the prior known relations of ideal pure DC-free sequences. A mathematical justification of this interpretation is to appear elsewhere [14]. Pure DC-free sequences, often called 'charge-constrained sequences', thus satisfy the $(0, \infty, N)$ constraint, and runlength limited (RLL) or (d, k) constrained sequences satisfy the (d, k, ∞) constraint.

In the following discussion, we will focus on ideal, 'maxentropic' (d, k, N) sequences. The reader should realize that a certain rate has to be sacrificed to convert arbitrary ('uncoded') data into a (d, k, N) constrained sequence. *Maxentropic* DCRLL sequences have the maximum code rate achievable to satisfy the (d, k, N) constraints. This maximum code rate is called the *Shannon capacity*. From an information-theoretic point of view, maxentropic DCRLL sequences carry the maximum amount of information possible under the (d, k, N) constraints. *Entropy* as a measure of the average information content present per symbol in a stochastic sequence was introduced by Shannon. Shannon further proved that the maximum entropy equals the capacity. In this paper, we do not explicitly consider the entropy of a DCRLL sequence, but it should be borne in mind that it is the relevant parameter for obtaining ideal, maxentropic DCRLL sequences. We are particularly interested in maxentropic DCRLL sequences, since they provide an upper bound in low-frequency suppression capability, given a certain (d, k) constraint and rate loss. The rate loss versus uncoded sequences is often called the *redundancy*. In the case of DCRLL sequences, we consider suppression of the low-frequency components of a sequence which is not uncoded but runlength-limited. So, it is reasonable to focus on the rate loss versus the corresponding maxentropic (d, k) constrained sequence. To this end, let the *extra redundancy* be defined as

$$\rho(d, k, N) = C(d, k, \infty) - C(d, k, N) \quad (> 0), \quad (1)$$

where $C(d, k, N)$ is the Shannon capacity of the (d, k, N) constraint. A method for computing the Shannon capacity $C(d, k, N)$ has been presented by Norris and Bloomberg [7]. This method involves the determination of the largest real eigenvalue of a certain nonnegative (real) square matrix having a size N. The extra redundancy as defined in (1) refers to maxentropic DCRLL sequences. If a DC-free code satisfying a certain (d, k) constraint is considered, $C(d, k, N)$ in (1) has to be replaced by the rate of the code.

As indicated by Kerpez et al. [4] and Justesen [5], analytical methods are not suitable for investigating dependencies between statistical properties of maxentropic DCRLL sequences. We will therefore perform a numerical analysis of these dependencies for a wide range of (d, k, N) constraints. In doing so, we will confine ourselves to practically interesting d constraints, i.e., $0 \leq d \leq 2$ [1]. A computationally efficient method for representing the (d, k, N) constraints involving *runlength graphs* has been described by Kerpez

et al. [4]. The runlength graph description of a specific (d, k, N) constraint can be given in the form of a square matrix with a size smaller than N. On the basis of a runlength graph, we will construct a stationary *Markov chain*, i.e., a specific kind of stochastic process including statistical dependencies between the random variables. Such a Markov chain forms the mathematical model of a DCRLL sequence. The fact that a DCRLL sequence can be modeled by a Markov chain is essential since it allows the evaluation of the entropy of the sequence. A maxentropic DCRLL sequence is obtained by carefully defining the Markov chain probabilities, represented by the entries of the underlying square matrix. This maxentropic Markov chain model is briefly described in Appendix A. A detailed description of this model can be found in [4]. The Markov chain description of maxentropic DCRLL sequences allows the computation of all relevant statistical properties of these sequences. Of these properties we will consider in particular the power spectral density function or power spectrum. We will denote the power spectrum of a maxentropic (d, k, N) constrained sequence by $H(\omega)$, where $\omega = 2\pi f T_c$. We will use the method described by Kerpez et al. [4] to evaluate the power spectrum $H(\omega)$. For each point of the frequency axis, this method mainly involves inversion of two complex valued square matrices of sizes $N - d - 1$. Another statistical property of maxentropic (d, k, N) constrained sequences to be considered is the sum variance, denoted by $\sigma_z^2(d, k, N)$, i.e., the variance of the RDS. The reason for investigating the sum variance $\sigma_z^2(d, k, N)$ stems from the fact that for maxentropic pure charge-constrained sequences it is a criterion of the low-frequency characteristic [5]. The computation of the sum variance is briefly described in Appendix A.

The outline of the next sections is as follows. In Section III, we will consider the trade-off between low-frequency suppression and rate loss for maxentropic DCRLL sequences. In order to quantify the low-frequency characteristic of these sequences, we will introduce an appropriate 'cut-off frequency'. For a wide range of (d, k, N) constraints, we will then numerically investigate the trade-off between cut-off frequency and extra redundancy. This trade-off turns out to be linear in good approximation. In Section IV, we will numerically investigate for a wide range of (d, k, N) constraints whether the sum variance of maxentropic DCRLL sequences can be used as a criterion of their low-frequency characteristic. For large DSVs, our numerical results reveal simple expressions for the sum variance of maxentropic DCRLL sequences and for the relation between sum variance and cut-off frequency. In Section V, we thus obtain the asymptotic behavior of the extra redundancy, which we will prove in Appendix B for the case that the k constraint remains finite as $N \to \infty$, i.e., for practically interesting k constraints [1]. All the relations of maxentropic DCRLL sequences presented in Sections III - V have corresponding relations in the theory of maxentropic pure charge-constrained sequences, as previously considered by Justesen [5] and Immink [1]. In Section VI, we will evaluate the performances of several implemented DCRLL

codes with respect to their low-frequency suppression capabilities. In particular, we will consider the EFM code [1] as applied in the CD system, EFMPlus [8] as applied in the DVD system, and the Miller Squared code [9]. In these performance evaluations, we will compare code properties obtained in computer simulations with the corresponding maxentropic performance bounds.

III. Trade-off between Redundancy and Low-frequency Suppression of Maxentropic DCRLL Sequences

We will start by considering a few examples of power spectra of maxentropic DCRLL sequences. As can be seen in Fig. 1, the spectral density of these sequences vanishes at zero frequency, and there is a region of frequencies, close to the zero frequency, where the spectral density is low. The width of this spectral notch can be quantified by a parameter called the *cut-off frequency*. Motivated by Fig. 1, we define the cut-off frequency of a maxentropic (d, k, N) constrained sequence, denoted by ω_0, by

$$H(\omega_0) = \frac{H_0(d, k)}{2}, \qquad (2)$$

where $H_0(d, k)$ denotes the *DC content* $H(\omega \to 0)$ of the corresponding maxentropic (d, k) constrained sequence. An expression for computing $H_0(d, k)$ has been presented by Immink [1], and an example of a cut-off frequency ω_0 is shown in Fig. 1.
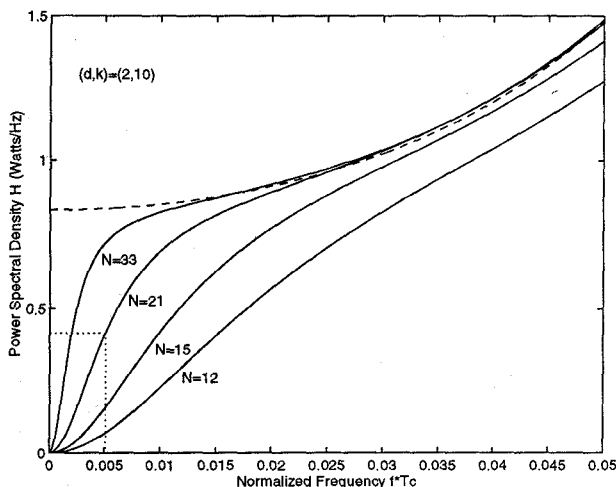

Figure 1: Examples of power spectra of maxentropic $(2, 10, N)$ constrained sequences, together with the power spectrum of a maxentropic $(2, 10)$ constrained sequence (dashed). By way of example, the cut-off frequency ω_0 is shown for a maxentropic $(2, 10, 21)$ constrained sequence (dotted).

From the power spectrum $H(\omega)$, we extract the cut-off frequency ω_0 for a wide range of (d, k, N) constraints, and relate it to the extra redundancy. It turns out that for maxentropic (d, k, N) constrained sequences, there exists in good approximation a linear trade-off between cut-off

frequency and extra redundancy, given by

$$\omega_0 \simeq \rho(d, k, N) \frac{2 \ln 2}{\pi^2/6 - 1}, \qquad N \gg 1. \qquad (3)$$

Relation (3) is fundamental in the sense that the constant of proportionality between cut-off frequency and extra redundancy is independent of the (d, k, N) constraints. This constant of proportionality has been found in accordance with the theory of maxentropic pure charge-constrained sequences [1], and will be mathematically justified in Section V and Appendix B. As an example, relation (3) is illustrated in Fig. 2 for maxentropic $(2, \infty, N)$ constrained sequences. Approximation (3) appears accurate to within 5% for $N > 20$ and, as Fig. 2 indicates, it is fairly reliable also for moderate DSVs. In the next section, we will focus on the sum variance of maxentropic DCRLL sequences.
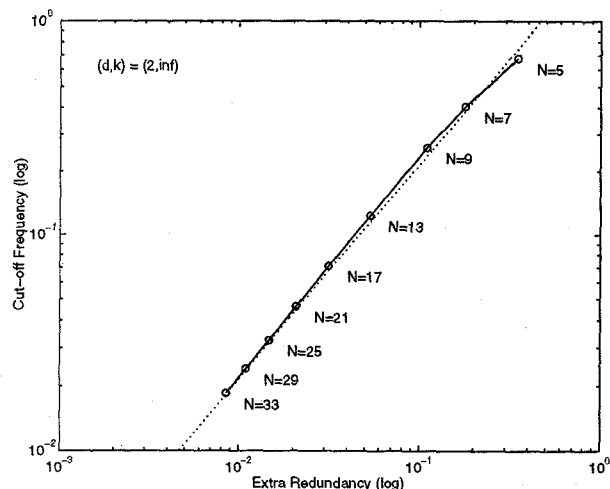

Figure 2: Cut-off frequencies of several maxentropic $(2, \infty, N)$ constrained sequences versus extra redundancy. The dotted line indicates equality in (3).

IV. The Sum Variance as a Criterion of the Low-frequency Characteristic of Maxentropic DCRLL Sequences

In the case of maxentropic pure charge-constrained sequences, the sum variance is related to the low-frequency characteristic [5]. In this section, we will investigate whether a corresponding relation also exists for maxentropic (d, k, N) constrained sequences. We will start by analyzing the influence of the (d, k, N) constraints on the sum variance of maxentropic DCRLL sequences. Subsequently, we will consider the relation between sum variance and cut-off frequency of these sequences. We will proceed by performing a numerical analysis of these properties for a wide range of (d, k, N) constraints.

A. The Influence of the Constraints on the Sum Variance

We will start by analyzing the sum variance of maxentropic DCRLL sequences in the absence of a specific k constraint (i.e. $k \to \infty$). It turns out that the influence of the d

constraint on the sum variance $\sigma_z^2(d, \infty, N)$ is negligible. Hence, we can approximate $\sigma_z^2(d, \infty, N)$ by the sum variance of maxentropic pure charge-constrained sequences, i.e., $\sigma_z^2(d, \infty, N) \simeq (1/12 - \pi^{-2}/2)(N+1)^2$ [1]. This approximation appears accurate to within 3% for $N > 9$. In a next step, we analyze the influence of the k constraint on the sum variance $\sigma_z^2(d, k, N)$. As an example, Fig. 3 shows
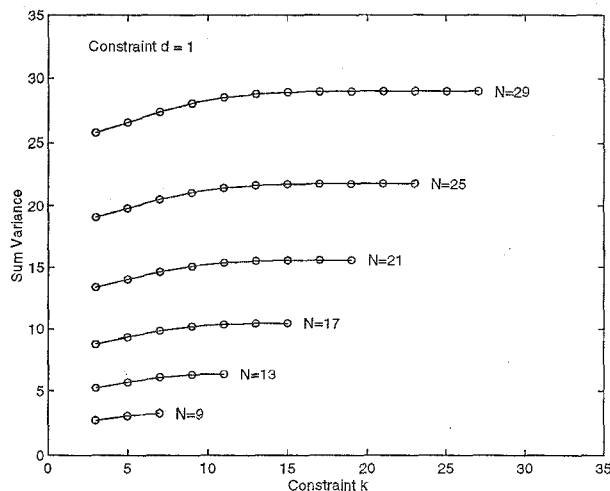

Figure 3: Sum variances of several maxentropic $(1, k, N)$ constrained sequences versus k. The DSV is denoted by N.

the sum variances of several maxentropic $(1, k, N)$ constrained sequences versus k. Apparently, the sum variance $\sigma_z^2(d, k, N)$ for fixed values of d and N decreases slightly with decreasing k. At a minimum k constraint, i.e., $k = d+1$, the sum variance can be relatively well approximated by $\sigma_z^2(d, d+1, N) \approx (1/12 - \pi^{-2}/2)(N-d)^2$. This approximation appears accurate to within 5% for $N > 21$. At fixed d and k constraints, we conclude that for asymptotically large values of DSV, irrespective of the d and k constraints, the sum variance of maxentropic DCRLL sequences is obtained as

$$\sigma_z^2(d, k, N) \simeq (\frac{1}{12} - \frac{1}{2\pi^2})N^2, \quad N \gg 1. \quad (4)$$

Below, we will consider whether the sum variance is related to the low-frequency characteristic for maxentropic DCRLL sequences.

B. The Relation between Sum Variance and Low-frequency Characteristic

We have investigated the relation between the sum variance $\sigma_z^2(d, k, N)$ and the cut-off frequency ω_0 for a wide range of (d, k, N) constraints. At fixed d and k constraints, we found that for large values of DSV there is a relation between sum variance and cut-off frequency of maxentropic DCRLL sequences according to

$$2\omega_0 \sigma_z^2(d, k, N) \approx H_0(d, k), \quad N \gg 1. \quad (5)$$

Approximation (5) appears accurate to within 10% for $N > 17$. For small values of DSV relation (5) is not reliable.

We conclude that for large values of DSV, the sum variance of maxentropic DCRLL sequences can serve as a criterion of their low-frequency characteristic. At fixed d and k constraints, we obtain from (3) and (5) that maxentropic DCRLL sequences for large values of DSV exhibit a constant sum variance-extra redundancy product, i.e.,

$$\rho(d, k, N)\sigma_z^2(d, k, N) \simeq \frac{\pi^2/6 - 1}{4\ln 2} H_0(d, k), \quad N \gg 1. \quad (6)$$

In the absence of a specific k constraint, relation (6) is illustrated in Fig. 4. Approximation (6) appears accurate to within 10% for $N > 25$.
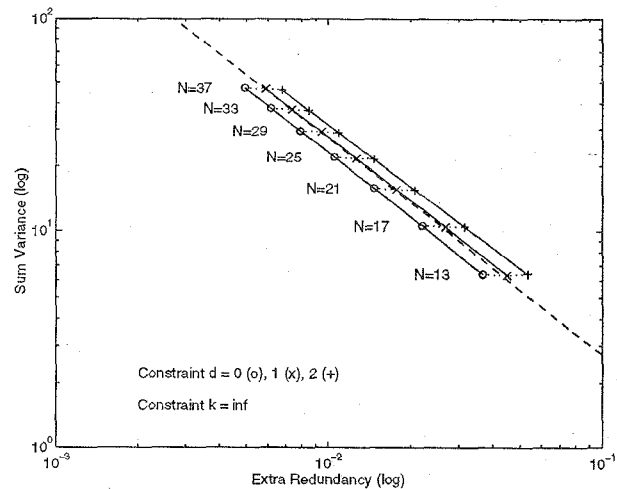

Figure 4: Sum variances of several maxentropic (d, ∞, N) constrained sequences versus extra redundancy, where $d = 0$ (o), 1 (x), and 2 (+). For $d = 1$, the dashed line indicates equality in (6). The dotted curves represent a constant value of DSV, denoted by N.

V. MATHEMATICAL FOUNDATION

All the properties of maxentropic DCRLL sequences presented in the preceding sections were obtained by numerically analyzing dependencies between statistical properties of these sequences. We will give a mathematical foundation for these properties by briefly showing the observed asymptotic behavior of the extra redundancy. From (4) and (6) we obtain that at fixed d and k constraints, the extra redundancy for asymptotically large values of DSV is given by

$$\rho(d, k, N) \simeq \frac{\pi^2}{2\ln 2} \frac{H_0(d, k)}{N^2}, \quad N \gg 1. \quad (7)$$

In Appendix B, we will briefly present a proof of this relation for the case that k remains finite as $N \to \infty$, i.e., for practically interesting k constraints [1]. We have proved relation (7) elsewhere [14] in the absence of a specific k constraint, but since this proof appears more involved than for bounded k constraints, it has been omitted here. Actually, this proof provides a mathematical indication for interpreting the DCRLL constraints in the form (d, ∞, N) in the absence of a specific k constraint.

VI. PERFORMANCES OF IMPLEMENTED CODES

In this section, we will assess the performances of several selected DCRLL codes with respect to their low-frequency suppression capabilities. In order to obtain a fair comparison between different DCRLL codes, both axes of the power spectra of these codes will be normalized for a fixed user bit rate, i.e., we will consider $H^*(fT_b) = RH(2\pi fT_c/R)$. As a performance criterion, we then determine the power spectral density of the coded sequence at a small fraction of the user bit rate, for example at $f_b/1000$. Maxentropic DCRLL sequences, where $R = C(d,k,N)$, provide an upper bound in low-frequency suppression capability, given a certain (d,k) constraint and redundancy. We will compare the low-frequency suppression capability of the coded sequence with the corresponding maxentropic performance bound. If there were no constraint imposed on the encoder/decoder complexity, the maxentropic bound could be reached arbitrarily closely. In practice however, a limited encoder/decoder complexity has an effect on the code rate and/or the low-frequency suppression capability of an implemented DCRLL code.

We will start by considering the EFM code [1] as applied in the CD player, EFMPlus [8] as applied in the DVD system, and two other EFM alternatives as described in [8]. These four codes all have finite values of DSV and satisfy the $(2,10)$ runlength constraint. Computer simulations were performed by Immink [8] to evaluate the power spectra of these codes. The low-frequency suppression performance for each of these codes is illustrated in Fig. 5. Apparently, all four codes allow further improvement in the suppression of the low frequency components. A strategy for improving the EFMPlus low-frequency suppression performance by about 3 dB is presented in [8].
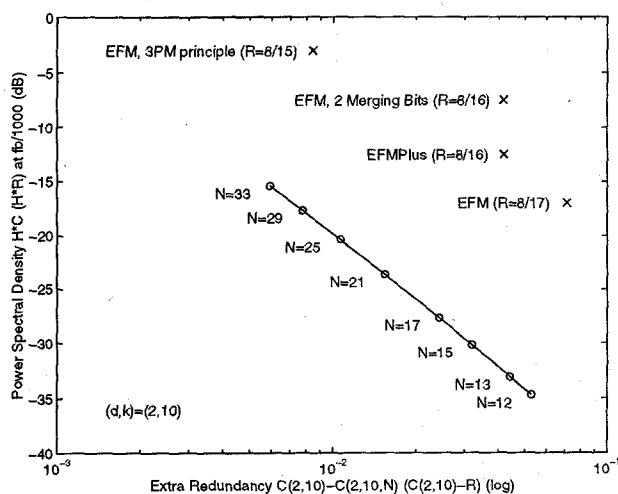
Figure 5:   Power spectral density $H^*(10^{-3})$ versus extra redundancy for maxentropic $(2,10,N)$ constrained sequences, for the EFM code, for EFMPlus, and for two other EFM alternatives.

As a further example, we will consider the Miller Squared code [9] satisfying the $(1,5,7)$ constraint. The power spectrum of the Miller Squared code has been determined in a computer simulation. As obtained from Fig. 6, the Miller Squared code has excellent performance with respect to low-frequency suppression. A current application of the Miller Squared code is in the D-2 digital video tape recorder [10].
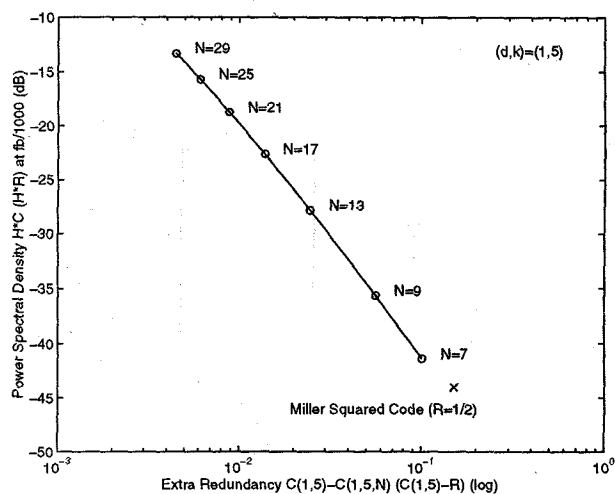
Figure 6: Power spectral density $H^*(10^{-3})$ versus extra redundancy for maxentropic $(1,5,N)$ constrained sequences and for the Miller Squared code.

VII. CONCLUSION

We have investigated dependencies between statistical properties of maxentropic DCRLL sequences. For these sequences, we introduced the extra redundancy and the cut-off frequency as appropriate measures for the rate loss and the low-frequency characteristic. Between cut-off frequency and extra redundancy, we found in good approximation a linear trade-off. For large values of DSV, we observed that the cut-off frequency is related to the variance of the RDS. All these properties of maxentropic DCRLL sequences were found by numerically analyzing dependencies between statistical properties of these sequences. We have mathematically founded our numerical results by proving the observed behavior of the Shannon capacity of the DCRLL constraints for asymptotically large values of DSV. By interpreting pure charge constraints in the form $(0,\infty,N)$, the theory of maxentropic pure charge-constrained sequences is obtained as a special instance of expressions (1)-(7) presented here. Knowledge of maxentropic DCRLL sequences enabled us to evaluate the performances of implemented DCRLL codes with respect to their low-frequency suppression capability. We showed that the Miller Squared code has excellent performance with respect to low-frequency suppression, whereas EFM and EFMPlus as applied in the CD and DVD systems, respectively, allow further improvements.

APPENDIX A
DERIVATION OF THE SUM VARIANCE

In this appendix, we will derive the sum variance $\sigma_z^2(d, k, N)$ of maxentropic (d, k, N) constrained sequences from the corresponding Markov chain description as presented by Kerpez et al. [4]. The (d, k, N) constraint is represented by a runlength graph. Let $\{s_m\}_{m \in I_M}$, $I_M = \{1, 2, \ldots, M\}$, denote the set of states of this graph, and recall that the number of states equals $M = N - d - 1$. To each state s_m is associated the value $U_m = d + m - (N-1)/2$, and we draw an edge of a length $l_{m,n} = U_m + U_n$ leading from state s_m to state s_n, if $d + 1 \leq l_{m,n} \leq k + 1$, where $m, n \in I_M$. Let $A_M = A_M(D)$ denote the adjacency matrix associated with this runlength graph, where $A_{m,n}$ denotes the mnth element of A_M. The nonzero elements of A_M are given by $A_{m,n} = D^{l_{m,n}}$, i.e, the adjacency matrix A_M can be expressed in the form

$$A_{m,n} = D^{d+1} \cdot \begin{cases} D^l, & \text{if } m + n = M + 1 + l, \\ 0, & \text{else,} \end{cases}$$

for $m, n \in I_M$, and for $l = 0, 1, \ldots, k - d$. Note that A_M is a Hankel matrix, i.e., A_M is constant on the antidiagonals. The Shannon capacity is given by $C(d, k, N) = \log_2 \lambda$, where λ denotes the largest real root of the characteristic polynomial $\det(A_M(\lambda^{-1}) - I)$, and I denotes the identity matrix. On the runlength graph there is based a stationary Markov chain which we describe by the state transition probability matrix P and the equilibrium distribution vector π. We define the (nonzero) elements of P by $P_{m,n} = v_n/v_m \lambda^{-l_{m,n}}$, where the vector v satisfies $A_M(\lambda^{-1})v = v$. The equilibrium distribution vector is defined by $\pi P = \pi$ together with $\sum_{m=1}^{M} \pi_m = 1$. The DCRLL sequences generated from the Markov chain thus defined are maxentropic [4].

As shown in [4], to each state s_m, with $m \in I_M$, there is associated the square of the RDS according to $(d + m - (N-1)/2)^2$. Further, recall that a binary symbol can only increase or decrease the RDS by one. So, to each edge of length l of the runlength graph, we can uniquely assign l values representing the squares of the RDS along this edge. In this way the sum variance of a maxentropic (d, k, N) constrained sequence can be evaluated according to

$$\sigma_z^2(d, k, N) = \frac{1}{L} \sum_{l=d+1}^{k+1} \sum_{m=l-d}^{M} \pi_m P_{m,(M+l-d-m)} z_{m,l}^2,$$

where L denotes the average runlength, and

$$z_{m,l}^2 = \sum_{i=0}^{l-1} (d + m - i - \frac{N-1}{2})^2.$$

APPENDIX B
ASYMPTOTIC BEHAVIOR OF THE EXTRA REDUNDANCY

In this appendix, we briefly prove the behavior of the extra redundancy for asymptotically large values of DSV. We consider relation (7) for small values of d in the case that k remains finite as $N \to \infty$. Recall that the (d, k, N) constraint is represented by the $M \times M$ Hankel type adjacency matrix $A_M = A_M(D)$ as defined in Appendix A, where $M = N - 1 - d$ [4]. To be able to determine the Shannon capacity $C(d, k, N) = -\log_2 D_{\min}$, we have to find the smallest $D = D_{\min} > 0$, such that the largest eigenvalue of $A_M(D)$ equals unity, i.e., $\lambda_{\max}(A_M) = 1$, as found in, for example, Proposition 4 of [11]. There is a vast amount of literature on the asymptotics of the eigenvalues of Toeplitz matrices, for example [12]. Further, we observe that the matrix A_M^2 almost has a Toeplitz structure. To be precise, the lower right $(M - p) \times (M - p)$ submatrix of A_M^2 is a symmetric Toeplitz matrix, whose elements on the jth diagonal are given by

$$c_j = D^{2d+2} \sum_{l=0}^{p-|j|} D^l D^{l+|j|},$$

where $c_j = c_{-j}$ $(j = 0, 1, \ldots, M - p - 1)$ and $p = k - d$. Note that $\lambda_{\max}(A_M^2) = \lambda_{\max}^2(A_M)$, since A_M is symmetric. Further, A_M^2 is irreducible, as found with the aid of, for example, Theorem 1.6. of [13]. We define a symmetric $M \times M$ Toeplitz matrix by $T_M = (c_{s-j})$, where $s, j = 0, 1, \ldots, M - 1$. As $A_M^2 \leq T_M$ entry by entry, we obtain

$$\lambda_{\max}(T_{M-p}) < \lambda_{\max}(A_M^2) < \lambda_{\max}(T_M),$$

as found with the aid of, for example, Lemma 2.4. (left hand side) and Theorem 2.1.3. (right hand side) of [13]. For $M \to \infty$, the maximum eigenvalue $\lambda_{\max}(T_{M-p}) \simeq \lambda_{\max}(T_M)$ is found with the aid of, for example, Theorem 2.1 of Widom [12]. From this theorem we obtain the maximum eigenvalue of A_M^2 as $M \to \infty$ according to

$$\lambda_{\max}(A_M^2) \simeq f(0) + \frac{\pi^2}{2} \frac{f^{(2)}(0)}{M^2},$$

where

$$f(\Theta) = \sum_{j=-\infty}^{\infty} c_j e^{ij\Theta} = D^{2d+2} |\sum_{l=0}^{p} D^l e^{il\Theta}|^2.$$

Using a Taylor expansion for $f(\Theta)$, we obtain

$$f^{(2)}(0) = 2D^{2d+2} \left((\sum_{l=0}^{p} l D^l)^2 - \sum_{l=0}^{p} l^2 D^l \cdot \sum_{l=0}^{p} D^l \right),$$

where $f^{(2)}(0) < 0$ (Cauchy-Schwarz inequality). Further, we introduce the definitions $\epsilon = \pi/M$, $\phi(D) = f(0) = D^{2d+2}(\sum_{l=0}^{p} D^l)^2$, and $\psi(D) = -f^{(2)}(0)/2$. Hence, we have to find the smallest root $D = D_\epsilon$ of

$$\lambda_{\max}(A_M^2) \simeq \phi(D) - \epsilon^2 \psi(D) = 1,$$

as $\epsilon \to 0$. In the case of $\epsilon = 0$, i.e., in the case of pure (d, k) constraints, we hence obtain $D_0 = D_{\epsilon=0}$ as the smallest root of the known characteristic equation $D^{d+1} \sum_{l=0}^{p} D^l = 1$, found in, for example, (5.15) of [1]. With a Taylor approximation, D_ϵ is approximately given by $D_\epsilon \simeq D_0 +$

$\epsilon^2 \psi(D_0)/\phi^{(1)}(D_0)$ as $\epsilon \to 0$. So, as $N \to \infty$, the extra redundancy results in

$$\rho(d, k, N) = \log_2 \frac{D_\epsilon}{D_0} \simeq \frac{\pi^2}{\ln 2} \frac{\psi(D_0)}{D_0 \phi^{(1)}(D_0)} \frac{1}{M^2}.$$

Using the characteristic equation $D_0^{d+1} \sum_{l=0}^{p} D_0^l = 1$, we find $\psi(D_0) = D_0^{d+1} \sum_{l=0}^{p} l^2 D_0^l - D_0^{2d+2}(\sum_{l=0}^{p} lD_0^l)^2$, and $D_0 \phi^{(1)}(D_0) = (2d + 2) + 2D_0^{d+1} \sum_{l=0}^{p} lD_0^l$. According to Immink [1], Chapter 5.4.3, the DC content of maxentropic (d, k) constrained sequences is given by

$$H_0(d, k) = \frac{\sum_{l=d+1}^{k+1} l^2 D_0^l - (\sum_{l=d+1}^{k+1} lD_0^l)^2}{\sum_{l=d+1}^{k+1} lD_0^l}.$$

Using the characteristic equation $D_0^{d+1} \sum_{l=0}^{p} D_0^l = 1$ and the substitution $l' = l - (d + 1)$, we obtain $\sum_{l=d+1}^{k+1} lD_0^l = (d + 1) + D_0^{d+1} \sum_{l'=0}^{p} l'D_0^{l'}$, and $\sum_{l=d+1}^{k+1} l^2 D_0^l = (d + 1)^2 + 2(d + 1)D_0^{d+1} \sum_{l'=0}^{p} l'D_0^{l'} + D_0^{d+1} \sum_{l'=0}^{p} l'^2 D_0^{l'}$. A comparison of the above expressions reveals that

$$\frac{\psi(D_0)}{D_0 \phi^{(1)}(D_0)} = \frac{H_0(d, k)}{2}.$$

Q.E.D.

REFERENCES

[1] K.A.S. Immink, *Coding Techniques for Digital Recorders*, Prentice Hall International (UK) Ltd, 1991.

[2] J. Watkinson, *The Art of Digital Audio*, Focal Press, London, 1988.

[3] P.H. Siegel, and J.K. Wolf, "Modulation and Coding for Information Storage," *IEEE Comm. Magazine*, vol. 29, no. 12, pp. 68-86, December 1991.

[4] K.J. Kerpez, A. Gallopoulos, and C. Heegard, "Maximum Entropy Charge-Constrained Run-Length Codes," *IEEE Journal on Selected Areas in Comm.*, vol. 10, no. 1, pp. 242-252, January 1992.

[5] J. Justesen, "Information Rates and Power Spectra of Digital Codes," *IEEE Trans. Inform. Theory*, vol. IT-28, no. 3, pp. 457-472, May 1982.

[6] G.L. Pierobon, "Codes for Zero Spectral Density at Zero Frequency," *IEEE Trans. Inform. Theory*, vol. IT-30, no. 2, pp. 435-439, March 1984.

[7] K. Norris, and D.S. Bloomberg, "Channel Capacity of Charge-Constrained Run-Length Limited Codes," *IEEE Trans. Magn.*, vol. MAG-17, no. 6, pp. 3542-3545, November 1981.

[8] K.A.S. Immink, "EFMPlus: The Coding Format of the Multimedia Compact Disc," *IEEE Trans. Consumer Electronics*, vol. CE-41, no. 3, pp. 491-497, August 1995.

[9] J.C. Mallinson, and J.W. Miller, "Optimal Codes for Digital Magnetic Recording," *Radio and Elec. Eng.*, vol. 47, pp. 172-176, 1977. See also J.W. Miller, U.S. Patent 4,027,335, May 1977.

[10] R. Brush, "Design Considerations for the D-2 PAL Composite DVTR," *Proc. Seventh Int. Conference on Video, Audio and Data Recording*, York, pp. 141-148, IERE Conference Publication no. 79, March 1988.

[11] R. Adler, J. Friedman, B. Kitchens, and B.H. Marcus, "State Splitting for Variable-Length Graphs," *IEEE Trans. Inform. Theory*, vol. IT-32, no. 1, pp. 108-113, January 1986.

[12] H. Widom, "On the Eigenvalues of Certain Hermitian Operators," *Trans. Amer. Math. Soc.*, vol. 88, pp. 491-522, July 1958.

[13] R.S. Varga, *Matrix Iterative Analysis*, Prentice-Hall, Inc., Englewood Cliffs, 1962.

[14] V. Braun, and A.J.E.M. Janssen, "manuscript to appear in the doctoral thesis of the first author," August 1995.
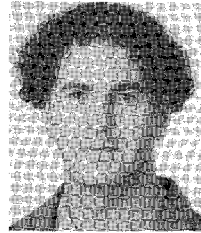
Volker Braun (S'93) was born in Neunkirchen, Germany, in 1967. He received the Dipl.-Ing. degree from the University of Kaiserslautern in 1992.

In 1992, he joined the Digital Communications Group at the Institute for Experimental Mathematics of the University of Essen. His research topics include modulation, coding and signal processing, especially for magnetic and optical recording systems.
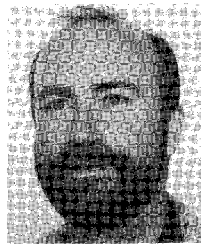
A.J.E.M. Janssen (SM'88) was born in Breda, The Netherlands, in 1953. He received the Eng. degree and Ph.D. degree in mathematics from the Eindhoven University of Technology, Eindhoven, The Netherlands, in 1976 and 1979, respectively.

From 1979 to 1981 he was a Bateman Research Instructor at the Mathematics Department of California Institute of Technology, and he joined the Philips Research Laboratories, Eindhoven, in 1981. His research interest is in mathematical analysis, in particular Fourier analysis and time-frequency analysis.