RZ 3677 (# 99687) 11/13/2006
Computer Science 35 pages

Research Report

Performance Evaluation of the Interleaved Parity-Check Intra-Disk Redundancy Scheme

Ilias Iliadis

IBM Research
Zurich Research Laboratory
8803 Rüschlikon
Switzerland

E-mail: ili@zurich.ibm.com

Performance Evaluation of the Interleaved Parity-Check Intra-Disk Redundancy Scheme

Ilias Iliadis
IBM Research
Zurich Research Laboratory
8803 Rüschlikon, Switzerland

Abstract

This report considers the interleaved parity-check intra-disk redundancy scheme proposed in [1], [2] to enhance the reliability of RAID systems. This scheme aims to protect the system against media-related unrecoverable errors. A detailed performance analysis of this as well as of traditional redundancy schemes based on Reed–Solomon codes and single-parity-check codes is conducted by analytical means. A new model is developed to capture the effect of correlated unrecoverable sector errors. The probability of an unrecoverable failure associated with these schemes is derived for the new correlated model as well as for the simpler independent error model. Furthermore, we derive closed-form expressions for the mean time to data loss of RAID 5 and RAID 6 systems in the presence of unrecoverable errors and disk failures. We then combine these results for a comprehensive characterization of the reliability of RAID systems that incorporate the proposed intradisk redundancy scheme. The impact on the mean time to data loss of a RAID 5 and RAID 6 systems is demonstrated.

I. Introduction

A current trend in the data storage industry is the increasing adoption of low-cost components, most notably SATA disk drives instead of FC and SCSI disk drives. SATA drives offer higher capacity per drive, but have a comparatively lower reliability. As the disk capacity grows, the total number of bytes that are read during a rebuild operation becomes very large. This increases the probability of encountering an unrecoverable error, i.e., an error that cannot be corrected by either the standard sector-associated error-control coding (ECC) or the re-read mechanism of the HDD. Unrecoverable media errors typically result in one or more sectors becoming unreadable. This is particularly problematic when combined with disk failures. For example, if a disk fails in a RAID 5 array, the rebuild process must read all the data on the remaining disks to rebuild the lost data on a spare disk. During this phase, a media error on any of the good disks would be unrecoverable and lead to data loss because there is no way to reconstruct the lost data sectors. A similar problem occurs when two disks fail in a RAID 6 scheme. In this case, any unrecoverable sectors encountered on the good disks during the rebuild process also lead to data loss.

A new XOR-based intra-disk redundancy scheme, called interleaved parity check (IPC), was proposed in [2] to enhance the reliability of RAID schemes. This scheme introduces an additional "dimension" of redundancy inside each disk that is orthogonal to the usual RAID dimension, which is based on redundancy across multiple disks. The RAID redundancy provides protection against disk failures, whereas the proposed intra-disk redundancy aims to protect against media-related unrecoverable failures. A key advantage of this new scheme, therefore, is that it can be applied to various RAID systems, including RAID 5 and RAID 6.

A new model capturing the effect of correlated unrecoverable sector errors was developed and subsequently used to analyze the proposed IPC scheme as well as traditional redundancy schemes based on Reed–Solomon (RS) codes and single-parity-check (SPC) codes. A first-order approximation of the probability of an unrecoverable failure associated with these schemes was derived for the new correlated model as well as for the simpler independent error model. The results showed that in the practical case of correlated sector errors where the maximum burst length exceeds the interleaving depth, the probability of an unrecoverable failure for the IPC scheme is roughly the same as for the optimum, albeit more complex, RS coding scheme. This, however, does not hold when the maximum burst length does not exceed the interleaving depth. In this report we derive the corresponding probabilities of an unrecoverable failure based on a second-order approximation.

Furthermore, suitable Markov models were developed in [2] to derive closed-form expressions for the mean time to data loss (MTTDL) of RAID 5 and RAID 6 systems in the presence of unrecoverable errors and disk failures. In particular, in the case of RAID 6, the expression derived yielded a lower bound of the actual MTTDL. In this report we show that the bound is tight in the range of interest, i.e. in the range of small sector error probabilities.

The remainder of the report is organized as follows. The basic intra-disk redundancy scheme is briefly reviewed in Section II. Section III presents the parameters affecting the performance of the RAID systems that use an intra-disk redundancy scheme. The IPC scheme as well as the traditional redundancy schemes based on RS and SPC codes are presented in Section IV. Also, the model that captures the effect of correlated unrecoverable sector errors is presented for the analysis of these schemes. The erasure correction capability

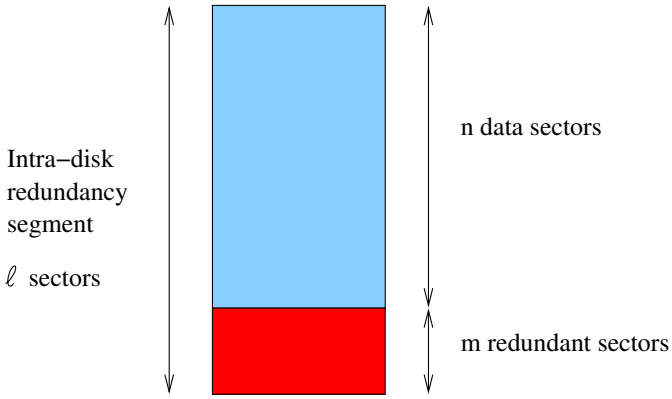


Fig. 1. Basic intra-disk redundancy scheme.

in the presence of correlated as well as independent unrecoverable sector errors, also in the case where the maximum burst length does not exceed the interleaving depth, is evaluated. Section V assesses the reliability of RAID 5 and RAID 6 storage systems that incorporate the coding schemes considered. Closed-form expressions are derived for the MTTDL of RAID 5 and RAID 6 systems in the presence of unrecoverable errors and disk failures. Section VI presents numerical results demonstrating the efficiency of the IPC coding scheme. Section VII concludes the report.

II. Intra-Disk Redundancy Scheme

In this section we briefly review the basic intra-disk redundancy scheme, which works as follows: each strip (stripe unit) is partitioned into segments, and within each segment, a portion of the storage, usually several sectors (called data sectors), is used for storing data, whereas the remainder is reserved for redundant sectors, which are computed based on an erasure code. A number of different schemes can be used to obtain the redundant parity sectors, including the traditional redundancy schemes based on RS and SPC codes. Furthermore, the redundant sectors are optimally placed within the segment to minimize the impact on the throughput performance. The entire segment, comprising ℓ data and parity sectors, is stored contiguously on the same disk, as shown in Fig. 1, where $\ell = n + m$.

The size of a segment should be chosen such that a sufficient degree of storage efficiency, performance and reliability are ensured. For practical reasons, the strip size should be a multiple of the data segment size. In addition, the number m of parity sectors in a segment is a design parameter that can be optimized based on the desired set of operating conditions. In general, more redundancy (large m) provides more protection against unrecoverable media errors. However, it also incurs more overhead in terms of storage space and computations required to obtain and update the parity sectors. Therefore, a judicious trade-off between these competing requirements needs to be made. The overhead $ov^{(\mathrm{IDR})}$ of the intra-disk redundancy scheme is given by

$$ov^{(\mathrm{IDR})} = \frac{m}{\ell - m}, \tag{1}$$

whereas the storage efficiency $se^{(\text{IDR})}$ is given by

$$se^{(\text{IDR})} = \frac{\ell - m}{\ell} . \tag{2}$$

III. System Analysis

The following notation is used for the purpose of our analysis. The parameters are divided into two sets, namely the set of independent and of dependent parameters.

N : number of disks per array group,
n_G : number of array groups in the system,
C_d : disk drive capacity,
S : sector size,
ℓ : number of sectors in a segment,
n_d : number of segments in a disk drive,
m : number of parity sectors in a segment, number of interleaves, or interleaving depth,
$1/\lambda$: mean time to failure for a disk,
P_{bit} : probability of an unrecoverable bit error,
$1/\mu$: mean time to rebuild in the critical mode for a RAID 5 array,
$1/\mu_1$: mean time to rebuild in the degraded mode for a RAID 6 array,
$1/\mu_2$: mean time to rebuild in the critical mode for a RAID 6 array,

$ov^{(\text{IDR})}$: overhead of the intra-disk redundancy scheme,
$se^{(\text{IDR})}$: storage efficiency of the intra-disk redundancy scheme,
$ov^{(\text{RAID})}$: overhead of the RAID scheme,
$se^{(\text{RAID})}$: storage efficiency of the RAID scheme,
$ov^{(\text{RAID+IDR})}$: overall overhead of the entire system,
$se^{(\text{RAID+IDR})}$: overall storage efficiency of the entire system,
P_s : probability of an unrecoverable sector error,
P_{seg} : probability of a segment encountering an unrecoverable sector error,
P_{uf} : probability of an unrecoverable failure,

Assuming that errors are independently occurring over successive bits, the unrecoverable sector error probability P_s is given by

$$P_s = 1 - (1 - P_{\text{bit}})^S , \tag{3}$$

with S expressed in bits.

Similarly, when no coding within the segment is applied ($m = 0$), the unrecoverable segment error probability P_{seg} is given by

$$P_{\text{seg}}(\text{no coding}) = 1 - (1 - P_s)^\ell = 1 - (1 - P_{\text{bit}})^{S\ell} . \tag{4}$$

As there are S bits per sector and ℓ sectors per segment, the number of segments in a disk drive, n_d, is given by

$$n_d = \frac{C_d}{\ell S} \, , \tag{5}$$

with S expressed in bits.

In the critical mode, an unrecoverable failure occurs when at least one out of the n_s segments that need to be read is in error. Consequently, the probability of an unrecoverable failure, P_{uf}, is given by

$$P_{\text{uf}} = 1 - (1 - P_{\text{seg}})^{n_s} \, . \tag{6}$$

For a RAID 5 and a RAID 6 system in the critical mode, the corresponding probabilities of an unrecoverable failure $P_{\text{uf}}^{(1)}$ and $P_{\text{uf}}^{(2)}$ are obtained by setting $n_s = (N-1)n_d$ and $n_s = (N-2)n_d$, as there are $N-1$ and $N-2$ operational disks, respectively. From (5) it follows that

$$P_{\text{uf}}^{(1)} = 1 - (1 - P_{\text{seg}})^{\frac{(N-1)C_d}{\ell S}} \, , \tag{7}$$

and

$$P_{\text{uf}}^{(2)} = 1 - (1 - P_{\text{seg}})^{\frac{(N-2)C_d}{\ell S}} \, . \tag{8}$$

The probability P_{seg} corresponding to the various coding schemes is evaluated in Section IV.

A. Overhead and Storage Efficiency

The overhead and storage efficiency of the RAID scheme chosen are given by

$$ov^{(\text{RAID})} \; = \; \frac{p}{N-p} \, , \tag{9}$$

and

$$se^{(\text{RAID})} \; = \; \frac{N-p}{N} \, , \tag{10}$$

respectively, with

$$p = \begin{cases} 1 & \text{for a RAID 5 system} \\ 2 & \text{for a RAID 6 system.} \end{cases} \tag{11}$$

Note that the above expressions hold if no intra-disk redundancy scheme is used. If an intra-disk redundancy scheme is used, the overall storage efficiency of the entire array (or system) is given by

$$ov^{(\text{RAID+IDR})} \; = \; (1 + ov^{(\text{RAID})})(1 + ov^{(\text{IDR})}) - 1 = \frac{1}{\left(1 - \frac{p}{N}\right)\left(1 - \frac{m}{\ell}\right)} - 1 \, , \tag{12}$$

and

$$se^{(\text{RAID+IDR})} \; = \; se^{(\text{RAID})}\, se^{(\text{IDR})} = \left(1 - \frac{p}{N}\right)\left(1 - \frac{m}{\ell}\right) \, , \tag{13}$$

respectively.

IV. Independent and Correlated Errors

The performance of the intra-disk redundancy scheme is analytically assessed based on two models. According to the first model (independent model), each sector encounters an unrecoverable error, independently of all other sectors, with probability P_s. This implies that the lengths (in number of sectors) of error-free intervals are independent and geometrically distributed with parameter P_s. In addition, we introduce a model for capturing error correlation effects in which sector errors are assumed to occur in bursts. We refer to this model as the correlated model. Let B and I denote the lengths (in number of sectors) of bursts and of the error-free intervals between successive bursts, respectively. Let $\bar{B}$ and $\bar{I}$ denote the corresponding average lengths. These lengths are assumed to be i.i.d., i.e. independently and identically distributed random variables. In particular, the error-free intervals are assumed to be geometrically distributed, as in the independent model, but with a parameter α. Therefore, the probability density function (pdf) $\{a_j\}$ of the length j of a typical error-free interval is given by $a_j = P(I = j) = \alpha(1-\alpha)^{j-1}$ for $j = 1, 2, \ldots$, such that $\bar{I} = 1/\alpha$, with $0 < \alpha \leq 1$. Let also $\{b_j\}$ denote the pdf of the length j of a typical burst of consecutive errors, i.e. $P(B = j) = b_j$ for $j = 1, 2, \ldots$. The average burst length is then given by $\bar{B} = \sum_{j=1}^{\infty} j b_j$ and is assumed to be bounded. Owing to ergodicity, the probability P_s that an arbitrary sector has an unrecoverable error is given by

$$P_s = \frac{\bar{B}}{\bar{B} + \bar{I}} \, . \tag{14}$$

From the above it follows that

$$\alpha = \frac{P_s}{\bar{B}(1 - P_s)} = \frac{P_s}{\bar{B}} + \frac{P_s^2}{\bar{B}} + \frac{P_s^3}{\bar{B}} + \cdots = \frac{P_s}{\bar{B}} + O(P_s^2) \, , \tag{15}$$

and that

$$P_s \leq \frac{\bar{B}}{\bar{B} + 1} \, , \tag{16}$$

given that $\alpha \leq 1$, or, equivalently, $\bar{I} \geq 1$.

Let $\{G_n\}$ denote the complementary cumulative density function (ccdf) of the burst size B. Then G_n denotes the probability that the length of a burst is greater than or equal to n, i.e. $G_n \triangleq \sum_{j=n}^{\infty} b_j$, for $n = 1, 2, \ldots$. Consequently, the probability that a burst of more than m consecutive errors occurs is equal to G_{m+1}.

Remark 1. Note that the independent model is a special case of the correlated model in which the $\{b_j\}$ distribution is geometric with parameter $1 - P_s$, i.e. $b_j = (1 - P_s)P_s^{j-1}$ for $j = 1, 2, \ldots$. Therefore, $\bar{B} = 1/(1 - P_s)$ and $G_j = P_s^{j-1}$ for $j = 1, 2, \ldots$. Also, in this case it holds that $\alpha = 1 - P_s$, with $0 \leq P_s < 1$.

Let us consider the sectors divided into groups of ℓ ($\ell > m$) successive sectors, with each such group constituting a segment. If no coding scheme is applied ($m = 0$), a segment is in error if there is an unrecoverable sector error. For the independent model, and according to (4), the probability P_{seg} that a segment is in error is then given by

$$P_{\text{seg}} = 1 - (1 - P_s)^{\ell} = \ell P_s + O(P_s^2) \, . \tag{17}$$

For the correlated model, the segment is correct if the first sector is correct and the subsequent $\ell - 1$ sectors are also correct. The probability of the first sector being correct is equal to $1 - P_s$, whereas from the

geometric assumption the probability of each subsequent sector being correct is equal to $1-\alpha$. By making use of (15) we obtain

$$\begin{aligned} P_{\text{seg}} &= 1-(1-P_s)(1-\alpha)^{\ell-1} = 1-(1-P_s)\left(1-\frac{P_s}{\bar{B}}-O(P_s^2)\right)^{\ell-1} = \\ &= \left(1+\frac{\ell-1}{\bar{B}}\right)P_s+O(P_s^2)\ . \end{aligned} \tag{18}$$

We now proceed with the evaluation of P_{seg} for various coding schemes. In particular, we consider P_{seg} expressed as a series expansion in powers of P_s, i.e. $P_{\text{seg}} = \sum_{i=1}^{\infty} c_i P_s^i$, with the coefficients c_i being independent of P_s. It turns out that in the case of the correlated model, the performance difference between the coding schemes considered can be demonstrated by considering the power series taken to the second order. That is, it suffices to make a power series expansion of P_{seg} in P_s of the form $P_{\text{seg}} = \sum_{i=1}^{2} c_i P_s^i + O(P_s^3)$. First we establish the following propositions which hold for the correlated model and independently of the coding scheme used.

Proposition 1: The probability $P_{\text{seg}}^{(k)}$ that a segment contains k ($k \leq \ell/2$) bursts of errors and is in error is of order $O(P_s^k)$.

Proof: See Appendix A. ■

Proposition 2: It holds that $P_{\text{seg}} = \sum_{i=1}^{\infty} c_i P_s^i$, with the coefficient c_i derived based only on $P_{\text{seg}}^{(1)}, \ldots, P_{\text{seg}}^{(i)}$.

Proof: By conditioning on the number of bursts of errors in a segment, and using Proposition 1 we obtain

$$\begin{aligned} P_{\text{seg}} &= \sum_{k=1}^{\ell/2} P(\text{segment contains } k \text{ bursts of errors and is in error}) = \\ &= \sum_{k=1}^{\ell/2} P_{\text{seg}}^{(k)} = \sum_{k=1}^{i} P_{\text{seg}}^{(k)} + \sum_{k=i+1}^{\ell/2} P_{\text{seg}}^{(k)} = \sum_{k=1}^{i} P_{\text{seg}}^{(k)} + O(P_s^{i+1})\ . \end{aligned} \tag{19}$$

■

A. Reed–Solomon (RS) Coding

Reed–Solomon (RS) coding is the standard choice for erasure correction when the implementation complexity is not a constraint. This is because these codes provide the best possible erasure correction capability for a given number of parity symbols, i.e. for a given storage efficiency (code rate). Essentially, for a code with m parity symbols in a codeword of n symbols, any m erasures in the block of n symbols can be corrected. RS codes are used in a wide variety of applications and are the primary mechanism that allows the stringent uncorrectable error probability specification of HDDs to be met. Note that the RS codes considered here provide an additional level of redundancy to that of the built-in ECC scheme.

The performance of the RS scheme is the best that can be achieved. With such a code, the probability of a segment being in error is equal to the probability of getting more than m unrecoverable sector errors per

segment and is given by

$$P_{\text{seg}}^{\text{RS}} \;=\; \sum_{j=m+1}^{\ell} \binom{\ell}{j} P_s^j (1-P_s)^{\ell-j} = \binom{\ell}{m+1} P_s^{m+1} + O(P_s^{m+2}) \; . \tag{20}$$

In the case of the correlated model, an approximate expression for the probability of a segment being in error is given by the following theorem.

Theorem 1: It holds that

$$P_{\text{seg}}^{\text{RS}} \;=\; c_1^{\text{RS}} \, P_s + c_2^{\text{RS}} \, P_s^2 + O(P_s^3) \; , \tag{21}$$

where

$$c_1^{\text{RS}} = 1 + \frac{(\ell-m-1)G_{m+1} - \sum_{j=1}^{m} G_j}{\bar{B}} \; , \tag{22}$$

$$c_2^{\text{RS}} = \left[\binom{\ell-m}{2} \, GG_{m+1} - \binom{\ell-m-1}{2} \, GG_{m+2} \right] \frac{1}{\bar{B}^2} \; , \tag{23}$$

with

$$GG_j \triangleq \sum_{\substack{(l_1,l_2)\in(\mathbb{N}\times\mathbb{N}) \\ l_1+l_2=j}} G_{l_1} G_{l_2} \; . \tag{24}$$

Proof: See Appendix B. ■

Corollary 1: The coefficient c_1^{RS} is equal to zero if and only if $G_{m+1} = 0$, i.e. the maximum burst length does not exceed m.

Proof: Note that c_1^{RS} can also be written as $[(\ell-m)G_{m+1} + \sum_{j=m+2}^{\infty} G_j]/\bar{B}$, which is equal to zero if and only if $G_{m+1} = 0$. ■

Corollary 2: Both coefficients c_1^{RS} and c_2^{RS} are equal to zero if and only if $G_{\lceil \frac{m+1}{2} \rceil} = 0$, i.e. the total number of errors of any two bursts does not exceed m.

Proof: See Appendix B. ■

Remark 2. According to Remark 1, the independent model is a special case of the correlated model in which the b_j distribution is geometric with parameter $1-P_s$, i.e. $b_j = (1-P_s)P_s^{j-1}$ for $j = 1, 2, \ldots$. Thus, $\bar{B} = 1/(1-P_s)$, $G_j = P_s^{j-1}$, and also $GG_j = (j-1)P_s^{j-2}$. Substituting these into (22) and (23) yields $c_1^{\text{RS}} = (\ell-m)P_s^m + O(P_s^{m+1})$ and $c_2^{\text{RS}} = [m(\ell-m)(\ell-m-1)/2]P_s^{m-1} + O(P_s^m)$, respectively. Thus, it now follows from (21) that $P_{\text{seg}}^{\text{RS}} = [(\ell-m)(m\ell - m^2 - m + 2)/2]P_s^{m+1} + O(P_s^{m+2})$. This expression, however, in general does not agree with (20) and therefore is not correct. The reason for this inconsistency is that although this expression is of order $O(P_s^{m+1})$, it is derived by considering a series expansion in powers of P_s taken to second order only.

B. Single-Parity Check (SPC) Coding

The simplest coding scheme is one in which a single parity sector is computed by using the XOR operation on $\ell - 1$ data sectors to form a segment with ℓ sectors in total. Such a scheme can tolerate a single erasure anywhere in the segment. In fact, the parity in a RAID 5 scheme is based on such a single parity-check (SPC) scheme, albeit with the redundancy along the RAID dimension. The probability of a segment being in error is equal to the probability of getting at least two unrecoverable sector errors. The independent model yields

$$P_{\text{seg}}^{\text{SPC}} \;=\; \sum_{j=2}^{\ell} \binom{\ell}{j} P_s^j (1-P_s)^{\ell-j} \;=\; \frac{\ell(\ell-1)}{2}\, P_s^2 + O(P_s^3)\;. \tag{25}$$

In the case of the correlated model, the probability of a segment being in error is given by the following theorem.

Theorem 2: It holds that

$$P_{\text{seg}}^{\text{SPC}} \;=\; 1-(1-P_s)(1-\alpha)^{\ell-1} - \frac{P_s}{\bar{B}}\,[2(1-\alpha)+(\ell-2)b_1](1-\alpha)^{\ell-3}\;. \tag{26}$$

Proof: See Appendix C. ■

An approximate expression for $P_{\text{seg}}^{\text{SPC}}$ based on a series expansion in powers of P_s is given by the following theorem.

Theorem 3: It holds that

$$P_{\text{seg}}^{\text{SPC}} \;=\; \left[1+\frac{(\ell-2)G_2-1}{\bar{B}}\right] P_s + \frac{(\ell-2)[\ell-1-2(\ell-3)G_2]}{2\bar{B}^2}\, P_s^2 + O(P_s^3)\;. \tag{27}$$

Proof: Note that the SPC coding scheme is a special case of the RS coding scheme in which only a single sector error can be corrected in a segment. Expression (27) is therefore derived from (21)–(24) by setting $m=1$. ■

Remark 3. Assuming a geometric distribution $\{b_j\}$ with parameter $1-P_s$, Eq. (27) yields $P_{\text{seg}}^{\text{SPC}} = [(\ell-1)\ell/2]P_s^2 + O(P_s^3)$, which is in agreement with the expression derived in (25) for the independent model.

C. Interleaved Parity-Check (IPC) Coding

Here we review the interleaved parity-check (IPC) coding scheme that has a simplicity akin to that of the SPC scheme but considerably better performance. In this scheme, n ($n=\ell-m$) contiguous data sectors are conceptually arranged in a matrix containing m columns. Data sectors in a column are XORed to obtain the parity sector and together form an *interleave*. An IPC scheme with m ($m \le \ell/2$) interleaves per segment, i.e. ℓ/m sectors per interleave, has the capability of correcting a single error per interleave. Consequently, a segment is in error if there is at least one interleave in which there are at least two unrecoverable sector errors. Note that this scheme can correct a single burst of m consecutive errors occurring in a segment. However, unlike the RS scheme, it in general does not have the capability of correcting any m sector errors in a segment, implying that $P_{\text{seg}}^{\text{IPC}} > P_{\text{seg}}^{\text{RS}}$.

According to the independent model, the probability $P_{\text{interleave}}$ of an interleave being in error is given by

$$\begin{aligned} P_{\text{interleave}} \;&=\; \sum_{j=2}^{\ell/m} \binom{\ell/m}{j} P_s^j (1-P_s)^{\ell/m-j} \;= \\ &=\; \frac{\frac{\ell}{m}\left(\frac{\ell}{m}-1\right)}{2}\, P_s^2 + O(P_s^3) \;=\; \frac{\ell(\ell-m)}{2m^2}\, P_s^2 + O(P_s^3)\;. \end{aligned} \tag{28}$$

Consequently,

$$P_{\text{seg}}^{\text{IPC}} \;=\; 1-(1-P_{\text{interleave}})^m \;=\; \frac{\ell(\ell-m)}{2m}\,P_s^2+O(P_s^3)\;. \tag{29}$$

In the case of the correlated model, an approximate expression for the probability of a segment being in error is given by the following theorem.

Theorem 4: It holds that

$$P_{\text{seg}}^{\text{IPC}} \;=\; c_1^{\text{IPC}}\,P_s+c_2^{\text{IPC}}\,P_s^2+O(P_s^3)\;, \tag{30}$$

where

$$c_1^{\text{IPC}}=1+\frac{(\ell-m-1)G_{m+1}-\sum_{j=1}^{m}G_j}{\bar{B}}\;, \tag{31}$$

$$\begin{aligned} c_2^{\text{IPC}}=\Bigg\{\frac{\ell-m}{2m}\quad\Bigg[&2-\ell+2\sum_{j=2}^{m}GG_j-2(m-1)\;GG_{m+1}+2(\ell-2)\sum_{j=1}^{m}G_j-\\ &-2(m-1)(\ell-m-2)\;G_{m+1}\Bigg]-\frac{(\ell-2m)(\ell-m-2)}{2m}\;GG_{m+2}\Bigg\}\;\frac{1}{\bar{B}^2}\;, \end{aligned} \tag{32}$$

and GG_j is given in (24).

Proof: See Appendix D. ■

Remark 4. Assuming a geometric distribution $\{b_j\}$ with parameter $1-P_s$, Eq. (30) yields $P_{\text{seg}}^{\text{IPC}}=[(\ell-m)\ell/(2m)]P_s^2+O(P_s^3)$, which is in agreement with the expression derived in (29) for the independent model.

Remark 5. From (22) and (31) it follows that $c_1^{\text{IPC}}=c_1^{\text{RS}}$. Thus, from (21)–(23) and (30)–(32) it follows that $P_{\text{seg}}^{\text{IPC}}\approx P_{\text{seg}}^{\text{RS}}$ given that $P_{\text{seg}}^{\text{IPC}}-P_{\text{seg}}^{\text{RS}}=O(P_s^2)$. Therefore, when the unrecoverable sector errors are known to occur in bursts whose length can exceed m with a nonnegligible likelihood, using an IPC check code is preferable because it is as efficient as the more complex RS code. This is because the interleaved coding scheme provides additional gain by recovering from consecutive unrecoverable sector errors, which can be as many as the interleaving depth. On the other hand, if the maximum burst length does not exceed m, Corollary 1 implies that $P_{\text{seg}}^{\text{RS}}$ and $P_{\text{seg}}^{\text{IPC}}$ are no longer of order $O(P_s)$. In this case, the two probabilities are of order $O(P_s^2)$ and significantly different.

D. Summary

TABLE I

APPROXIMATE P_{seg} (FOR $P_s\ll$).

Coding Scheme	**Model for Errors**	
	Independent	**Correlated** (for $G_{m+1}>0$)
None	ℓP_s	$(1+\frac{\ell-1}{\bar{B}})P_s$
RS	$\binom{\ell}{m+1}P_s^{m+1}$	$\left[1+\frac{(\ell-m-1)G_{m+1}-\sum_{j=1}^{m}G_j}{\bar{B}}\right]P_s$
SPC	$\frac{\ell(\ell-1)}{2}P_s^2$	$\left[1+\frac{(\ell-2)G_2-1}{\bar{B}}\right]P_s$
IPC	$\frac{\ell(\ell-m)}{2m}P_s^2$	$\left[1+\frac{(\ell-m-1)G_{m+1}-\sum_{j=1}^{m}G_j}{\bar{B}}\right]P_s$

TABLE II

APPROXIMATE P_{seg}.

Coding Scheme	Model for Errors	
	Independent	**Correlated**
None	5.2×10^{-9}	5.0×10^{-9}
RS	6.2×10^{-81}	2.5×10^{-12}
SPC	1.3×10^{-17}	9.5×10^{-11}
IPC	1.6×10^{-18}	2.5×10^{-12}

TABLE III

APPROXIMATE $P_{\text{uf}}^{(1)}$ FOR RAID 5 WITH $N = 8$.

Coding Scheme	Model for Errors	
	Independent	**Correlated**
None	1.5×10^{-1}	1.5×10^{-1}
RS	2.0×10^{-73}	7.9×10^{-5}
SPC	4.3×10^{-10}	3.1×10^{-3}
IPC	5.1×10^{-11}	7.9×10^{-5}

TABLE IV

APPROXIMATE $P_{\text{uf}}^{(2)}$ FOR RAID 6 WITH $N = 16$.

Coding Scheme	Model for Errors	
	Independent	**Correlated**
None	2.8×10^{-1}	2.7×10^{-1}
RS	3.9×10^{-73}	1.6×10^{-4}
SPC	8.7×10^{-10}	6.1×10^{-3}
IPC	1.0×10^{-10}	1.7×10^{-4}

Table I summarizes the results obtained for the probability P_{seg} that a segment is in error for the various models and coding schemes, assuming that the sector error probability is small.

E. Numerical Results

We consider SATA drives with $C_d = 300$ GB and $P_{\text{bit}} = 10^{-14}$. Assuming a sector size of 512 bytes and according to (3), the equivalent unrecoverable sector error probability is $P_s \approx P_{\text{bit}} \times 4096$, which is 4.096×10^{-11}. We also consider a segment comprised of $\ell = 128$ sectors with $m = 8$ interleaves. We now consider the following error-burst length distribution:

$$\begin{aligned}\mathbf{b} \;=\; & [0.9812\ 0.016\ 0.0013\ 0.0003; 0.0003\ 0.0002\ 0.0001\ 0.0001\ 0\ 0.0001\ 0 \\ & 0.0001\ 0.0001\ 0\ 0.0001\ 0.0001] \,. \end{aligned} \tag{33}$$

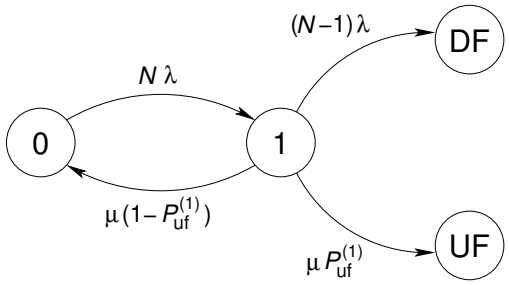


Fig. 2. Reliability model for a RAID 5 array.

Then, we have bursts of at most 16 sectors with $\bar{B} = 1.0291$, $G_2 = 0.0188$, and $G_9 = 0.0005$. These values are based on actual data collected from the field for a product that is currently being shipped. The results for P_{seg} are listed in Table II. The corresponding unrecoverable failure probabilities for a RAID 5 array with $N = 8$ and a RAID 6 array with $N = 16$ are listed in Table III and Table IV, respectively. From the results it follows that in the case of correlated errors, the proposed IPC scheme improves the unrecoverable failure probability by two orders of magnitude compared with the SPC scheme. This is also the improvement we would get by using the more complex RS code.

V. Continuous-Time Markov Chain (CTMC) Models

Here we derive the MTTDL for a RAID 5 and a RAID 6 disk array. Assuming independent and exponentially distributed disk failures and rebuild times, the MTTDLs for the two disk arrays are obtained using CTMC models. The numbered states of the Markov models represent the number of failed disks. The DF and UF states represent a data loss due to a disk failure and an unrecoverable sector failure, respectively.

Assuming that the MTTDL of a single array is exponentially distributed, the MTTDL of a RAID system, $MTTDL_{\text{sys}}$, comprising n_G arrays is subsequently obtained as follows:

$$MTTDL_{\text{sys}} = \frac{MTTDL}{n_G}\,. \tag{34}$$

A. Intra-Disk Redundancy with RAID 5

The CTMC model for a RAID 5 disk array is shown in Fig. 2. The infinitesimal generator matrix $\mathbf{Q}$ is given by

$$\begin{bmatrix} -N\lambda & N\lambda & 0 & 0 \\ \mu(1-P_{\text{uf}}^{(1)}) & -\mu-(N-1)\lambda & (N-1)\lambda & \mu P_{\text{uf}}^{(1)} \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}.$$

In particular, the submatrix corresponding to the transient states 0 and 1 is

$$\mathbf{Q}_T = \begin{bmatrix} -N\lambda & N\lambda \\ \mu(1-P_{\text{uf}}^{(1)}) & -\mu-(N-1)\lambda \end{bmatrix}.$$

The vector τ of the average time spent in the transient states before a failure occurs, i.e. before the Markov chain enters either one of the absorbing states DF and UF, is obtained based on the following relation [3]

$$\tau \mathbf{Q}_T = -\mathbf{P}_T(0) \ ,$$

where $\tau = [\tau_0 \ \ \tau_1]$ and $\mathbf{P}_T(0) = [1 \ \ 0]$. Solving the above equation for τ yields

$$\tau_0 = \frac{(N-1)\lambda + \mu}{N\lambda[(N-1)\lambda + \mu P_{\text{uf}}^{(1)}]}, \quad \tau_1 = \frac{N\lambda}{N\lambda[(N-1)\lambda + \mu P_{\text{uf}}^{(1)}]}. \tag{35}$$

Finally, the mean time to data loss is given by

$$MTTDL = \tau_0 + \tau_1 = \frac{(2N-1)\lambda + \mu}{N\lambda[(N-1)\lambda + \mu P_{\text{uf}}^{(1)}]} \ , \tag{36}$$

where $P_{\text{uf}}^{(1)}$ is given by (7).

Note that for $P_{\text{uf}}^{(1)} = 0$ (which holds when $P_s = 0$) and $\lambda \ll \mu$, Eq. (36) can be approximated as follows:

$$MTTDL \ \approxeq \ \frac{\mu}{N(N-1)\lambda^2} \ , \tag{37}$$

which is the same result as derived in [4].

B. Intra-Disk Redundancy with RAID 6

A RAID 6 array can tolerate up to two disk failures; thus it is in the critical mode when the disk array has two disk failures. When the first disk fails, the disk array enters into the degraded mode, in which the rebuild of the failing disk takes place while still serving I/O requests. The rebuild of a segment of the failed drive is performed based on up to $N-1$ corresponding segments residing on the remaining disks. When the rebuild fails, then two or more of these segments are in error. Note, however, that the converse does not hold. It may well be that two segments are in error and the corresponding sectors in error are in such positions that the RAID 6 reconstruction mechanism can correct all of them. Consequently, the probability P_{recf} that a given segment of the failed disk cannot be reconstructed is upper-bounded by the probability that two or more of the corresponding segments residing in the remaining disks are in error. As segments residing in different disks are independent, the upper bound $P_{\text{recf}}^{\text{UB}}$ of the probability P_{recf} is given by

$$P_{\text{recf}}^{\text{UB}} = \sum_{j=2}^{N-1} \binom{N-1}{j} P_{\text{seg}}^j (1 - P_{\text{seg}})^{N-1-j} \approx \binom{N-1}{2} P_{\text{seg}}^2 \ . \tag{38}$$

Furthermore, the reconstruction of each of the n_d segments of the failed disk is independent of the reconstruction of the other segments of this disk. Consequently, the upper bound $P_{\text{uf}}^{(\text{r})}$ of the probability that an unrecoverable failure occurs because the rebuild of the failed disk cannot be completed is given by

$$P_{\text{uf}}^{(\text{r})} = 1 - (1 - P_{\text{recf}}^{\text{UB}})^{n_d} \ , \tag{39}$$

where n_d is given by (5).

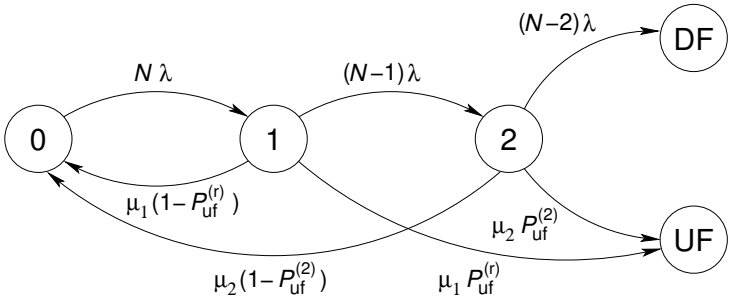


Fig. 3. Reliability model for a RAID 6 array.

Assuming that the rebuild times in the degraded and the critical mode are exponentially distributed with parameters μ_1 and μ_2, respectively, we obtain the CTMC model shown in Fig. 3. Note that, in contrast to the case of a RAID 5 array, the rate from state 1 to UF is $\mu_1 P_{\rm uf}^{(\rm r)}$ instead of $\mu_1 P_{\rm uf}^{(1)}$.

The infinitesimal generator submatrix $\mathbf{Q}_T$, restricted to the transient states 0, 1 and 2, is given by

$$\left[\begin{array}{ccc} -N\lambda & N\lambda & 0 \\ \mu_1(1-P_{\rm uf}^{(\rm r)}) & -(N-1)\lambda-\mu_1 & (N-1)\lambda \\ \mu_2(1-P_{\rm uf}^{(2)}) & 0 & -(N-2)\lambda-\mu_2 \end{array}\right] .$$

Solving the equation $\tau\mathbf{Q}_T = -\mathbf{P}_T(0)$ for $\tau = [\tau_0 \;\; \tau_1 \;\; \tau_2]$, with $\mathbf{P}_T(0) = [1 \;\; 0 \;\; 0]$, we get

$$\tau_0 = \frac{[(N-1)\lambda+\mu_1]\,[(N-2)\lambda+\mu_2]}{N\lambda V} , \tag{40}$$

$$\tau_1 = \frac{(N-2)\lambda+\mu_2}{V} , \qquad \tau_2 = \frac{(N-1)\lambda}{V} , \tag{41}$$

where

$$V \triangleq [(N-1)\lambda+\mu_1 P_{\rm uf}^{(\rm r)}][(N-2)\lambda+\mu_2 P_{\rm uf}^{(2)}] + \mu_1\mu_2 P_{\rm uf}^{(\rm r)}(1-P_{\rm uf}^{(2)}) , \tag{42}$$

and $P_{\rm uf}^{(\rm r)}$ and $P_{\rm uf}^{(2)}$ are given by (39) and (8), respectively.

Then, we have

$$MTTDL = \tau_0 + \tau_1 + \tau_2 . \tag{43}$$

Note that for $P_{\rm uf}^{(\rm r)} = P_{\rm uf}^{(2)} = 0$ (which holds when $P_s = 0$) and $\lambda \ll \mu_1 = \mu_2 = \mu$, Eq. (43) can be approximated as follows:

$$MTTDL \;\approxeq\; \frac{\mu^2}{N(N-1)(N-2)\lambda^3} , \tag{44}$$

which is the same result as reported in [5].

VI. Numerical Examples

Here we assess the reliability of various schemes considered above through illustrative examples. We consider different systems using SATA 300GB disk drives and storing a user data base of 10 PB. The

TABLE V

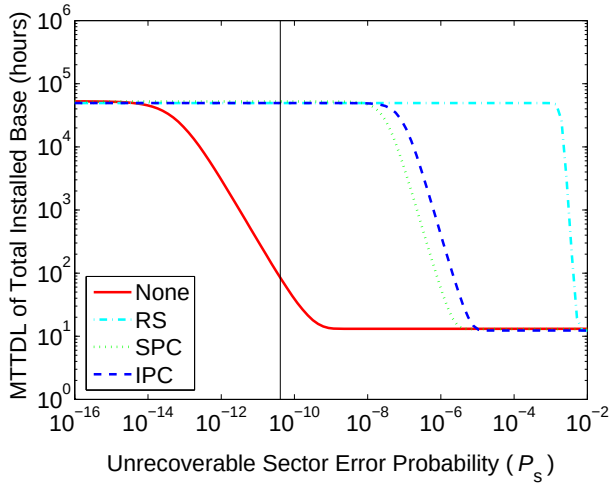
NUMERICAL VALUES

Parameter	Value	
	SATA	**SCSI**
C_d	300 GB	146 GB
P_{bit}	10^{-14}	10^{-15}
λ^{-1}	500 000 h	1 000 000 h
μ^{-1}	17.8 h	9.3 h
μ_1^{-1}	17.8 h	9.3 h
μ_2^{-1}	17.8 h	9.3 h
N	8 for RAID 5 16 for RAID 6	
S	512 bytes = 4096 bits	
ℓ	128 sectors	
m	8 interleaves per segment	

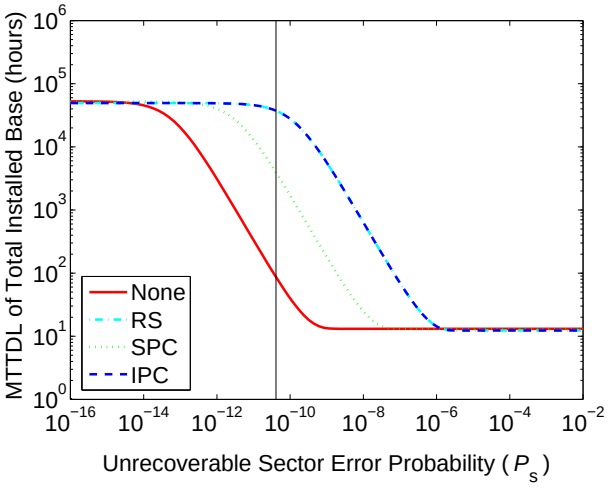
disk drive parameters are summarized in Table V. In particular, for a sector size of 512 bytes, we have $P_s = 4.096 \times 10^{-11}$.

From (10), (11) and (13) it follows that the storage efficiency of the entire system is independent of the RAID configuration if the arrays in a RAID 6 system are twice the size of the arrays in a RAID 5 system. For a RAID 5 system with $N = 8$, when there is no intra-disk redundancy, the required number of arrays to store the user data is equal to 4762 (i.e. 10 PB/(7×300 GB)), whereas for a RAID 6 system with $N = 16$, it is equal to 2381 (i.e. 10 PB/(14×300 GB)). The corresponding storage efficiency is equal to 7/8, i.e. 0.875. For the RS, SPC, and IPC redundancy schemes, the intra-disk storage efficiency is obtained from (2) by setting $m = 8, 1,$ and 8, respectively. For $\ell = 128$, the storage efficiency is equal to 0.94, 0.99, and 0.94, respectively. Furthermore, the required number of arrays for a RAID 5 configuration is obtained as the ratio of 4762 to the intra-disk storage efficiency and is equal to 5080, 4800, and 5080, respectively. Similarly, for a RAID 6 configuration, the required number of arrays is equal to 2540, 2400, and 2540, respectively. The overall storage efficiency is obtained by (13) and is equal to 0.82, 0.87, and 0.82, respectively.

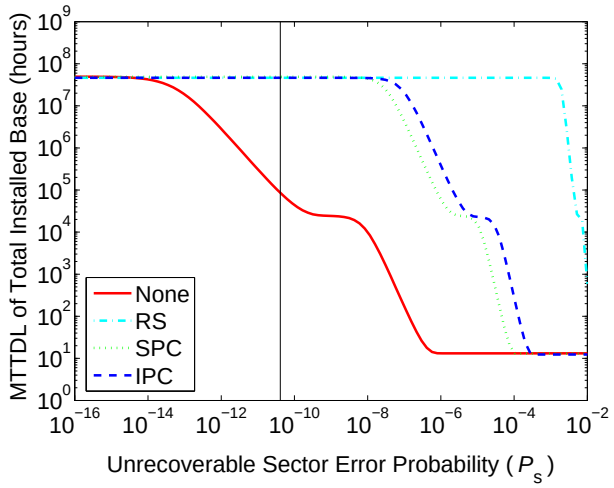
The combined effects of disk and unrecoverable failures can be seen in Fig. 4 as a function of the unrecoverable sector error probability. The vertical line in the figures indicates the SATA drive specification for unrecoverable sector errors. Note that for small sector error probabilities, the MTTDL remains unaffected because data is lost owing to a disk rather than an unrecoverable failure. In particular, the MTTDL of a RAID 6 system is three orders of magnitude higher than that of a RAID 5 system. However, as the sector error probability increases, the probability of an unrecoverable failure in the critical mode P_{uf} also increases and therefore the MTTDL decreases. This decrease ends when the sector error probability is such that the corresponding P_{uf} is extremely high, i.e. close to one. In this case the rebuild process in critical



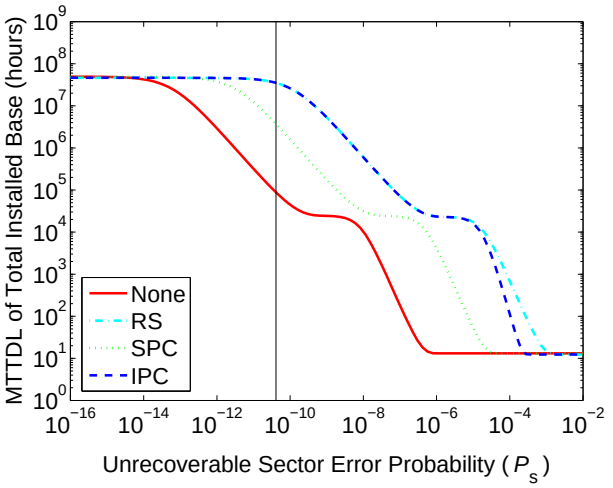
(a) RAID 5 with independent unrecoverable sector errors.



(b) RAID 5 with correlated unrecoverable sector errors.



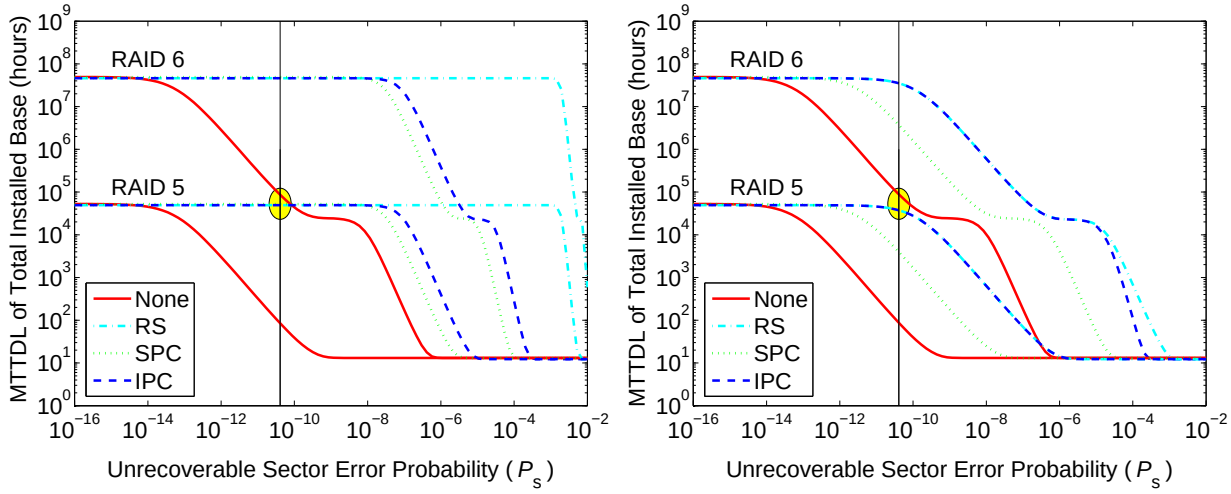
(c) RAID 6 with independent unrecoverable sector errors.



(d) RAID 6 with correlated unrecoverable sector errors.

Fig. 4. MTTDL for RAID 5 and RAID 6 systems with unrecoverable sector errors ($\ell = 128$, $m = 8$).

mode cannot be successfully completed because of an unrecoverable failure. Consequently, the MTTDL is the mean time until the system (i.e. any of the disk arrays) enters the critical mode. In a RAID 5 system, this occurs when the first disk fails after an expected time of $1/(n_G N\lambda)$. In a RAID 6 system, this occurs when a second disk fails while the system is in the degraded mode. Note that this corresponds to the MTTDL of a RAID 5 system without unrecoverable sector errors. This also explains why the RAID 6 curves become flat at the height of a RAID 5 system, as can be seen in Fig. 5. This range of sector error probabilities is of primary interest because it includes the SATA drive specification. Note that in this range the upper



(a) Independent unrecoverable sector errors.

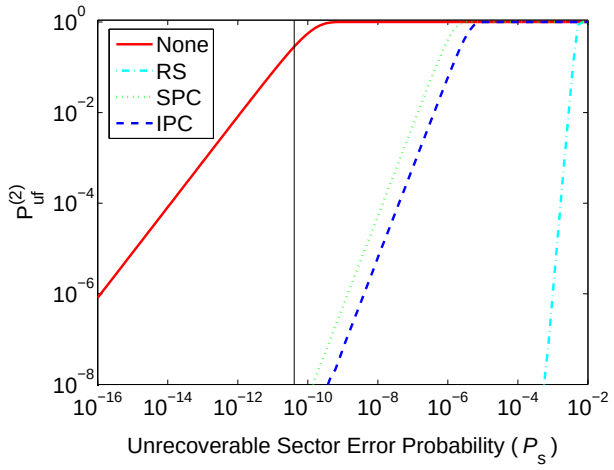
(b) Correlated unrecoverable sector errors.

Fig. 5. RAID 5 vs. RAID 6 systems with independent or correlated unrecoverable sector errors ($\ell = 128$, $m = 8$).

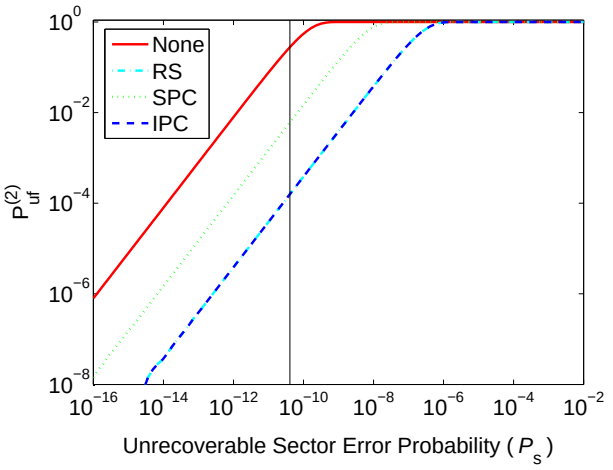
bound $P_{\text{uf}}^{(\text{r})}$ of the probability (as well as the probability itself) of an unrecoverable failure in the degraded mode is negligible, as shown in Fig. 6. Consequently, in this range of sector error probabilities, called the first range, the RAID 6 curves are tight lower bounds of the actual MTTDL. We subsequently consider the second range of the remaining sector error probabilities. As the sector error probability further increases, the upper bound $P_{\text{uf}}^{(\text{r})}$ of the probability of an unrecoverable failure in the degraded mode starts becoming significant, as shown in Fig. 6, resulting in a further decrease of the MTTDL. This decrease ends when the sector error probability is such that the corresponding $P_{\text{uf}}^{(\text{r})}$ is extremely high, i.e. close to one. In this case the rebuild process in degraded mode cannot be successfully completed because of an unrecoverable failure. Consequently, the MTTDL is the mean time until the system (i.e. any of the disk arrays) enters the degraded mode. In a RAID 6 system, this occurs when the first disk fails after an expected time of $1/(n_G N \lambda)$, which is the same as for a RAID 5 system.

In all cases, the intra-disk redundancy schemes considerably improve the reliability over a wide range of sector error probabilities. In particular, in the case of correlated errors, the IPC coding scheme offers the maximum possible improvement that is also achieved by the RS coding scheme. Furthermore, for large sector error probabilities, the gain from the use of the intra-disk redundancy schemes is smaller for correlated errors than for independent errors. Note that according to Remark 5, in the case of correlated errors the MTTDL for the IPC scheme is roughly the same as for the optimum, albeit more complex, RS coding scheme. This is because for both the IPC and RS schemes, and for small sector error probabilities, the probability of an unrecoverable failure is essentially determined by the event of encountering a single burst of more than 8 consecutive errors.

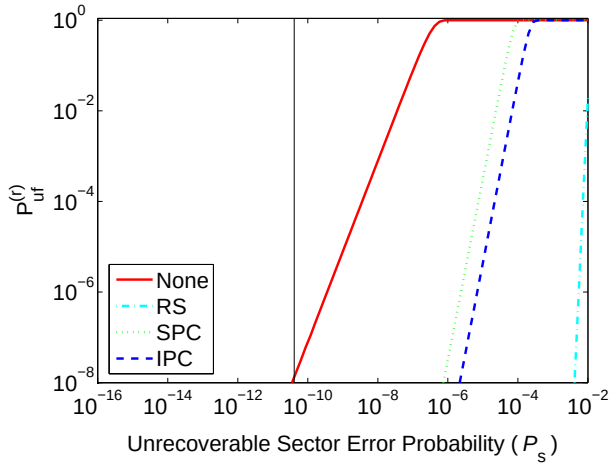
Both the plain RAID 6 and the RAID 5 + IPC system improve the reliability over the plain RAID 5 system,



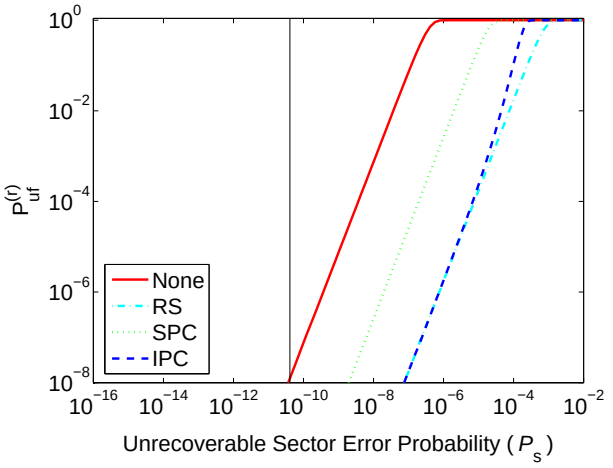
(a) $P_{\text{uf}}^{(2)}$ with independent unrecoverable sector errors.



(b) $P_{\text{uf}}^{(2)}$ with correlated unrecoverable sector errors.



(c) $P_{\text{uf}}^{(\text{r})}$ with independent unrecoverable sector errors.

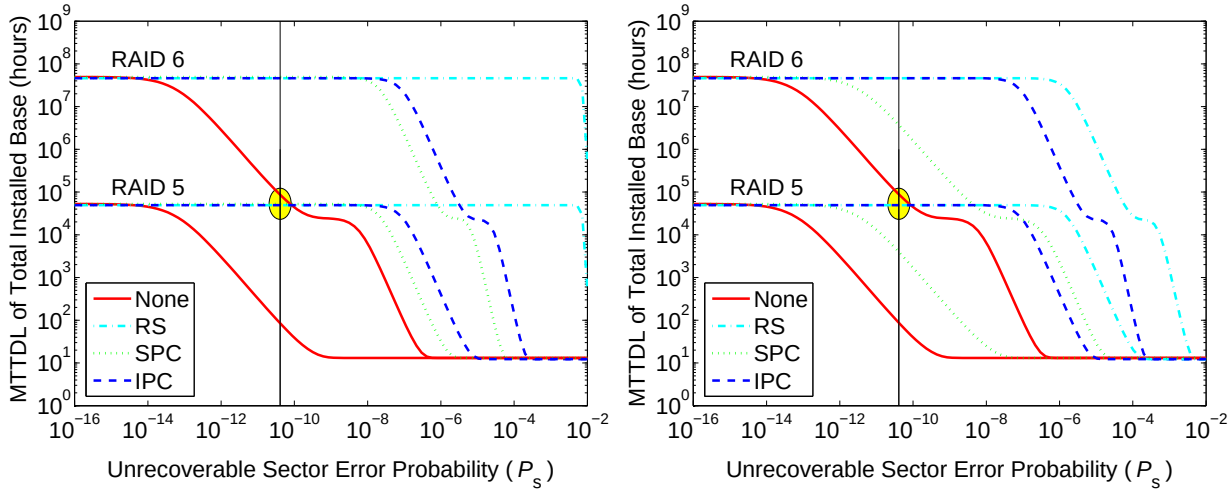


(d) $P_{\text{uf}}^{(\text{r})}$ with correlated unrecoverable sector errors.

Fig. 6. Probabilities $P_{\text{uf}}^{(2)}$ and $P_{\text{uf}}^{(\text{r})}$ for a RAID 6 system ($\ell = 128, m = 8$).

with the respective gains shown in Fig. 5. Note that in the case of SATA drives the resulting MTTDLs for these two systems are of the same order (depicted by the ellipse) for both independent and correlated errors. Therefore, the RAID 5 + IPC system is an attractive alternative to a RAID 6 system, in particular because its I/O performance is better than that of a RAID 6 system [2].

Next we consider a system in which both the segment size and the interleaving length are twice as long, i.e. $\ell = 256$ and $m = 16$, such that the number of sectors per interleave remains the same. Also the storage efficiency and the required number of arrays for the IPC and RS redundancy schemes remain the same. The results obtained are shown in Fig. 7. For the IPC scheme, and in the case of independent sector errors, the



(a) Independent unrecoverable sector errors.

(b) Correlated unrecoverable sector errors.

Fig. 7. RAID 5 vs. RAID 6 systems with independent or correlated unrecoverable sector errors ($\ell = 256$, $m = 16$).

MTTDL is not affected because (28) implies that the probability $P_{\text{interleave}}$ of an interleave being in error remains unaffected. The MTTDL for the RS scheme, however, improves. In the case of correlated errors, the MTTDLs for both the IPC and RS schemes are significantly better, because now all single bursts in a segment can be corrected. Furthermore, the MTTDL for the IPC scheme is no longer the same as for the RS coding scheme. It is worse because, unlike the RS scheme, the IPC scheme in general does not have the capability of correcting any two bursts in a segment having a total number of errors smaller than 16.

VII. Conclusions

Owing to increasing disk capacities and the adoption of low-cost disks in modern data storage systems, unrecoverable errors are becoming a significant cause of user data loss. To cope with this issue, we consider the XOR-based intra-disk redundancy scheme called interleaved parity-check (IPC) coding scheme. A new model capturing the effect of correlated unrecoverable sector errors was developed to analyze this scheme. Traditional redundancy schemes based on Reed–Solomon (RS) codes and single-parity-check codes were also analyzed. Closed-form expressions were derived for the mean time to data loss of RAID 5 and RAID 6 systems in the presence of unrecoverable errors and disk failures.

The results obtained demonstrate that the proposed IPC scheme considerably improves the reliability over a wide range of sector error probabilities. In particular, in the case of correlated errors where the maximum burst length exceeds the interleaving depth, the IPC coding scheme offers the maximum possible improvement that is also achieved by the RS coding scheme. Furthermore, in the case of SATA disk drives, the IPC scheme applied to a RAID 5 system offers the same level of reliability as a RAID 6 system.

APPENDIX A
NUMBER OF BURSTS OF ERRORS IN A SEGMENT

Proof of Proposition 1.

Let us consider an instance of k bursts in a segment and let us denote by $\vec{L}$ the vector $(L_1, \ldots, L_k)$ of the corresponding burst lengths and by $\vec{S}$ the vector $(S_1, \ldots, S_k)$ of their corresponding starting sector positions with $1 \leq S_1 < \cdots < S_k \leq \ell$. The length of the error-free interval I_j following the j-th burst is then given by $S_{j+1} - S_j - L_j$, for $j = 1, 2, \ldots, k-1$, implying that $S_{j+1} \geq S_j + L_j + 1$. Also, the length of the error-free interval I_0 preceding the first burst is at least $S_1 - 1$, and the length of the error-free interval I_k following the k-th burst is at least $\ell + 1 - S_k - L_k$.

Let us now consider the following realization in terms of burst lengths $\vec{l} = (l_1, \ldots, l_k)$ and starting sector positions $\vec{s} = (s_1, \ldots, s_k)$. Let us denote by $\mathcal{R}_k$ the set of all possible realizations $\{(\vec{l}, \vec{s})\}$, and by $\mathcal{E}_k$ its subset containing those realizations that lead to a segment error. Next we proceed to calculating the probability $P(\vec{L} = \vec{l}, \vec{S} = \vec{s})$. Depending on the value of s_1, two cases are considered:

Case 1) $s_1 = 1$.

As the first sector of the segment has an error, the corresponding burst may have started in the preceding segment. Therefore, the length R_1 of the remaining consecutive errors is distributed according to the residual burst size $\hat{B}$, i.e. $P(R_1 = j) = \hat{b}_j$, where $\hat{b}_j \triangleq P(\hat{B} = j) = G_j/\bar{B}$ for $j = 1, 2, \ldots$ [6]. Note that the length L_1 of consecutive errors within the segment is equal to $\min(R_1, \ell)$, and therefore its pdf is given by $P(L_1 = j) = P(R_1 = j) = \hat{b}_j$ for $j = 1, 2, \ldots, \ell - 1$, and $P(L_1 = \ell) = P(R_1 \geq \ell) = \sum_{j=\ell}^{\infty} \hat{b}_j$. Depending on whether I_k exists, two cases are considered:

Case 1.a) $\exists\ I_k$. This is equivalent to the condition $s_k + l_k \leq \ell$.

As in this case the length of the interval I_k is at least $\ell + 1 - s_k - l_k$, it holds that

$P(\vec{L} = \vec{l}, \vec{S} = \vec{s})$ =

$P(\text{first sector in error}, L_1 = l_1,\ I_1 = s_2 - s_1 - l_1,\ L_2 = l_2,\ \ldots\ , L_k = l_k\ , I_k \geq \ell + 1 - s_k - l_k)$ =

$P_s\, P(L_1 = l_1)\, P(I_1 = s_2 - s_1 - l_1)\, P(L_2 = l_2) \cdots P(L_k = l_k)\, P(I_k \geq \ell + 1 - s_k - l_k)$ =

$P_s\, \frac{G_{l_1}}{\bar{B}}\, \alpha(1-\alpha)^{s_2 - s_1 - l_1 - 1}\, b_{l_2} \cdots b_{l_k}\, (1-\alpha)^{\ell - s_k - l_k}$ =

$P_s \frac{G_{l_1}}{\bar{B}}\, b_{l_2} \cdots b_{l_k}\, (1-\alpha)^{\ell - k - (l_1 + \cdots + l_k)}\, \alpha^{k-1}$ =

$\frac{G_{l_1}\, b_{l_2} \cdots b_{l_k}}{\bar{B}^k} \left\{ P_s^k + \left[k - 1 - \frac{\ell - k - (l_1 + \cdots + l_k)}{\bar{B}} \right] P_s^{k+1} \right\} + O(P_s^{k+2})$.

Case 1.b) $\nexists\ I_k$. This is equivalent to the condition $s_k + l_k = \ell + 1$.

Depending on the value of k, two cases are considered:

Case 1.b.i) $k = 1$.

In this case it holds that $l_1 = \ell$. Thus,

$P(L_1 = \ell, S_1 = 1) = P(\text{first sector in error}, R_1 \geq \ell) = P_s\, P(R_1 \geq \ell) = \frac{\sum_{j=\ell}^{\infty} G_j}{\bar{B}}\, P_s$.

Case 1.b.ii) $k \geq 2$.

As the last sector of the segment has an error, the corresponding burst may extend into the next segment. Therefore, the pdf of the length L_k of consecutive errors within the segment is distributed according to the complementary cumulative density function of the burst size B, i.e. $P(L_k = n) = \sum_{j=n}^{\infty} b_j = G_n$ for $n = 1, 2, \ldots$. In this case it holds that $s_k + l_k = \ell + 1$. Thus,

$P(\vec{L} = \vec{l}, \vec{S} = \vec{s})$ =

$P(\text{first sector in error, } L_1 = l_1,\ I_1 = s_2 - s_1 - l_1,\ L_2 = l_2,\ \ldots\ , L_k = l_k) =$
$P_s\, P(L_1 = l_1)\, P(I_1 = s_2 - s_1 - l_1)\, P(L_2 = l_2) \cdots P(I_{k-1} = s_k - s_{k-1} - l_{k-1})\, P(L_k = l_k) =$
$P_s\, \frac{G_{l_1}}{\bar{B}}\, \alpha(1-\alpha)^{s_2 - s_1 - l_1 - 1}\, b_{l_2}\ \cdots\ \alpha(1-\alpha)^{s_k - s_{k-1} - l_{k-1} - 1}\, G_{l_k} =$
$P_s\, \frac{G_{l_1}}{\bar{B}}\, b_{l_2}\ \cdots\ b_{l_{k-1}}\, G_{l_k}\, (1-\alpha)^{\ell - (k-1) - (l_1 + \cdots + l_k)}\, \alpha^{k-1} =$
$\frac{G_{l_1}\, b_{l_2} \cdots b_{l_{k-1}}\, G_{l_k}}{\bar{B}^k} \left\{ P_s^k + \left[k - 1 - \frac{\ell + 1 - k - (l_1 + \cdots + l_k)}{\bar{B}} \right] P_s^{k+1} \right\} + O(P_s^{k+2}).$

Case 2) $s_1 \geq 2$.

Let P_{bs} be the probability that a burst of errors starts at a given sector position. This is equal to the product of the probability of the sector being in error and the probability of an erroneous sector being the first of its corresponding burst, i.e. $P_{\text{bs}} = P_s/\bar{B}$. Depending on whether I_k exists, two cases are considered:

Case 2.a) $\exists\ I_k$. This is equivalent to the condition $s_k + l_k \leq \ell$.

Similarly to Case 1.a, it holds that
$P(\vec{L} = \vec{l}, \vec{S} = \vec{s}) =$
$P(I_0 \geq s_1 - 1,\ \text{burst of errors starts at } s_1,\ L_1 = l_1,\ I_1 = s_2 - s_1 - l_1,\ \ldots\ , L_k = l_k\ , I_k \geq \ell + 1 - s_k - l_k) =$
$P(I_0 \geq s_1 - 1)\, P_{\text{bs}}\, P(L_1 = l_1)\, P(I_1 = s_2 - s_1 - l_1) \cdots P(L_k = l_k)\, P(I_k \geq \ell + 1 - s_k - l_k) =$
$(1-\alpha)^{s_1 - 2}\, \frac{P_s}{\bar{B}}\, b_{l_1}\, \alpha(1-\alpha)^{s_2 - s_1 - l_1 - 1}\ \cdots\ b_{l_k}\, (1-\alpha)^{\ell - s_k - l_k} =$
$\frac{P_s}{\bar{B}}\, b_{l_1}\ \cdots\ b_{l_k}\, (1-\alpha)^{\ell - k - (l_1 + \cdots + l_k) - 1}\, \alpha^{k-1} =$
$\frac{b_{l_1} \cdots b_{l_k}}{\bar{B}^k} \left\{ P_s^k + \left[k - 1 - \frac{\ell - 1 - k - (l_1 + \cdots + l_k)}{\bar{B}} \right] P_s^{k+1} \right\} + O(P_s^{k+2}).$

Case 2.b) $\nexists\ I_k$. This is equivalent to the condition $s_k + l_k = \ell + 1$.

Similarly to Case 1.b.ii, and for all values of k, it holds that
$P(\vec{L} = \vec{l}, \vec{S} = \vec{s}) =$
$P(I_0 \geq s_1 - 1,\ \text{burst of errors starts at } s_1,\ L_1 = l_1,\ I_1 = s_2 - s_1 - l_1,\ \ldots\ , L_k = l_k) =$
$P(I_0 \geq s_1 - 1)\, P_{\text{bs}}\, P(L_1 = l_1)\, P(I_1 = s_2 - s_1 - l_1) \cdots P(I_{k-1} = s_k - s_{k-1} - l_{k-1})\, P(L_k = l_k) =$
$(1-\alpha)^{s_1 - 2}\, \frac{P_s}{\bar{B}}\, b_{l_1}\, \alpha(1-\alpha)^{s_2 - s_1 - l_1 - 1}\ \cdots\ \alpha(1-\alpha)^{s_k - s_{k-1} - l_{k-1} - 1}\, G_{l_k} =$
$\frac{P_s}{\bar{B}}\, b_{l_1}\ \cdots\ b_{l_{k-1}}\, G_{l_k}\, (1-\alpha)^{\ell - k - (l_1 + \cdots + l_k)}\, \alpha^{k-1} =$
$\frac{b_{l_1} \cdots b_{l_{k-1}}\, G_{l_k}}{\bar{B}^k} \left\{ P_s^k + \left[k - 1 - \frac{\ell - k - (l_1 + \cdots + l_k)}{\bar{B}} \right] P_s^{k+1} \right\} + O(P_s^{k+2}).$

From the above it follows that $P(\vec{L} = \vec{l}, \vec{S} = \vec{s})$ is of order $O(P_s^k)$ because for every $(\vec{l}, \vec{s})$ it holds that $P(\vec{L} = \vec{l}, \vec{S} = \vec{s}) = \frac{F(\vec{l},\vec{s})}{\bar{B}^k}\ P_s^k + \frac{F(\vec{l},\vec{s}) H(\vec{l},\vec{s})}{\bar{B}^k}\ P_s^{k+1} + O(P_s^{k+2})$, with $F(\vec{l}, \vec{s})$ and $H(\vec{l}, \vec{s})$ being functions of $\vec{l}$, $\vec{s}$ and $\{b_j\}$. Consequently, $P_{\text{seg}}^{(k)} = \sum_{(\vec{l},\vec{s}) \in \mathcal{E}_k} P(\vec{L} = \vec{l}, \vec{S} = \vec{s}) = \frac{\sum_{(\vec{l},\vec{s}) \in \mathcal{E}_k} F(\vec{l},\vec{s})}{\bar{B}^k}\ P_s^k + O(P_s^{k+1})$. ■

APPENDIX B
REED–SOLOMON (RS) CODING SCHEME

Proof of Theorem 1.

Let us consider an arbitrary segment. For $k = 1$ and using the terminology of Appendix A, the segment is in error for all realizations (l, s) such that $l \geq m + 1$. Thus,

$$
\begin{aligned}
P_{\text{seg}}^{(1)} &= \sum_{\substack{l\geq m+1\\ 1\leq i\leq \ell}} P(L=l,S=i) = \\
&= \sum_{l=m+1}^{\ell-1} P(L=l,S=1) + P(L=\ell,S=1) + \\
&\quad + \sum_{i=2}^{\ell-m-1}\sum_{l=m+1}^{\ell-i} P(L=l,S=i) + \sum_{l=m+1}^{\ell-1} P(L=l,S=\ell+1-l)\ , \qquad (45)
\end{aligned}
$$

with the four summation terms corresponding to the Cases 1.a), 1.b.i), 2.a) and 2.b) of Appendix A, respectively. Thus,

$$
\begin{aligned}
P_{\text{seg}}^{(1)} &= \sum_{l=m+1}^{\ell-1} \frac{G_l}{\bar{B}}\,(P_s - \frac{\ell-1-l}{\bar{B}}P_s^2) + \frac{\sum_{j=\ell}^{\infty} G_j}{\bar{B}}\,P_s + \sum_{i=2}^{\ell-m-1}\sum_{l=m+1}^{\ell-i} \frac{b_l}{\bar{B}}\,(P_s - \frac{\ell-2-l}{\bar{B}}P_s^2) + \\
&\quad + \sum_{l=m+1}^{\ell-1} \frac{G_l}{\bar{B}}\,(P_s - \frac{\ell-1-l}{\bar{B}}P_s^2) + O(P_s^3) = \\
&= \left(\sum_{l=m+1}^{\infty} G_l + \sum_{i=2}^{\ell-m-1}\sum_{l=m+1}^{\ell-i} b_l + \sum_{i=2}^{\ell-m} G_{\ell+1-i}\right)\frac{P_s}{\bar{B}} - \\
&\quad - \left[2\sum_{l=m+1}^{\ell-1}(\ell-1-l)G_l + \sum_{i=2}^{\ell-m-1}\sum_{l=m+1}^{\ell-i}(\ell-2-l)b_l\right]\frac{P_s^2}{\bar{B}^2} + O(P_s^3) = \\
&= \left(\sum_{l=1}^{\infty} G_l - \sum_{l=1}^{m} G_l + \sum_{l=m+1}^{\ell-2}\sum_{i=2}^{\ell-l} b_l + \sum_{i=m+1}^{\ell-1} G_i\right)\frac{P_s}{\bar{B}} - \\
&\quad - \left[2\sum_{l=m+1}^{\ell-1}(\ell-1-l)G_l + \sum_{l=m+1}^{\ell-2}\sum_{i=2}^{\ell-l}(\ell-2-l)b_l\right]\frac{P_s^2}{\bar{B}^2} + O(P_s^3) = \\
&= \left[\bar{B} - \sum_{l=1}^{m} G_l + \sum_{l=m+1}^{\ell-2}(\ell-1-l)b_l + \sum_{i=m+1}^{\ell-1} G_i\right]\frac{P_s}{\bar{B}} - \\
&\quad - \left[2\sum_{l=m+1}^{\ell-1}(\ell-1-l)G_l + \sum_{l=m+1}^{\ell-2}(\ell-1-l)(\ell-2-l)b_l\right]\frac{P_s^2}{\bar{B}^2} + O(P_s^3) = \\
&= \left[\bar{B} - \sum_{l=1}^{m} G_l + (\ell-m-1)G_{m+1} - \sum_{i=m+1}^{\ell-1} G_i + \sum_{i=m+1}^{\ell-1} G_i\right]\frac{P_s}{\bar{B}} - \\
&\quad - \left[(\ell-m-1)(\ell-m-2)\,G_{m+1}\right]\frac{P_s^2}{\bar{B}^2} + O(P_s^3) = \\
&= \left[1 + \frac{(\ell-m-1)G_{m+1} - \sum_{j=1}^{m} G_j}{\bar{B}}\right]P_s - \left[2\binom{\ell-m-1}{2}G_{m+1}\right]\frac{P_s^2}{\bar{B}^2} + O(P_s^3)\ .
\end{aligned}
\qquad (46)
$$

For $k = 2$, and using the terminology used in Appendix A, the segment is in error for all realizations

(l_1, l_2) such that $l_1 + l_2 \geq m + 1$. Thus,

$$
\begin{aligned}
P_{\text{seg}}^{(2)} &= \sum_{\substack{(\vec{l},\vec{s})\in\mathcal{R}_2 \\ l_1+l_2\geq m+1}} P(L=\vec{l}, S=\vec{s}) = \\
&= \sum_{\substack{(l_1,l_2) \\ m+1\leq l_1+l_2\leq \ell-2}} \sum_{s_2=l_1+2}^{\ell-l_2} P(\vec{L}=(l_1,l_2), \vec{S}=(1,s_2)) + \\
&\quad + \sum_{\substack{(l_1,l_2) \\ m+1\leq l_1+l_2\leq \ell-1}} P(\vec{L}=(l_1,l_2), \vec{S}=(1,\ell+1-l_2)) + \\
&\quad + \sum_{\substack{(l_1,l_2) \\ m+1\leq l_1+l_2\leq \ell-3}} \sum_{s_1=2}^{\ell-1-(l_1+l_2)} \sum_{s_2=s_1+l_1+1}^{\ell-l_2} P(\vec{L}=(l_1,l_2), \vec{S}=(s_1,s_2)) + \\
&\quad + \sum_{\substack{(l_1,l_2) \\ m+1\leq l_1+l_2\leq \ell-2}} \sum_{s_1=2}^{\ell-(l_1+l_2)} P(\vec{L}=(l_1,l_2), \vec{S}=(s_1,\ell+1-l_2)) \ ,
\end{aligned}
$$

with the four summation terms corresponding to the Cases 1.a), 1.b.ii), 2.a) and 2.b) of Appendix A, respectively. Thus,

$$
\begin{aligned}
P_{\text{seg}}^{(2)} &= \sum_{m+1\leq l_1+l_2\leq \ell-2} \sum_{s_2=l_1+2}^{\ell-l_2} \frac{G_{l_1}\, b_{l_2}}{\bar{B}^2}\, P_s^2 \ + \sum_{m+1\leq l_1+l_2\leq \ell-1} \frac{G_{l_1}\, G_{l_2}}{\bar{B}^2}\, P_s^2 \ + \\
&\quad + \sum_{m+1\leq l_1+l_2\leq \ell-3} \sum_{s_1=2}^{\ell-1-(l_1+l_2)} \sum_{s_2=s_1+l_1+1}^{\ell-l_2} \frac{b_{l_1}\, b_{l_2}}{\bar{B}^2}\, P_s^2 \ + \\
&\quad + \sum_{m+1\leq l_1+l_2\leq \ell-2} \sum_{s_1=2}^{\ell-(l_1+l_2)} \frac{b_{l_1}\, G_{l_2}}{\bar{B}^2}\, P_s^2 \ + O(P_s^3) = \\
&= \left\{ \sum_{m+1\leq l_1+l_2\leq \ell-2} [\ell-1-(l_1+l_2)]\, G_{l_1} b_{l_2} + \right. \\
&\quad + \sum_{m+1\leq l_1+l_2\leq \ell-1} G_{l_1}\, G_{l_2} \ + \\
&\quad + \sum_{m+1\leq l_1+l_2\leq \ell-3} \sum_{s_1=2}^{\ell-1-(l_1+l_2)} [\ell-(l_1+l_2)-s_1]\, b_{l_1} b_{l_2} + \\
&\quad \left. + \sum_{m+1\leq l_1+l_2\leq \ell-2} [\ell-1-(l_1+l_2)]\, b_{l_1} G_{l_2} \right\} \frac{P_s^2}{\bar{B}^2} + O(P_s^3) \ ,
\end{aligned}
\tag{47}
$$

or

$$P_{\text{seg}}^{(2)} = \left\{ \sum_{j=m+1}^{\ell-2} (\ell-1-j)\Big(\sum_{l_1+l_2=j} G_{l_1} b_{l_2}\Big) + \right.$$
$$+ \sum_{j=m+1}^{\ell-1} \Big(\sum_{l_1+l_2=j} G_{l_1}\, G_{l_2}\Big) +$$
$$+ \sum_{j=m+1}^{\ell-3} \sum_{s_1=2}^{\ell-1-j} (\ell-j-s_1)\Big(\sum_{l_1+l_2=j} b_{l_1} b_{l_2}\Big) +$$
$$\left. + \sum_{j=m+1}^{\ell-2} (\ell-1-j)\Big(\sum_{l_1+l_2=j} b_{l_1} G_{l_2}\Big) \right\} \frac{P_s^2}{\bar{B}^2} + O(P_s^3)\ . \tag{48}$$

Introducing the following notation:

$$BB_j \triangleq \sum_{\substack{(l_1,l_2)\\ l_1+l_2=j}} b_{l_1} b_{l_2}\ , \qquad GB_j \triangleq \sum_{\substack{(l_1,l_2)\\ l_1+l_2=j}} G_{l_1} b_{l_2}\ , \qquad GG_j \triangleq \sum_{\substack{(l_1,l_2)\\ l_1+l_2=j}} G_{l_1} G_{l_2}\ , \tag{49}$$

Eq. (48) can be written as follows:

$$P_{\text{seg}}^{(2)} = \left\{ 2 \sum_{j=m+1}^{\ell-2} (\ell-1-j)\, GB_j + \sum_{j=m+1}^{\ell-1} GG_j + \sum_{j=m+1}^{\ell-3} \binom{\ell-1-j}{2} BB_j \right\} \frac{P_s^2}{\bar{B}^2} +$$
$$+ O(P_s^3)\ . \tag{50}$$

Next we express GB_j and BB_j as functions of G_j:

$$GB_j = \sum_{\substack{(l_1,l_2)\\ l_1+l_2=j}} G_{l_1} b_{l_2} = \sum_{\substack{(l_1,l_2)\\ l_1+l_2=j}} G_{l_1}(G_{l_2} - G_{l_2+1}) = \sum_{\substack{(l_1,l_2)\\ l_1+l_2=j}} G_{l_1} G_{l_2} - \sum_{\substack{(l_1,l_2)\\ l_1+l_2=j}} G_{l_1} G_{l_2+1} =$$
$$= GG_j - [\sum_{\substack{(l_1,l_2')\\ l_1+l_2'=j+1}} G_{l_1} G_{l_2'} - G_j G_1] = GG_j - GG_{j+1} + G_j\ . \tag{51}$$

$$BB_j = \sum_{\substack{(l_1,l_2)\\ l_1+l_2=j}} b_{l_1} b_{l_2} = \sum_{\substack{(l_1,l_2)\\ l_1+l_2=j}} (G_{l_1} - G_{l_1+1}) b_{l_2} = \sum_{\substack{(l_1,l_2)\\ l_1+l_2=j}} G_{l_1} b_{l_2} - \sum_{\substack{(l_1,l_2)\\ l_1+l_2=j}} G_{l_1+1} b_{l_2} =$$
$$= GB_j - [\sum_{\substack{(l_1',l_2)\\ l_1'+l_2=j+1}} G_{l_1'} b_{l_2} - G_1 b_j] = GB_j - GB_{j+1} + b_j =$$
$$\overset{(51)}{=} (GG_j - GG_{j+1} + G_j) - (GG_{j+1} - GG_{j+2} + G_{j+1}) + (G_j - G_{j+1}) =$$
$$= GG_j - 2\, GG_{j+1} + GG_{j+2} + 2(G_j - G_{j+1})\ . \tag{52}$$

Substituting (51) and (52) into (50), yields the following:

$$
\begin{aligned}
P_{\text{seg}}^{(2)} &= \left\{ 2 \sum_{j=m+1}^{\ell-2} (\ell-1-j)\,(GG_j - GG_{j+1} + G_j) \;+\; \sum_{j=m+1}^{\ell-1} GG_j \;+ \right. \\
&\quad \left. + \sum_{j=m+1}^{\ell-3} \binom{\ell-1-j}{2} (GG_j - 2\,GG_{j+1} + GG_{j+2} + 2G_j - 2G_{j+1}) \right\} \frac{P_s^2}{\bar{B}^2} + O(P_s^3) = \\
&= \left\{ \sum_{j=m+1}^{\ell-3} \left[2(\ell-1-j) + 1 + \binom{\ell-1-j}{2} \right] GG_j + 3GG_{\ell-2} + GG_{\ell-1} - \right. \\
&\quad - 2 \sum_{j=m+1}^{\ell-3} \left[(\ell-1-j) + \binom{\ell-1-j}{2} \right] GG_{j+1} \;-\; 2\,GG_{\ell-1} \;+ \\
&\quad + \sum_{j=m+1}^{\ell-3} \binom{\ell-1-j}{2} \; GG_{j+2} \;+ \\
&\quad + 2 \sum_{j=m+1}^{\ell-3} \left[(\ell-1-j) + \binom{\ell-1-j}{2} \right] G_j \;+\; 2G_{\ell-2} \;- \\
&\quad \left. - 2 \sum_{j=m+1}^{\ell-3} \binom{\ell-1-j}{2} \; G_{j+1} \right\} \frac{P_s^2}{\bar{B}^2} + O(P_s^3) = \\
&= \left\{ \sum_{j=m+1}^{\ell-3} \binom{\ell+1-j}{2} \; GG_j + 3GG_{\ell-2} - GG_{\ell-1} - \right. \\
&\quad - 2 \sum_{i=m+2}^{\ell-2} \binom{\ell+1-i}{2} \; GG_i \;+ \sum_{i=m+3}^{\ell-1} \binom{\ell+1-i}{2} \; GG_i \;+ \\
&\quad \left. + \;2 \sum_{j=m+1}^{\ell-3} \binom{\ell-j}{2} \; G_j \;+\; 2G_{\ell-2} \;-\; 2 \sum_{i=m+2}^{\ell-2} \binom{\ell-i}{2} \; G_i \right\} \frac{P_s^2}{\bar{B}^2} + O(P_s^3) = \\
&= \left\{ \binom{\ell-m}{2} \; GG_{m+1} - \binom{\ell-m-1}{2} \; GG_{m+2} \;+\; 2 \binom{\ell-m-1}{2} \; G_{m+1} \right\} \frac{P_s^2}{\bar{B}^2} + \\
&\quad + O(P_s^3)\,. \qquad (53)
\end{aligned}
$$

Owing to Proposition 2, and by making use of (46) and (53), we get (21) – (23). ■

Proof of Corollary 2.

First we show that condition $G_{\lceil \frac{m+1}{2} \rceil} = 0$ is necessary. From (23), and by making use of (51), it follows that

$$
c_2^{\text{RS}} = \left\{ (\ell-m-1)\,GG_{m+1} + \binom{\ell-m-1}{2} \; GB_{m+1} \;-\; \binom{\ell-m-1}{2} \; G_{m+1} \right\} \frac{1}{\bar{B}^2}\,. \qquad (54)
$$

As the coefficient c_1^{RS} is equal to zero, from the above and Corollary 1 it follows that the coefficient c_2^{RS}

is equal to zero if and only if $GG_{m+1} = GB_{m+1} = 0$. As $0 \leq b_j \leq G_j\ ,\ \forall\, j \in \mathbb{N}$, from definitions (49) it follows that $0 \leq GB_{m+1} \leq GG_{m+1}$. Consequently, $GG_{m+1} = 0$ implies $GB_{m+1} = 0$ and therefore $c_2^{\text{RS}} = 0$ if and only if $GG_{m+1} = 0$. Moreover, from the definition (24) it follows that $GG_{m+1} = 0$ if and only if $G_{\lceil \frac{m+1}{2} \rceil} = 0$.

Next we show that condition $G_{\lceil \frac{m+1}{2} \rceil} = 0$ is sufficient. This condition implies that $G_{m+1} = 0$, which in turn by virtue of Corollary 1 implies that $c_1^{\text{RS}} = 0$. From the above it also follows that condition $G_{\lceil \frac{m+1}{2} \rceil} = 0$ implies that $c_2^{\text{RS}} = 0$. ■

APPENDIX C
SPC CODING SCHEME

Proof of Theorem 2.

Let us consider an arbitrary segment. According to the SPC coding scheme, the segment is not in error if either there are no sectors in error or if there is one sector in error, i.e.

$$1 - P_{\text{seg}}^{\text{SPC}} \ = \ P_{\text{no}} + P_{\text{single}}\ , \tag{55}$$

where P_{no} and P_{single} denote the probabilities of the former and the latter event, respectively. From (18) it follows that

$$P_{\text{no}} \ = \ (1 - P_s)(1 - \alpha)^{\ell - 1}\ . \tag{56}$$

For $k = 1$, and using the terminology used in Appendix A, the segment contains a single sector in error for all realizations (l, s) such that $l = 1$. Therefore,

$$\begin{aligned} P_{\text{single}} \ &= \ \sum_{1 \leq i \leq \ell} P(L = 1, S = i) = \\ &= \ P(L = 1, S = 1) + \sum_{i=2}^{\ell - 1} P(L = 1, S = i) + P(L = 1, S = \ell)\ , \end{aligned} \tag{57}$$

with the three summation terms corresponding to the Cases 1.a), 2.a) and 2.b) of Appendix A, respectively. Thus,

$$\begin{aligned} P_{\text{single}} \ &= \ P_s \frac{G_1}{\bar{B}}\ (1 - \alpha)^{\ell - 2} \ + \sum_{i=2}^{\ell - 1} \frac{P_s}{\bar{B}}\ b_1\ (1 - \alpha)^{\ell - 3} \ + \frac{P_s}{\bar{B}}\ G_1\ (1 - \alpha)^{\ell - 2} = \\ &= \ \frac{P_s}{\bar{B}} \left[2G_1 (1 - \alpha)^{\ell - 2} + \sum_{i=2}^{\ell - 1} b_1\ (1 - \alpha)^{\ell - 3} \right] = \frac{P_s}{\bar{B}}\ (1 - \alpha)^{\ell - 3}\ [2(1 - \alpha) + (\ell - 2) b_1]\ . \end{aligned} \tag{58}$$

Equation (26) follows immediately from (55), (56) and (58). ■

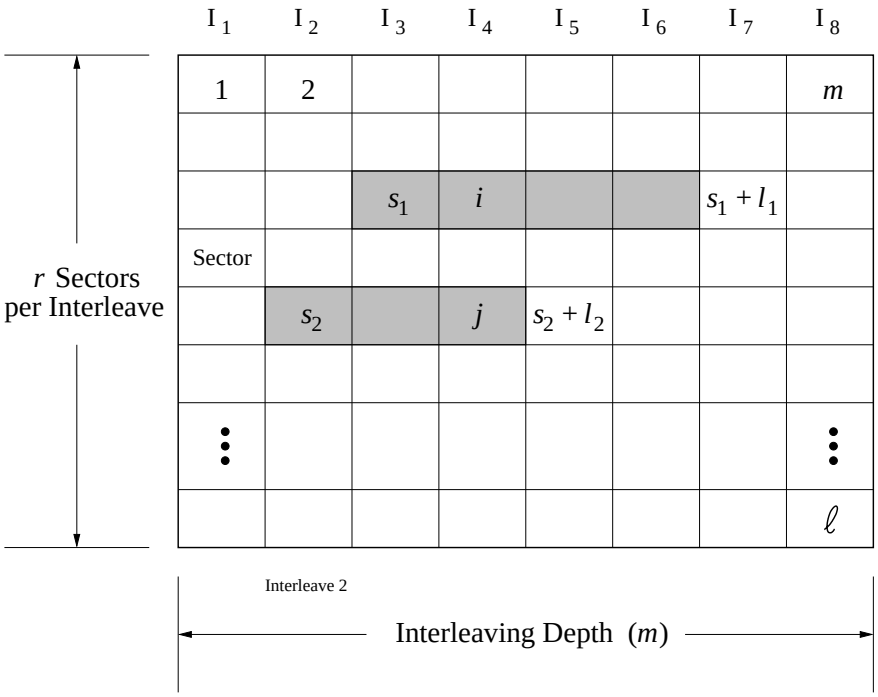


Fig. 8. Segment in error under IPC coding scheme.

APPENDIX D
IPC CODING SCHEME

Proof of Theorem 4.

According to the IPC coding scheme, a segment is in error if there is at least one interleave in which there are at least two unrecoverable sector errors. In the case of a single burst, this occurs when the burst length exceeds the interleaving depth. Consequently, the probability of the segment being in error is the same as in the case of the RS coding scheme given by (46). In the case of two bursts, this occurs if the sum of the two burst lengths exceeds m, but also occurs if the sum of the two burst lengths is less or equal to m and the bursts are positioned in a way such that there is at least one interleave in which there are two unrecoverable sector errors. Let $P_{\text{seg}}^{(2,a)}$ and $P_{\text{seg}}^{(2,b)}$ denote the probabilities of these two events, respectively, such that

$$P_{\text{seg}}^{(2)} = P_{\text{seg}}^{(2,a)} + P_{\text{seg}}^{(2,b)} \ . \tag{59}$$

The probability $P_{\text{seg}}^{(2,a)}$ of the former event is equal to the one derived in (53), whereas the probability $P_{\text{seg}}^{(2,b)}$ of the latter event will be evaluated next.

For $k = 2$, and using the terminology of Appendix A, the segment is in error for all realizations (l_1, l_2, s_1, s_2) with $l_1 + l_2 \leq m$, and s_1, s_2 such that there are sectors i and j, with $s_1 \leq i \leq s_1 + l_1 - 1$, $s_2 \leq j \leq s_2 + l_2 - 1$ and $i \overset{m}{=} j$. The symbol $\overset{m}{=}$ has the following meaning:

$$a \overset{m}{=} b \ \Leftrightarrow \ (a \mod m) = (b \mod m) \ \Leftrightarrow \ a - b = nm \ , \quad \text{with } n \in \mathbb{Z} \text{ and } m \in \mathbb{N} \ . \tag{60}$$

A typical such realization is shown in Fig. 8 with $m = 8$, $l_1 = 4$, $l_2 = 3$ and the segment being in error because each of interleaves I_3 and I_4 contain two sector errors. Also, r denotes the number of sectors per

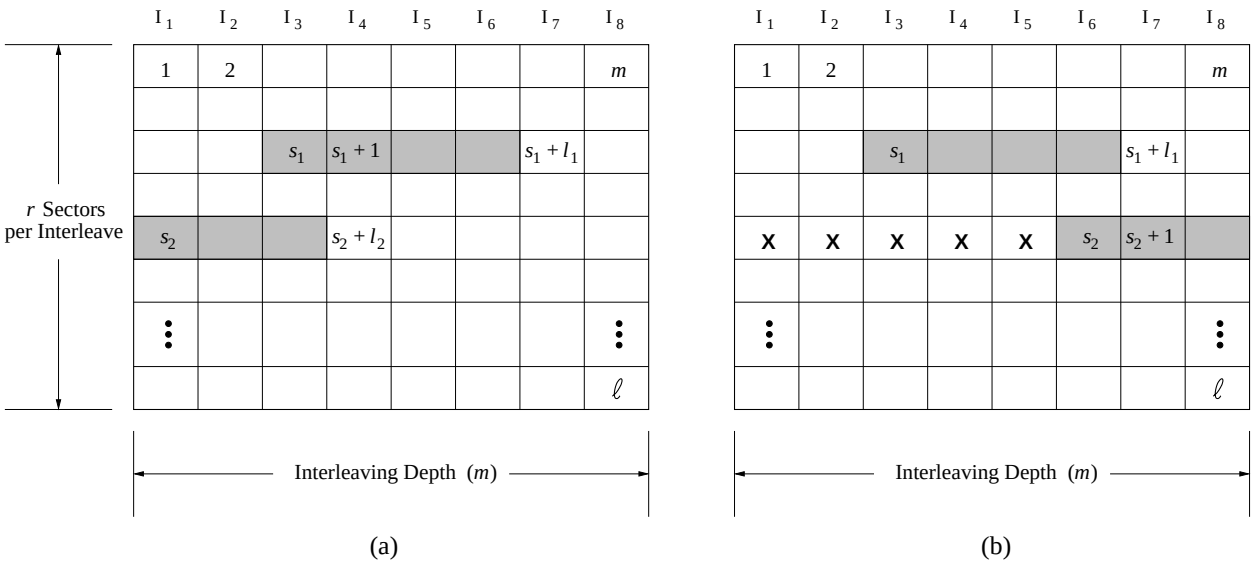


Fig. 9. Sequence of positions of second burst leading to a segment error.

interleave, i.e.

$$r \triangleq \frac{\ell}{m} . \tag{61}$$

Let us now examine the positions of the second burst that lead to a segment error, given its burst length l_2 and for a given first burst, i.e. for fixed s_1 and l_1. As demonstrated in Fig. 9a, when $s_2 + l_2 \overset{m}{=} s_1 + 1$ there is a single interleave (I_3) containing two sectors in error, namely, the first sector of the first burst and the last sector of the second burst. When the second burst is shifted forward, there are multiple interleaves with two sector errors until the position shown in Fig. 9b, where $s_2 + 1 \overset{m}{=} s_1 + l_1$ and a single interleave (I_6) containing two sectors in error, namely, the last sector of the first burst and the first sector of the second burst. Therefore, the sequences of positions s_2 that lead to a segment error are the intervals

$$s_1 + 1 - l_2 + nm \leq s_2 \leq s_1 + l_1 - 1 + nm , \qquad \text{with } n \in \mathbb{N} . \tag{62}$$

We now proceed by considering the various positions of the two bursts leading to a segment error, and in particular the following cases:

Case 1: $s_1 = 1$.

According to (62), the sequences of positions s_2 are the following: $2 - l_2 + nm \leq s_2 \leq l_1 + nm$, for $n = 1, 2, \ldots, r-1$, as shown in Fig. 10b. Note that $s_2 + l_2 \leq l_1 + l_2 + (r-1)m \leq m + (r-1)m = \ell$, which implies that there is always an interval I_2. Thus, unconditioning on lengths l_1 and l_2, and using Case 1.a of Appendix A and (48) we get

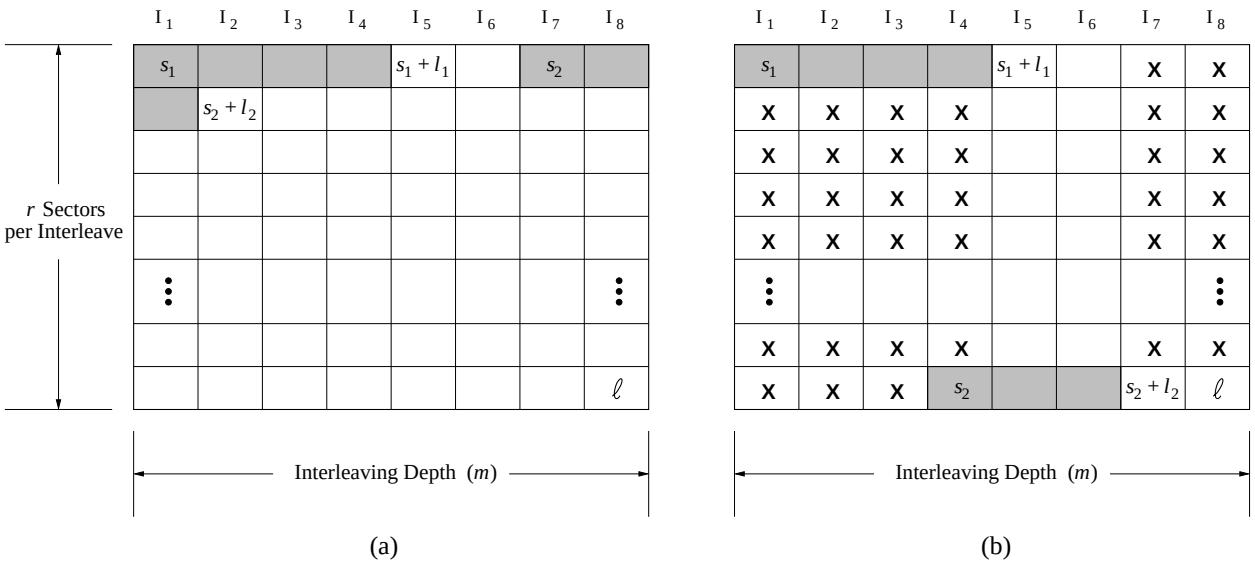


Fig. 10. Position of second burst leading to a segment error ($s_1 = 1$).

$$P_{\text{seg}}^{(2,b.1)} = \sum_{\substack{(l_1,l_2)\\ 2\le l_1+l_2\le m}} \sum_{n=1}^{r-1} \sum_{s_2=nm+2-l_2}^{nm+l_1} P(\vec{L}=(l_1,l_2),\vec{S}=(1,s_2)) =$$

$$= \sum_{\substack{(l_1,l_2)\\ 2\le l_1+l_2\le m}} \sum_{n=1}^{r-1} \sum_{s_2=nm+2-l_2}^{nm+l_1} \frac{G_{l_1} b_{l_2}}{\bar{B}^2}\ P_s^2\ + O(P_s^3)\ , \tag{63}$$

or

$$P_{\text{seg}}^{(2,b.1)} = \sum_{\substack{(l_1,l_2)\\ 2\le l_1+l_2\le m}} \sum_{n=1}^{r-1} (l_1+l_2-1)\ G_{l_1} b_{l_2} \frac{P_s^2}{\bar{B}^2}\ + O(P_s^3)\ =$$

$$= \sum_{j=2}^{m} \sum_{\substack{(l_1,l_2)\\ l_1+l_2=j}} (r-1)\,(j-1)\,G_{l_1} b_{l_2}\ \frac{P_s^2}{\bar{B}^2}\ + O(P_s^3)\ =$$

$$= (r-1)\left[\sum_{j=2}^{m} (j-1)\ \left(\sum_{\substack{(l_1,l_2)\\ l_1+l_2=j}} G_{l_1} b_{l_2}\right)\right] \frac{P_s^2}{\bar{B}^2}\ + O(P_s^3)\ =$$

$$= (r-1)\left[\sum_{j=2}^{m} (j-1)\ GB_j\right] \frac{P_s^2}{\bar{B}^2}\ + O(P_s^3)\ . \tag{64}$$

Case 2: $s_2 + l_2 = \ell + 1$.

In this case the second burst runs until the end of the segment. As shown in Fig. 11b, for a given l_1 and l_2, the sequence of positions s_1 that lead to a segment error are the following: $nm + 2 - l_1 - l_2 \le s_1 \le nm$, for $n = 1, 2, \ldots, r-1$. Thus, unconditioning on lengths l_1 and l_2, and using Case 2.b of Appendix A (as $s_1 \ge 2$ and $\nexists\ I_2$) we get

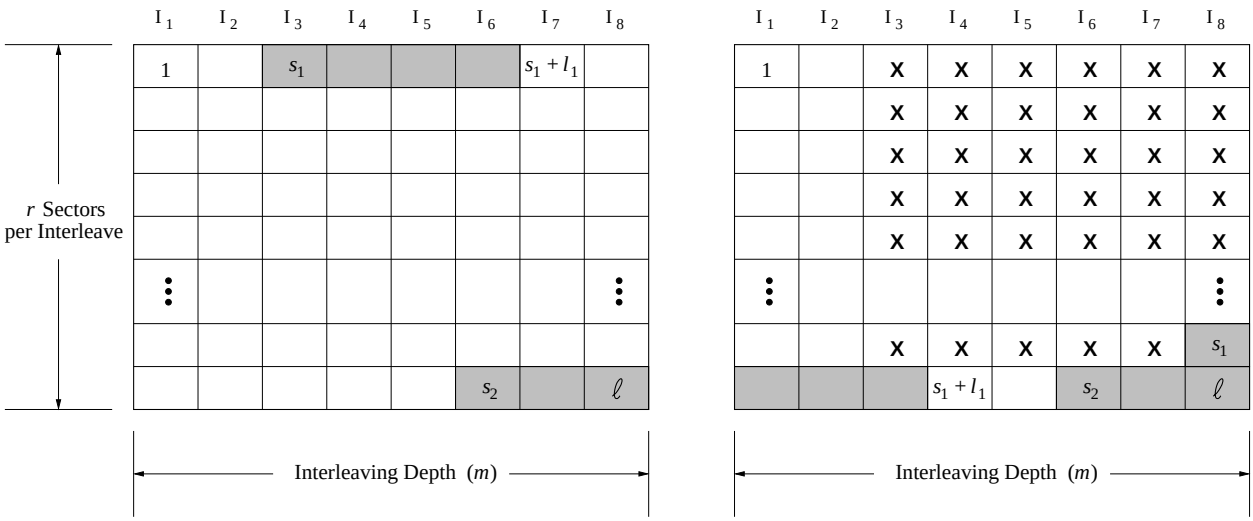


Fig. 11. Position of first burst leading to a segment error ($s_2 + l_2 = \ell + 1$).

$$
\begin{aligned}
P_{\text{seg}}^{(2,b.2)} &= \sum_{\substack{(l_1,l_2)\\ 2\le l_1+l_2\le m}} \sum_{n=1}^{r-1} \sum_{s_1=nm+2-l_1-l_2}^{nm} P(\vec{L}=(l_1,l_2),\vec{S}=(s_1,\ell+1-l_2)) = \\
&= \sum_{\substack{(l_1,l_2)\\ 2\le l_1+l_2\le m}} \sum_{n=1}^{r-1} \sum_{s_1=nm+2-l_1-l_2}^{nm} \frac{b_{l_1} G_{l_2}}{\bar{B}^2}\ P_s^2 + O(P_s^3) = \\
&= \sum_{\substack{(l_1,l_2)\\ 2\le l_1+l_2\le m}} \sum_{n=1}^{r-1} (l_1+l_2-1)\ b_{l_1} G_{l_2} \frac{P_s^2}{\bar{B}^2} + O(P_s^3) = \\
&= \sum_{j=2}^{m} \sum_{\substack{(l_1,l_2)\\ l_1+l_2=j}} (r-1)\,(j-1)\, b_{l_1} G_{l_2}\ \frac{P_s^2}{\bar{B}^2} + O(P_s^3) = \\
&= (r-1) \left[\sum_{j=2}^{m} (j-1) \left(\sum_{\substack{(l_1,l_2)\\ l_1+l_2=j}} b_{l_1} G_{l_2} \right) \right] \frac{P_s^2}{\bar{B}^2} + O(P_s^3) = \\
&= (r-1) \left[\sum_{j=2}^{m} (j-1)\ GB_j \right] \frac{P_s^2}{\bar{B}^2} + O(P_s^3)\ . \qquad (65)
\end{aligned}
$$

Note that this case is the symmetric of case 1 because it holds that $P(\vec{L}=(l_1,l_2),\vec{S}=(s_1,\ell+1-l_2)) = P(\vec{L}=(l_2,l_1),\vec{S}=(1,\ell+2-s_1-l_1))$. Consequently, $P_{\text{seg}}^{(2,b.1)} = P_{\text{seg}}^{(2,b.2)}$.

Case 3: $s_1 \geq 2$ and $s_2 + l_2 \leq \ell$.

This corresponds to Case 2.a of Appendix A. As shown in Fig. 12, the boundary positions for the first burst are $s_1 = 2$ and $s_1 = (r-2)m + m - 1$, which implies that s_1 can be represented as follows:

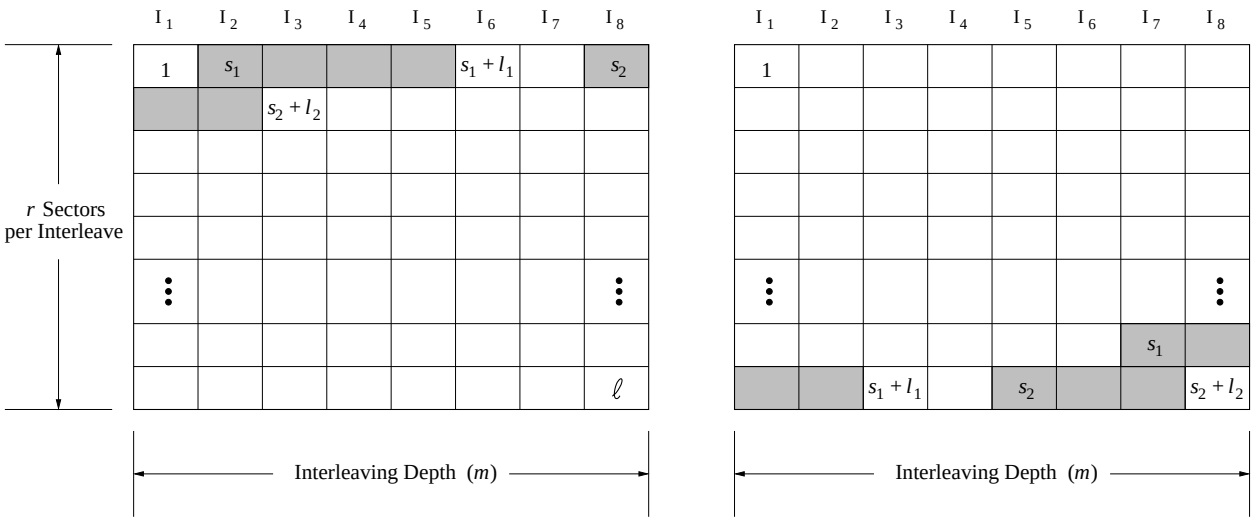


Fig. 12. Boundary positions of first burst leading to a segment error ($s_1 \geq 2$ and $s_2 + l_2 \leq \ell$).

$$s_1 = \begin{cases} i\,, & 2 \leq i \leq m-1\,, \quad \text{for } q = 0\,, \\ qm + i\,, & 0 \leq i \leq m-1\,, \quad \text{for } q = 1, \ldots, r-2\,. \end{cases} \tag{66}$$

Let us now derive the sequence of positions s_2 for the second burst that lead to a segment error. From (61), (62), (66), and owing to the restriction $s_2 + l_2 \leq \ell$, it follows that $qm + i + 1 + nm = s_1 + 1 + nm \leq s_2 + l_2 \leq \ell = rm$, or $nm \leq (r-q)m - 1 - i$. This, together with the fact that $(r-q-1)m \leq (r-q)m - 1 - i < (r-q)m$, implies that $n \leq r - q - 1$. On the other hand, owing to the restriction $s_2 + l_2 \leq \ell$, or $s_2 \leq \ell - l_2$, the upper bound for the sequences of positions s_2 given in (62) should not exceed $\ell - l_2$, i.e. $s_1 + l_1 - 1 + nm \leq \ell - l_2 \;\Leftrightarrow\; nm \leq \ell - s_1 - l_1 - l_2 + 1 = rm - qm - i - (l_1 + l_2) + 1 \;\Leftrightarrow\; nm \leq (r-q)m - i - (l_1 + l_2) + 1$. Note that this inequality holds for all n with $n \leq r - q - 2$. However, for $n = r - q - 1$, this inequality holds only if $(r-q-1)m \leq (r-q)m - i - (l_1 + l_2) + 1 \;\Leftrightarrow\; i \leq m + 1 - (l_1 + l_2)$. From the above it follows that the sequence of positions s_2 for the second burst that lead to a segment error are given by

$$\begin{aligned} &\{qm + i + 1 - l_2 + nm \leq s_2 \leq qm + i + l_1 - 1 + nm\,, \quad \text{with } n \in \{1, 2, \ldots, r-q-1\}\}\,, \\ &\quad \text{for } 0 \leq i \leq m + 1 - (l_1 + l_2)\,, \end{aligned} \tag{67}$$

and

$$\begin{aligned} &\left\{ \begin{array}{l} qm + i + 1 - l_2 + nm \leq s_2 \leq qm + i + l_1 - 1 + nm\,, \quad \text{with } n \in \{1, 2, \ldots, r-q-2\} \\ (r-1)m + i + 1 - l_2 \leq s_2 \leq rm - l_2\,, \quad (n = r - q - 1) \end{array} \right\} \\ &\quad \text{for } m + 2 - (l_1 + l_2) \leq i \leq m - 1\,. \end{aligned} \tag{68}$$

Unconditioning on burst lengths l_1 and l_2, and noting that condition (68) applies when $l_1 + l_2 \geq 3$ and that $m + 2 - (l_1 + l_2) \geq 2$, we get

$$
\begin{aligned}
P_{\text{seg}}^{(2,b.3)} &= \sum_{(l_1,l_2)} \sum_{s_1} \sum_{s_2} P(\vec{L} = (l_1, l_2), \vec{S} = (s_1, s_2)) = \\
&= \sum_{i=2}^{m-1} \sum_{n=1}^{r-1} P(\vec{L} = (1,1), \vec{S} = (i, i + nm)) + \\
&\quad + \sum_{q=1}^{r-2} \sum_{i=0}^{m-1} \sum_{n=1}^{r-q-1} P(\vec{L} = (1,1), \vec{S} = (qm + i, qm + i + nm)) + \\
&\quad + \sum_{\substack{(l_1,l_2) \\ 3 \leq l_1 + l_2 \leq m}} \\
&\quad \left\{ \sum_{i=2}^{m+1-(l_1+l_2)} \sum_{n=1}^{r-1} \sum_{s2=qm+i+1-l_2+nm}^{qm+i+l_1-1+nm} P(\vec{L} = (l_1, l_2), \vec{S} = (i, s_2)) + \right. \\
&\quad + \sum_{q=1}^{r-2} \sum_{i=0}^{m+1-(l_1+l_2)} \sum_{n=1}^{r-q-1} \sum_{s2=qm+i+1-l_2+nm}^{qm+i+l_1-1+nm} P(\vec{L} = (l_1, l_2), \vec{S} = (qm + i, s_2)) + \\
&\quad + \sum_{q=0}^{r-2} \sum_{i=m+2-(l_1+l_2)}^{m-1} \sum_{n=1}^{r-q-2} \sum_{s2=(r-1)m+i+1-l_2}^{(r-1)m+i+l_1-1} P(\vec{L} = (l_1, l_2), \vec{S} = (qm + i, s_2)) + \\
&\quad \left. + \sum_{q=0}^{r-2} \sum_{i=m+2-(l_1+l_2)}^{m-1} \sum_{s2=(r-1)m+i+1-l_2}^{rm-l_2} P(\vec{L} = (l_1, l_2), \vec{S} = (qm + i, s_2)) \right\} = \\
&= \sum_{i=2}^{m-1} \sum_{n=1}^{r-1} \frac{b_1^2}{\bar{B}^2} P_s^2 + \sum_{q=1}^{r-2} \sum_{i=0}^{m-1} \sum_{n=1}^{r-q-1} \frac{b_1^2}{\bar{B}^2} P_s^2 + \\
&\quad + \sum_{\substack{(l_1,l_2) \\ 3 \leq l_1 + l_2 \leq m}} \left\{ \sum_{i=2}^{m+1-(l_1+l_2)} \sum_{n=1}^{r-1} \sum_{s2=qm+i+1-l_2+nm}^{qm+i+l_1-1+nm} \frac{b_{l_1} b_{l_2}}{\bar{B}^2} P_s^2 + \right. \\
&\qquad + \sum_{q=1}^{r-2} \sum_{i=0}^{m+1-(l_1+l_2)} \sum_{n=1}^{r-q-1} \sum_{s2=qm+i+1-l_2+nm}^{qm+i+l_1-1+nm} \frac{b_{l_1} b_{l_2}}{\bar{B}^2} P_s^2 + \\
&\qquad + \sum_{q=0}^{r-2} \sum_{i=m+2-(l_1+l_2)}^{m-1} \sum_{n=1}^{r-q-2} \sum_{s2=(r-1)m+i+1-l_2}^{(r-1)m+i+l_1-1} \frac{b_{l_1} b_{l_2}}{\bar{B}^2} P_s^2 + \\
&\qquad \left. + \sum_{q=0}^{r-2} \sum_{i=m+2-(l_1+l_2)}^{m-1} \sum_{s2=(r-1)m+i+1-l_2}^{rm-l_2} \frac{b_{l_1} b_{l_2}}{\bar{B}^2} P_s^2 \right\} + O(P_s^3) \, .
\end{aligned}
\tag{69}
$$

Thus,

$$
\begin{aligned}
P_{\text{seg}}^{(2,b.3)} &= \left[(m-2)(r-1)+m\sum_{q=1}^{r-2}(r-q-1)\right]\frac{b_1^2}{\bar{B}^2}\,P_s^2 + \\
&\quad +\sum_{j=3}^{m}\sum_{\substack{(l_1,l_2)\\ l_1+l_2=j}}\left\{\sum_{i=2}^{m+1-(l_1+l_2)}\sum_{n=1}^{r-1}(l_1+l_2-1)+\right. \\
&\qquad +\sum_{q=1}^{r-2}\sum_{i=0}^{m+1-(l_1+l_2)}\sum_{n=1}^{r-q-1}(l_1+l_2-1)+ \\
&\qquad +\sum_{q=0}^{r-2}\sum_{i=m+2-(l_1+l_2)}^{m-1}\sum_{n=1}^{r-q-2}(l_1+l_2-1)+ \\
&\qquad \left.+\sum_{q=0}^{r-2}\sum_{i=m+2-(l_1+l_2)}^{m-1}(m-i)\right\}\frac{b_{l_1}b_{l_2}}{\bar{B}^2}\,P_s^2\;+O(P_s^3)\;= \\
&= \left[(m-2)(r-1)+m\frac{(r-1)(r-2)}{2}\right]\frac{b_1^2}{\bar{B}^2}\,P_s^2 + \\
&\quad +\sum_{j=3}^{m}\sum_{\substack{(l_1,l_2)\\ l_1+l_2=j}}\left\{(m-j)(r-1)(j-1)+\sum_{q=1}^{r-2}(m+2-j)(r-q-1)(j-1)+\right. \\
&\qquad \left.+\sum_{q=0}^{r-2}(j-2)(r-q-2)(j-1)\;+\sum_{q=0}^{r-2}\frac{(j-1)(j-2)}{2}\right\}\frac{b_{l_1}b_{l_2}}{\bar{B}^2}\,P_s^2\;+O(P_s^3)\;= \\
&= \frac{(r-1)(\ell-4)b_1^2}{2}\,\frac{P_s^2}{\bar{B}^2}\;+ \\
&\quad +\sum_{j=3}^{m}(j-1)\left[(m-j)(r-1)+\frac{(m+2-j)(r-1)(r-2)}{2}+\right. \\
&\qquad \left.+\frac{(j-2)(r-1)(r-2)}{2}\;+\frac{(r-1)(j-2)}{2}\right]\left(\sum_{\substack{(l_1,l_2)\\ l_1+l_2=j}}b_{l_1}b_{l_2}\right)\,\frac{P_s^2}{\bar{B}^2}\;+O(P_s^3)\;= \\
&= \left\{\frac{r-1}{2}\;\left[(\ell-4)b_1^2\;+\sum_{j=3}^{m}(j-1)(\ell-2-j)\;BB_j\right]\right\}\,\frac{P_s^2}{\bar{B}^2}\;+O(P_s^3)\;= \\
&= \left\{\frac{r-1}{2}\;\left[\sum_{j=2}^{m}(j-1)(\ell-2-j)\;BB_j\right]\right\}\,\frac{P_s^2}{\bar{B}^2}\;+O(P_s^3)\,. \qquad (70)
\end{aligned}
$$

The probability $P_{\text{seg}}^{(2,b)}$ of encountering two bursts of errors that lead to a segment error, while the total number of errors does not exceed m, is given by $P_{\text{seg}}^{(2,b)} = P_{\text{seg}}^{(2,b.1)} + P_{\text{seg}}^{(2,b.2)} + P_{\text{seg}}^{(2,b.3)}$. Combining (64), (65) and (70) yields

$$P_{\text{seg}}^{(2,b)} = \frac{r-1}{2} \left\{ \sum_{j=2}^{m} (j-1)[4\,GB_j + (\ell-2-j)\,BB_j] \right\} \frac{P_s^2}{\bar{B}^2} + O(P_s^3)\,. \tag{71}$$

Making use of (51) and (52), after some manipulations, the term in braces in (71) is written as follows:

$$\begin{aligned}
&\sum_{j=2}^{m}(j-1)[4\,GB_j + (\ell-2-j)\,BB_j] = \\
=\;& \sum_{j=2}^{m}(j-1)[4\,(GG_j - GG_{j+1} + G_j) + (\ell-2-j)\,(GG_j - 2\,GG_{j+1} + GG_{j+2} + 2G_j - 2G_{j+1})] = \\
=\;& \sum_{j=2}^{m}(j-1)[(\ell+2-j)\,GG_j - 2(\ell-j)\,GG_{j+1} + (\ell-2-j)\,GG_{j+2} + 2(\ell-j)G_j - 2(\ell-2-j)G_{j+1}] = \\
=\;& \sum_{j=2}^{m}(j-1)(\ell+2-j)\,GG_j - 2\sum_{i=3}^{m+1}(i-2)(\ell+1-i)\,GG_i + \sum_{i=4}^{m+2}(i-3)(\ell-i)\,GG_i + \\
&\qquad +2\sum_{j=2}^{m}(j-1)(\ell-j)G_j - 2\sum_{i=3}^{m+1}(i-2)(\ell-1-i)\,G_i = \\
=\;& \left[\sum_{j=4}^{m} 2\,GG_j\right] + \ell\,GG_2 + 2\,GG_3 - (m\ell - m^2 + m - 2)\,GG_{m+1} + (m-1)(\ell-m+2)\,GG_{m+2} + \\
&\qquad + \left[\sum_{j=3}^{m} 2(\ell-2)G_j\right] + 2(\ell-2)G_2 - 2(m-1)(\ell-2-m)G_{m+1} = \\
=\;& \ell + 2\sum_{j=3}^{m} GG_j - (m\ell - m^2 + m - 2)\,GG_{m+1} + (m-1)(\ell-m+2)\,GG_{m+2} + \\
&\qquad +2(\ell-2)\sum_{j=2}^{m} G_j - 2(m-1)(\ell-2-m)G_{m+1} = \\
=\;& \ell + 2\left(\sum_{j=2}^{m} GG_j - 1\right) - (m\ell - m^2 + m - 2)\,GG_{m+1} + (m-1)(\ell-m+2)\,GG_{m+2} + \\
&\qquad +2(\ell-2)\left(\sum_{j=1}^{m} G_j - 1\right) - 2(m-1)(\ell-2-m)G_{m+1}\,.
\end{aligned} \tag{72}$$

Substituting (61) and (72) into (71) yields

$$\begin{aligned}
P_{\text{seg}}^{(2,b)} = \;& \frac{\ell-m}{2m}\Bigg[2-\ell+(m-1)(\ell-m-2)\;GG_{m+2} - (m\ell-m^2+m-2)\;GG_{m+1} + \\
&+2\sum_{j=2}^{m} GG_j - 2(m-1)(\ell-2-m)\;G_{m+1} + 2(\ell-2)\sum_{j=1}^{m} G_j\Bigg]\;\frac{P_s^2}{\bar{B}^2} + O(P_s^3)\,.
\end{aligned} \tag{73}$$

Substituting (53) and (73) into (59), yields the following:

$$P_{\text{seg}}^{(2)} = \left\{ \frac{\ell - m}{2m} \quad \left[2 - \ell + 2 \sum_{j=2}^{m} GG_j - 2(m-1)\ GG_{m+1} + 2(\ell - 2) \sum_{j=1}^{m} G_j \right] + \right. \\ \left. \frac{(\ell - 2m)(\ell - m - 2)}{2m}\ (\ 2G_{m+1} - GG_{m+2}\ ) \right\} \frac{P_s^2}{\bar{B}^2}\ + O(P_s^3)\ . \tag{74}$$

Owing to Proposition 2, and by making use of (46) and (74), we get (30) – (32). ■

REFERENCES

[1] A. Dholakia, E. Eleftheriou, X.–Y. Hu, I. Iliadis, J. Menon, and KK Rao, "Analysis of a New Intra-Disk Redundancy Scheme for High-Reliability RAID Storage Systems in the Presence of Unrecoverable Errors," *ACM SIGMETRICS Performance Evaluation Review (Proc. ACM SIGMETRICS/IFIP Performance 2006, Saint Malo, France)*, vol.34, no. 1, pp. 373 - 374, June 2006.

[2] A. Dholakia, E. Eleftheriou, X.–Y. Hu, I. Iliadis, J. Menon, and KK Rao, "Analysis of a New Intra-Disk Redundancy Scheme for High-Reliability RAID Storage Systems in the Presence of Unrecoverable Errors," *IBM Research Report RZ 3652*, March 16, 2006.

[3] K. S. Trivedi, *Probabilistic and Statistics with Reliability, Queueing and Computer Science Applications, 2nd Ed.* Wiley, New York, 2002.

[4] D.A. Patterson, G. Gibson, and R.H. Katz, "A Case for Redundant Arrays of Inexpensive Disks (RAID)," *Proc. ACM SIGMOD-88*, pp. 109-116, 1988.

[5] P. M. Chen, E. K. Lee, G. A. Gibson, R. H. Katz, and D. A. Patterson, "RAID: High-Performance, Reliable Secondary Storage," *ACM Computing Surveys*, vol. 26, no. 2, pp. 145-185, June 1994.

[6] L. Kleinrock, *Queueing Systems, Volume 2: Computer Applications*, Wiley, New York, 1976.