Coding Efficiency and Reliability in Probe-based Storage Devices

Maria Varsamou and Theodore Antonakopoulos
Department of Electrical and Computer Engineering
University of Patras, Rio-Patras, 26500, Greece
Email: mtvars@upatras.gr, antonako@upatras.gr

Abstract—High density probe-based storage devices use multiple, simultaneously accessed parallel channels for achieving high I/O data rates. This paper presents an analytical methodology for evaluating the performance of coding and interleaving schemes in such devices, when they are affected by burst errors. Markov processes are used to describe the burstiness of errors due to external disturbances and analytical formulas are provided to estimate the system reliability for various system parameters. Using this approach, the most appropriate system configuration, in terms of number of Reed-Solomon codewords, interleaving depth and coding rate can be determined for a given system reliability and storage efficiency.

Index Terms—Probe storage, Reed-Solomon codes, Burst errors, Performance evaluation.

Fig. 1. Storage system with multiple data channels.

I. INTRODUCTION

The constantly increasing need for higher data rates and storage densities requires the use of multiple, simultaneously accessed channels in storage systems. In this case, the user data block (sector) is partitioned into several sub-blocks, which are accessed concurrently, each one over a different storage channel. The general architecture of a system with parallel channels is shown in Fig. 1. When the storage in a new data area is initiated, a preamble is written initially, for synchronization purposes, and then a number of data sub-blocks is appended. A characteristic example of such a parallel system is found in storage on probe-based devices [1], where ultrahigh storage densities and high data rates can be achieved by using atomic force microscope (AFM) techniques to write and read back data in very thin polymer films with the parallel operation of arrays with multiple tips.

Several coding schemes for such a system have been proposed and studied. One scheme with both random and burst error correction capabilities, which is based on Reed-Solomon (RS) codes along with proper interleaving is shown in Fig. 1. The original data block is partitioned into a number of datawords, which are then encoded using an RS code. If L is the sector size and an $RS(n,k)$ code is used, then $M = \left\lceil \frac{L}{k} \right\rceil$ codewords are created and the RS code can correct up to $t = \left\lfloor \frac{n-k}{2} \right\rfloor$ symbol errors per codeword. The encoded data are symbol-interleaved and then split into smaller blocks, each of which is stored over a separate channel. If N is the number of channels, then $L_t = M \cdot \left\lceil \frac{n}{N} \right\rceil$ symbols are stored over each channel per data block.

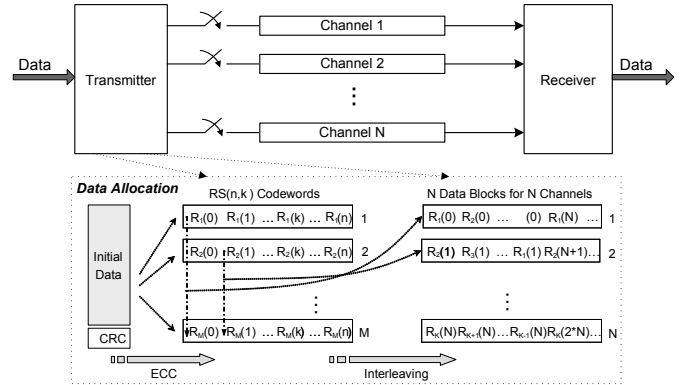An application of this coding scheme is found in the data storage system presented in [2]. If the channels are statistically independent, then a burst of errors in a single channel will be spread across multiple codewords and the decoder stands good chances of correcting it. When a number of channels is affected by the same burst, the system's capability to correct all occurred errors depends on the used error correcting code, the interleaving scheme and the statistical association between the different data channels.

The main goal when designing the dataflow configuration of such devices is to achieve the maximum possible storage efficiency for a given system reliability, or to achieve the maximum possible system reliability for a given storage efficiency. The storage efficiency depends on the redundancy added for error correction purposes and the preservation of the interleaver depth, while the system reliability expresses the capability to retrieve correctly a sector already stored in the device. In [2], the effect of the dataflow parameters on the storage efficiency of probe-based storage devices has been studied and a data allocation procedure that achieves the maximum possible storage efficiency for a given RS code rate has been presented. In this work, we study the effect of the coding/interleaving scheme parameters on the system's reliability in devices with parallel channels. Since these devices may consist of a large number of parallel channels, questions arise regarding the structure of the coding scheme that maximizes the reliability of the device. For example, certain aspects of the system that need to be studied are whether a sector should be allocated in all channels, or to a smaller number of fields, which is the most appropriate
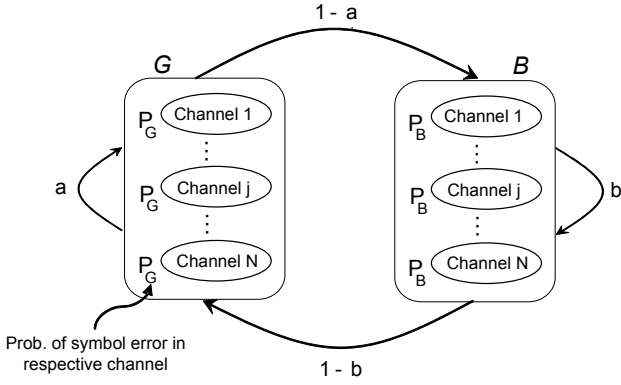
Fig. 2. Markov model for bursts affecting all channels concurrently

sector size, or, for a given code rate, how the number of codewords affects the device's reliability. We study these issues under the most severe possible error conditions, i.e. when an external disturbance is applied to the system. In this case, all channels are affected simultaneously and with the same statistical characteristics. Although the interleaving mechanism spreads the codeword symbols among all channels, depending on the number of channels and the duration of the noise effect, a great number of burst errors appear in all codewords, probably leading to sector decoding failure.

We present a burst error model using Markov processes to describe the mechanism of burst errors that appear in a set of simultaneously accessed channels, along with an analytical methodology for calculating the probability of sector decoding failure for various system parameters. This way the system reliability, in terms of error correction capability, can be evaluated.

II. SYSTEM RELIABILITY ANALYSIS

A. Burst Errors Model

In order to study the correlated symbol errors that appear in the various channels and affect all data codewords, due to an external noise source, and to determine their effects on the error correction capability of the coding scheme, the use of a reliable burst error model is needed. Since all channels are affected by the same noise source with the same probabilistic characteristics, we use the Markov model shown in Fig. 2, which is based on the well-known Gilbert-Elliot model [3]. All N channels enter the G ('good') or B ('bad') state at the same RS symbol offset from the synchronization preamble. When a channel is in the G state, an error occurs with probability P_G, whereas when it is in the B state, an error occurs with probability P_B, and $P_G \ll P_B$. The occupancy times for states B and G are both geometrically distributed with respective means $(1-a)^{-1}$ and $(1-b)^{-1}$. In each state there are $N+1$ distinct events that can be observed. There may be N channels without errors, 1 channel in error and $N-1$ channels without errors, up to all N channels with errors. More analytically, the characteristics of the Markov model are:

1) Number of states: 2, $\{S_G, S_B\}$ where $S_G = G$ and $S_B = B$.
2) The state transition probabilities $q = \{q_{ij}\}$, where $i, j \in \{S_G, S_B\}$. The respective state transition probability matrix is:

$$Q = \begin{bmatrix} q_{GG} & q_{BG} \\ q_{GB} & q_{BB} \end{bmatrix} = \begin{bmatrix} a & 1-b \\ 1-a & b \end{bmatrix} \quad (1)$$

3) The steady state probabilities $\pi = \{\pi_i\}$, i.e. the probabilities of being in state i. The respective steady state probability matrix is:

$$\Pi = \begin{bmatrix} \pi_G \\ \pi_B \end{bmatrix} = \begin{bmatrix} \frac{(1-b)}{(1-a)+(1-b)} \\ \frac{(1-a)}{(1-a)+(1-b)} \end{bmatrix} \quad (2)$$

4) The number of distinct events that can be observed in each state: $N+1$. These events are called observations. So, the set of possible observations in each state is:

$$V = \{V_0, V_1, V_2 \ldots V_N\}$$

where V_j denotes the event that j out of N fields are in error.

5) The observation probabilities $\beta = \{\beta_{i,j}\}$ with $i \in \{S_G, S_B\}$ and $j \in V$, i.e. the probability of observing V_j when the model is in state i.
Let us assume that the model is in the B state. Since there are $\binom{N}{j}$ different ways that errors occur in j out of the N channels, the probability of this event is:

$$\beta_{B,V_j} = \binom{N}{j} \cdot P_B^j \cdot (1 - P_B)^{N-j} \quad (3)$$

The observation probabilities β_{G,V_j} are defined accordingly. Then, the total probability, P_{V_j}, that errors occur in j out of N channels is given by:

$$P_{V_j} = \sum_{i \in \{S_G, S_B\}} \beta_{i,V_j} \cdot \pi_i \quad (4)$$

B. Sector Decoding Failure Probability

To evaluate the system reliability, we need to estimate the probability that a data block (sector) was not retrieved correctly. We denote as sector decoding failure, P_{sdf}, the probability that at least one codeword is not decoded correctly, meaning that either the decoder fails to find a codeword or finds a codeword other than the originally stored [4]. Since an RS codeword can be decoded correctly if up to t errors have occurred, the probability P_{cdf} that a codeword cannot be decoded correctly is given by:

$$P_{cdf} = 1 - \sum_{i=0}^{t} P_i \quad (5)$$

where P_i is the probability that i errors have occurred. When an external disturbance is applied to the device, thus affecting all channels with the same statistical characteristics, then the codewords are also affected by noise with the same statistical characteristics. So, we have to study the effect of the burst

errors and to calculate P_{cdf} for a single codeword. Then, the sector decoding failure, P_{sdf}, is given by:

$$P_{sdf} = 1 - (1 - P_{cdf})^M \qquad (6)$$

This means that, in order to calculate the sector decoding failure P_{sdf}, we need a method for calculating the probabilities P_i, for $0 \leq i \leq t$. For this purpose, we have to use a modified version of the aforementioned Markov model.

The effect of interleaving: Since the codeword symbols are interleaved not only inside a single channel but also among the N channels, as shown in Fig. 1, the effect of the interleaving mechanism and the interleaver depth has to be incorporated in the Markov model. The used interleaving scheme, named *shifted interleaving*, is the following: the symbols of a codeword are allocated so that successive symbols are located in successive offset locations in consecutive channels up to the number of the used codewords. There are two cases that need to be studied:

- $N \geq M$: In this case, in a single symbol duration per channel, only $K = N/M$ symbols of each codeword are affected by an external disturbance applied in all channels. Note that since the parameters of the coding/interleaving scheme affect significantly both the hardware complexity and the processing time, N is chosen to be a multiple of M, and so, K is an integer. An illustrative example of this case is given in Fig. 3, where $K = 2$. Since we are interested in finding the probability of having more than t errors in a codeword, the number of distinct events that can be observed in each state of the Markov model are $K + 1$, i.e. the probability of having 0 errors in a codeword, or 1 out of K symbols of the codeword in error and so on, up to having all K symbols in error. So, the set of possible observations in each state is:

$$V = \{V_0, V_1, V_2 \ldots V_K\}$$

and the observation probability in (3) has to be modified, by replacing N with K.

Using shifted interleaving, in each symbol duration there are always K symbols of the same codeword which are affected simultaneously. This means that although in a single channel the consecutive symbols of the same codeword are M places apart from one another, when all channels are observed, the interleaving depth inside a codeword becomes $1/K$. When $K > 1$ means that K symbols are affected per symbol duration during a burst. As a result, the state transition probability matrix of (1) describes the possible transitions of the K codeword symbols between G and B states without any modification. To compute the probability P_i that i errors occur in a codeword, i.e. i errors in a group of n symbols, and since in each symbol duration K symbols from a codeword are affected, we need to calculate P_i in a sequence of $T = (n/K)$ symbol durations. The i errors may be spread in this sequence in all possible ways the number i can be partitioned and allocated in T smaller
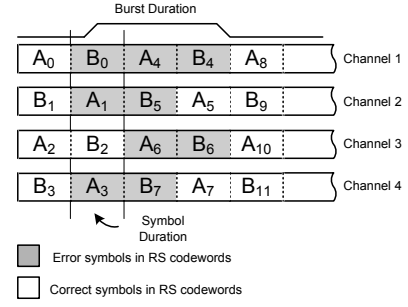

Fig. 3. Example case of burst errors appearance in a device with $N = 4$ channels when the sector is partitioned in $M = 2$ codewords.
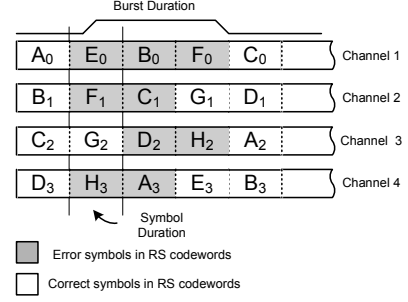

Fig. 4. Example case of burst errors appearance in a device with $N = 4$ channels when the sector is partitioned in $M = 8$ codewords.

integers, each of which is less or equal to K. To specify the number of all possible ways, at first we find all the partitions of i into smaller integers and then, for each partition we calculate the number of possible allocations of the elements that comprise the partition in T symbol durations, as presented in the Appendix. Each one of these allocations constitutes an observation sequence of the Markov model.

If λ denotes the probabilistic characteristics of the Markov model, then to compute the probability of occurrence of an observation sequence $O = O_1, O_2, \ldots, O_j, \ldots, O_T$, where $O_j \in V$, given the Markov model λ, we use the following Forward-Backward procedure [5], [6]:

We define the forward variable $\alpha_t(i)$ as:

$$\alpha_t(i) = P(O_1, O_2, \ldots, O_t, S = i | \lambda)$$

i.e the probability of the partial observation sequence up to time t and state $S = i$, given the model λ. $\alpha_t(i)$ can be computed by induction as follows:

1) for $i \in \{S_G, S_B\}$:

$$\alpha_1(i) = \pi_i \cdot \beta_{i,O_1}$$

2) for $t = 1, 2, \ldots, T - 1, j \in \{S_G, S_B\}$:

$$\alpha_{t+1}(j) = \left[\sum_{i \in \{S_G, S_B\}} \alpha_t(i) \cdot q_{ij} \right] \cdot \beta_{j,O_{t+1}}$$
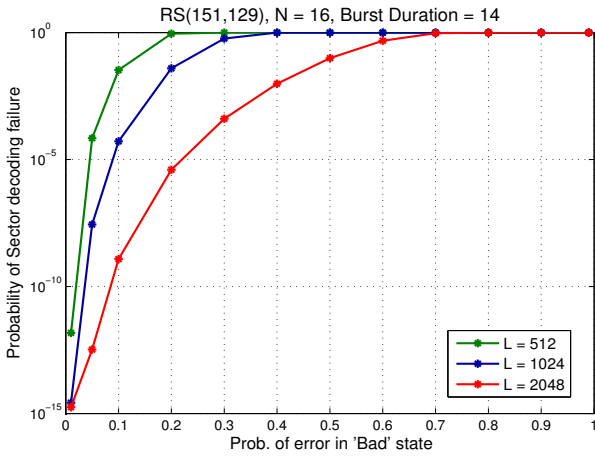
Fig. 5. Probability of sector decoding failure for different values of sector size.
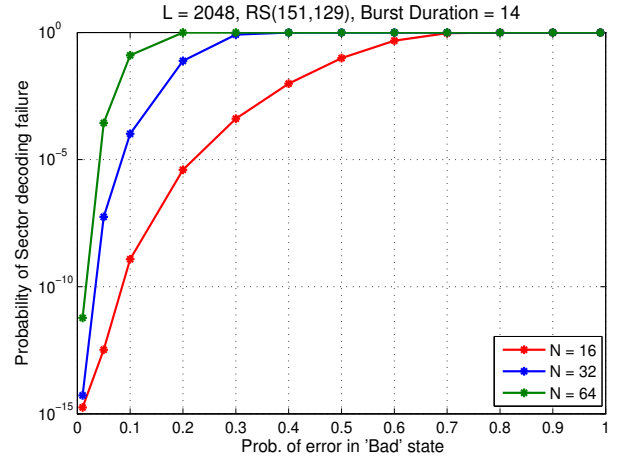

Fig. 6. Probability of sector decoding failure for different values of the number of fields used to store a sector.

3) then we have:

$$P(O|\lambda) = \sum_{i \in \{S_G, S_B\}} \alpha_{\tau}(i)$$

Finally, the probability P_i is given by:

$$P_i = \sum_{\text{for all } O} P(O|\lambda) \qquad (7)$$

- $N < M$:

  In this case, in a single symbol duration, at most 1 symbol per codeword is affected when an external disturbance causes burst errors in all channels. An illustrative example of this case is shown in Fig. 4. Since we study the system on a codeword basis, the set of possible observations in each state is $V = \{V_0, V_1\}$, i.e. 0 or 1 symbol error in a codeword for a single symbol duration, and the observation probability in (3) is simplified for $N = 1$.

  Due to the used interleaving mechanism, the interleaving depth in a codeword, when considering the whole set of N channels, is M/N. M/N is also an integer, since M is a multiple of N for keeping the system complexity at a reasonable level. This means that the same codeword may be affected by a burst every M/N RS symbols duration, which can also be verified by the example in Fig. 4. As a result, the state transition probability matrix (1) has to be modified. According to the proof by Yee and Weldon [7], the (M/N)-step transition probabilities can be obtained by raising the one-step transition matrix to the (M/N)th power. In this case, to compute the probability P_i that i errors appear in a codeword we need to calculate P_i in a sequence of n symbols duration. For this, we use the aforementioned procedure for calculating the probability of an observation sequence.

III. NUMERICAL RESULTS

An application of the aforementioned model can be found in the probe-based data storage device presented in [1]. In this device, the information is stored by means of thermomechanical formation of indentations in thin polymer films, using nanometer-sharp tips, similar to those used in atomic force microscopy (AFM) [8]. To increase the achievable data rates, the use of a large array of probes has been proposed. In this case, each probe performs read/write/erase operations on a dedicated area, named a data field, while the storage medium is placed in the x/y plane.

As in conventional storage devices, the data are stored in the form of sectors of fixed length. If N is the number of probes operating in parallel, then each sector is encoded as shown in Fig. 1. Finally N smaller blocks are formed, and each one is stored in a single storage field. During data read, the microscanner moves the storage fields under their associated tips, such that each tip operates in the center of the line with the sequence of indentations corresponding to the specific sector. Since each probe operates on a distinct storage area, the N parallel read channels are statistically independent, only when there is no external source of noise. However, an external shock or vibration applied to the device, and consequently to the microscanner, while reading or writing a sector, will cause the same displacement to all tips, thus producing burst errors in all codewords.

We apply the proposed analytical methodology to investigate the effect of various dataflow parameters on the system's reliability, when burst errors appear in all storage fields due to an external disturbance. We assume that burst errors appear only when the device enters the *B* state, which means that $P_G = 0$. Fig. 5 shows the effect of the sector size, while Fig. 6 shows the effect of the number of storage fields used for storing one sector. In all cases, an RS$(151, 129)$ code with 0.8543 code rate is considered, which has been used in [1] according to device specifications and measurements regarding the random error rate. According to these results, for a given coding rate, large sector sizes and small number of storage fields lead to better performance, in terms of system
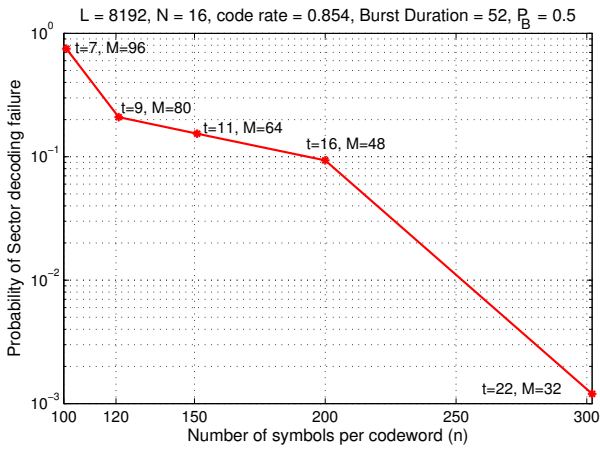
Fig. 7. Probability of sector decoding failure for different combinations of the number of codewords M and the $\mathrm{RS}(n, k)$ code for a given code rate.

Fig. 8. Probability of sector decoding failure for different combinations of the number of codewords M and the $\mathrm{RS}(n, k)$ code for various code rates.

reliability. This means that, although a device may include a very large number of storage fields that can operate in parallel, a single sector should be allocated in a small group of fields and multiple sectors have to be stored in parallel.

We also study the effect of the number of codewords, and consequently the $\mathrm{RS}(n, k)$ code and the number of errors that it can correct, when a given code rate is used in the device. There exist many RS codes that satisfy a specific code rate. Smaller values of n and k, which also result to a smaller error correction capability per codeword, mean that the sector will be partitioned into a larger number of codewords. Fig. 7 shows how the selection of several combinations of M and $\mathrm{RS}(n, k)$ affect the performance of the device, when a large sector size of $8K$ is used, the sector is allocated in 16 fields and the RS code operates in $GF(2^9)$. The use of a small number of codewords in a sector, which favors the error correction capability of the code, is highly preferable. Finally, Fig. 8 demonstrates how the code rate can significantly alter the device's reliability performance.

IV. CONCLUSIONS

We presented a Markov model that describes the burst errors that appear in a device that employs multiple simultaneously-accessed channels, when it is affected by an external noise source. Additionally, we presented an analytical methodology for deriving statistics about the probability of failure of the stored data, which enables the evaluation of the system reliability for various coding/interleaving schemes. This way, the most appropriate combination of data structure parameters can be specified, given the device requirements on reliability and storage efficiency.

APPENDIX

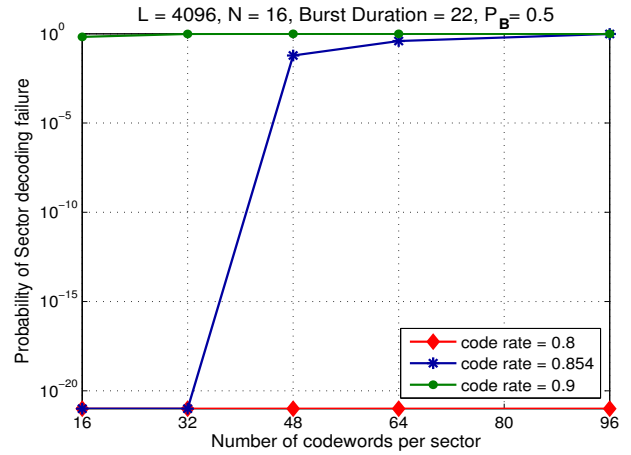As partition of an integer we call any representation of that integer as a sum of smaller integers. Generally, a partition of an integer n can be described using the following sum:

$$n = \sum_{i=1}^{n} i \cdot z_i$$

where the coefficient z_i denotes how many times the number i appears in the partition. A common algorithm that is used for generating the partitions of an integer in anti-lexicographic order (from the biggest integer to the smallest) is the one proposed by Stockmal [9].

The number of all possible allocations of a specific partition of an integer n into exactly m bins is given by:

$$p(m; m_1, m_2, \ldots, m_n, m_0) = \frac{m!}{m_1! \cdot m_2! \cdots m_n! \cdot (m - M)!}$$

where $m_1, m_2, \ldots, m_n$ denote how many times that the respective integer appears in the partition (they can also have zero value), $M = m_1 + m_2 + \ldots + m_n$ and $m_0 = m - M$.

REFERENCES

[1] A. Pantazi, A. Sebastian, *et al.*, "Probe-based ultrahigh-density storage technology," *IBM J. Res. and Dev.*, vol. 52, no. 4/5, pp. 493–511, 2008.
[2] M. Varsamou and T. Antonakopoulos, "A new data allocation method for parallel probe-based storage devices," *IEEE Transactions on Magnetics*, vol. 44, no. 4, pp. 547–554, Apr. 2008.
[3] E. O. Elliot, "Estimates of error rates for codes on burst-noise channels," *Bell Syst. Tech. J.*, vol. 42, pp. 1977–1997, Sep. 1963.
[4] R. J. McEliece and L. Swanson, "On the Decoder Error Probability for Reed-Solomon Codes," *IEEE Transactions on Information Theory*, vol. IT-32, no. 5, pp. 701–703, Sep. 1986.
[5] L. R. Rabiner, "A tutorial on Hidden Markov Models and Selected Applications in Speech Recognition," *Proc. IEEE*, vol. 77, pp. 257–273, Feb. 1989.
[6] L. N. Kanal and A. R. K. Sastry, "Models for channels with memory and their applications to error control," *Proc. IEEE*, vol. 66, pp. 724–744, Jul. 1978.
[7] J. R. Yee and E. J. Weldon, "Evaluation of the Performance of Error-Correcting Codes on a Gilbert Channel," *IEEE Transactions on Communications*, vol. 43, pp. 2316–2323, Aug. 1995.
[8] G. Binnig, C. F. Quate, and C. Gerber, "Atomic force microscope," *Phys. Rev. Lett*, vol. 56, no. 9, pp. 930–933, 1986.
[9] F. Stockmal, "Generation of partitions in part-count form, Algorithm 95," *Communications of the ACM*, vol. 5, p. 344, 1962.