

Reliable Memories Built from Unreliable Components Based on Expander Graphs

Shashi Kiran Chilappagari, *Student Member, IEEE*, and Bane Vasic, *Member, IEEE*

Abstract

In this paper, memories built from components subject to transient faults are considered. A fault-tolerant memory architecture based on low-density parity-check codes is proposed and the existence of reliable memories for the adversarial failure model is proved. The proof relies on the expansion property of the underlying Tanner graph of the code. An equivalence between the Taylor-Kuznetsov (TK) scheme and Gallager B algorithm is established and the results are extended to the independent failure model. It is also shown that the proposed memory architecture has lower redundancy compared to the TK scheme. The results are illustrated with specific numerical examples.

I. INTRODUCTION

During the past four decades, the decrease in transistor size and the increase in integration factor have led to very small, fast, and power efficient chips. As the demand for power efficiency continues, a wide range of new nano-scale technologies is being actively investigated for processing and storage of digital data. Although it is difficult to discern which of these approaches will become a technological basis for computers in the future, it is widely recognized that due to their miniature size and variations in technological process, the nano-components will be inherently unreliable. Even in more traditional

Manuscript received April 30, 2007. This work is funded by the NSF under grant CCF-0634969

S. K. Chilappagari is with the Department of Electrical and Computer Engineering, University of Arizona, Tucson, AZ, 85721 USA (e-mail: shashic@ece.arizona.edu).

B. Vasic is with the Department of Electrical and Computer Engineering, University of Arizona, Tucson, AZ, 85721 USA (e-mail: vasic@ece.arizona.edu).

semiconductor technologies, reducing transistor size has already started affecting circuit reliability, and it is widely believed that transistor failures (both transient and permanent) will become one of the main technological obstacles as the trend of increasing the integration factor continues. In this paper, we consider storage circuits built from such unreliable (faulty) components. We consider an unreliable component (a logic gate or a memory element) to be a component that is subject to *transient faults*, i.e., faults that manifest themselves at particular time steps but do not necessarily persist for later times [1].

Von Neumann [2] was the first to study computation using faulty gates. In [2], he showed that, under certain conditions, increased gate redundancy can lead to increased reliability of a circuit. However, it was shown that, in general, computation by faulty gates with non-zero computational capacity is not possible (see [3], [4]). The study of storage circuits made of unreliable components led to much more optimistic results. Taylor in [5] proved that a memory has an associated information storage capacity, C , such that arbitrarily reliable information storage is possible for all memory redundancies greater than $1/C$. The methodology of the proof, however, does not allow one to explicitly calculate the storage capacity. Taylor considered two models of component failures and proposed construction of fault-tolerant memories based on low-density parity-check (LDPC) codes. In the first model, the failures of a particular component are assumed to be statistically independent from one use to another and is referred to as the independent failure model. In the second model, the components fail permanently but bad components are replaced with good ones at regular intervals. The failures in different components are assumed to be independent in both the models. This construction was further studied by Kuznetsov in [6] and we will refer to it as the Taylor-Kuznetsov (TK) scheme. Hadjicostis [1] was able to generalize Taylor's scheme to fault tolerant linear finite state machines. Spielman [7] obtained the best result for a general model of computation, by marrying the ideas of von Neumann with Reed-Solomon (RS) codes.

The fundamental contribution of this paper is to show existence of reliable memories built entirely from unreliable components and which have finite redundancies. We consider the adversarial failure model in which only a fixed fraction of the components fail at any given time and extend our results to the

independent failure model using Chernoff bounds [8]. Our memory architecture has lower redundancy compared to the TK scheme. Our fault tolerant memory architecture is also based on LDPC codes but differs from the TK scheme in the decoding algorithm employed. The TK scheme can be shown to be an implementation of the Gallager B decoding algorithm for LDPC codes (the proof will be given in Section V). We use the parallel bit flipping decoding algorithm proposed in the context of expander codes by Sipser and Spielman [9]. Expander codes are a class of asymptotically good error correcting codes with linear time decoding algorithms which can correct a linear fraction of errors. Expander graph based arguments have been successfully applied for message passing algorithms by Burshtein and Miller in [10] as well as for linear programming decoding by Feldman *et.al* in [11]. At the time of their discovery, explicit construction of graphs with expansion required for parallel bit flipping algorithm were not known. Capalbo *et al.* [12] recently gave an explicit construction of expander graphs based on randomness constructors. Hence, our method can be seen as a constructive proof in contrast to Taylor's method which is an existence proof.

The rest of the paper is organized as follows. In Section II we provide the necessary definitions and a brief overview of LDPC codes. We explain the proposed memory architecture and characterize it in terms of complexity and redundancy. In Section III we introduce the model of failure of the components and prove our main result showing the existence of memories which can tolerate failures in all the components. In Section IV we provide a few numerical examples. In Section V we establish an equivalence between the TK scheme and Gallager B algorithm and extend our results to the independent failure model. In Section VI we discuss open questions and conclude with some interesting remarks.

II. THE SYSTEM DESCRIPTION

In this section, we give a detailed description of the memory system. We start by introducing the terminology used to characterize memories and proceed to discuss the importance of LDPC codes. We explain the coding scheme and the error correction scheme employed in the proposed memory architecture. We then calculate the redundancy and complexity associated with the memories.

A. Definitions

A memory is a device in which information is stored at some time and retrieved at a later time [5]. The memories under consideration store information in form of bits and are built from registers (memory elements) each of which can store a single bit. The information storage capability of a memory is the number of information bits it stores. Consider a memory built out of reliable registers. To build a memory with information storage capability of k bits requires k registers. Such a memory is termed as an irredundant memory. Now, consider the problem of information storage with unreliable memory elements. Due to the component failures, the information read out of the memory may not be identical to the information stored originally. Hence, to ensure reliable storage, the information needs to be stored in coded form (see [5] for an excellent discussion on the importance of coded form). Initially, a codeword from some error correcting code is stored in the memory. The unreliable nature of the memory elements introduces errors in the registers and the contents of the memory differ from the initial state. To ensure reliability, a correcting circuit is employed which performs error correction and updates the contents of the registers with an estimate of the original codeword. Hence, a fault-tolerant memory system (referred to as memory system or simply memory henceforth) consists of memory elements (referred to as storage circuit) and a correcting circuit. The correcting circuit is also built of unreliable components. The coding of information along with the correcting circuit introduce redundancy into the memory system. Such redundant memories are characterized by two closely related parameters, namely, complexity and redundancy. The *complexity* of a memory is the number of components within the memory (a component is a device which either performs an elementary operation or stores a single bit where an elementary operation is any Boolean function of two binary operands [5]). The *redundancy* of a memory is the ratio of the complexity of the memory to the complexity of an irredundant memory which has the same information storage capability [5]. It should be noted that there can be many memory architectures with different complexities but the same information storage capability.

Another important characteristic of a memory is reliability. We say that arbitrarily reliable information

storage is possible in a memory if the probability of memory failure can be made arbitrarily small. To quantify the reliability of a memory system, it is important to first define what constitutes a memory failure. Let a memory failure be defined as an event in which the word read out of memory is not equal to the original codeword. Arbitrarily reliable information storage is not possible with such a definition of memory failure. This is due to the fact that the probability of failure is lower bounded by the probability of failure of components in the final step of extracting the information bits. Hence, we define a failure in the following manner. Associated with each codeword in a code is a decoding equivalence class, i.e., the set of words which decode to that particular codeword when decoded with a decoder built of reliable components. If the contents of the memory do not belong to the decoding equivalence class of the original codeword, we say a memory failure has occurred. The *storage capacity*, C , of a memory is a number such that for all memory redundancies greater than $1/C$, arbitrarily reliable information storage is possible [5].

B. LDPC Codes

The memories under consideration store information in form of bits and therefore we restrict our attention to binary codes in this paper. An (n, k) binary block code maps a message block of k information bits to a binary n -tuple [13]. The rate r of the code is given by $r = k/n$. An (n, k) binary linear block code, \mathcal{C} , is a subspace of $GF(2)^n$ of dimension k [13]. A parity check matrix H of \mathcal{C} is a matrix whose columns generate the orthogonal complement of \mathcal{C} , i.e., an element \mathbf{w} of $GF(2)^n$ is a codeword of \mathcal{C} iff $\mathbf{w}H^T = \mathbf{0}$ [14]. The information storage capability of a memory depends on the type of the code employed in the correcting circuit. Hence, a memory employing an (n, k) block code has information storage capability of k bits.

Taylor in [5] argues that no decoding scheme other than iterative decoding of LDPC codes can achieve non-zero storage capacity. LDPC codes [15] are a class of linear block codes which can be defined by sparse bipartite graphs [16]. Let \mathcal{G} be a bipartite graph with two sets of nodes: n variable (bit) nodes and m check (constraint) nodes. The check nodes (variable nodes) connected to a variable node (check node) are referred to as its neighbors. The degree of a node is the number of its neighbors. This graph defines

a linear block code of length n and dimension at least $n - m$ in the following way: The n variable nodes are associated to the n coordinates of codewords. A vector $\mathbf{v} = (v_1, v_2, \dots, v_n)$ is a codeword if and only if for each check node, the sum of its neighbors is zero. Such a graphical representation of an LDPC code is called the Tanner graph [17] of the code. The adjacency matrix of \mathcal{G} gives H , a parity check matrix of \mathcal{C} . An (n, γ, ρ) regular LDPC code has a Tanner graph with n variable nodes each of degree γ and $n\gamma/\rho$ check nodes each of degree ρ . This code has length n and rate $r \geq 1 - \gamma/\rho$ [16]. It should be noted that the Tanner graph is not uniquely defined by the code and when we say the Tanner graph of an LDPC code, we only mean one possible graphical representation.

C. The Proposed Fault Tolerant Memory Architecture

The complexity and redundancy of a fault-tolerant memory depend on the coding scheme as well as the decoding algorithm employed in updating the contents of the memory. We now explain our memory architecture in detail.

At time $t = 0$, a codeword from an (n, γ, ρ) regular binary LDPC code is written into the storage circuit consisting of n registers each of which can store a single bit. The n bits of the codeword correspond to the n variable nodes in the Tanner graph, \mathcal{G} , of the code. The contents of the registers are updated at times $\tau, 2\tau, \dots, L\tau$, $L \in \mathbb{N}$. The update rules can be explained by defining messages along the edges in \mathcal{G} . For a variable node v (check node c), let $E(v)$ ($E(c)$) denote the edges incident on v (c). Each edge e is associated with a variable node v and a check node c . Let $\overrightarrow{m_t}(e)$ and $\overleftarrow{m_t}(e)$ represent the messages passed on an edge e from variable node to check node and check node to variable node at time t respectively. Let $v(t)$ denote the value of variable node v at time t . Then the update at time t is given by the following algorithm:

Algorithm A

- For each edge e and corresponding variable node v

$$\overrightarrow{m_t}(e) = v(t^-)$$

- For each edge e and corresponding check node c

$$\overleftarrow{m}_t(e) = \left(\sum_{e' \in E(c) \setminus \{e\}} \overrightarrow{m}_t(e') \right) \bmod 2$$

- For each variable node v

$$v(t^+) = \begin{cases} 1, & \sum_{e \in E(v)} \overleftarrow{m}_t(e) > \lfloor \gamma/2 \rfloor \\ 0, & \gamma - \sum_{e \in E(v)} \overleftarrow{m}_t(e) > \lfloor \gamma/2 \rfloor \\ v(t^-), & \text{otherwise} \end{cases}$$

The algorithm can be interpreted in the following manner. Every variable node sends an estimate of its value to the neighboring check nodes. A check node calculates an estimate of a neighboring variable node by computing the modulo two sum of all the remaining $(\rho - 1)$ neighboring variable nodes. Each variable node receives γ estimates, one from each neighboring check and the majority of these estimates is the updated value of the node.

Remarks: We assume that the update is instantaneous and use $v(t^-)$ and $v(t^+)$ to denote the value of variable v just before and after the update respectively. We note that the algorithm presented above is a slight modification of the parallel bit flipping algorithm proposed in [9].

D. Complexity and Redundancy

LDPC codes can achieve non-zero capacity due to the fact that the redundancy of the LDPC codes memory increases linearly with the information storage capability. The complexity of the logic gates needed to perform decoding depend only on γ and ρ and not on the length of the code. So the redundancy remains bounded even as the code length tends to infinity.

We now calculate the complexity and redundancy associated with our fault-tolerant memory architecture. The storage circuit consists of n registers each of which can store a single bit and hence has complexity n . The correcting circuit consists of logic gates (built from components) needed to implement the update

algorithm. The message sent from a check node to variable node involves computing the modulo two sum of $(\rho - 1)$ bits which requires a $(\rho - 1)$ -input XOR gate which can be implemented using $(\rho - 2)$ two input XOR gates (a two input XOR gate calculates modulo two sum of two bits). Each check node needs to compute ρ such estimates. Therefore, the total number of two input XOR gates is

$$(n\gamma/\rho) \times \rho \times (\rho - 2) = n\gamma(\rho - 2)$$

Each variable node is updated based on the majority of the γ estimates received from its neighbors. This requires a γ -input majority logic gate for every variable node whose complexity we denote by D_γ . Hence, the complexity of the memory system is

$$\mathcal{S} = n(1 + D_\gamma + \gamma(\rho - 2))$$

The memory has information storage capability of rn bits and the complexity of an irredundant memory with the same information storage capability is rn . The redundancy of the fault-tolerant memory is therefore

$$\begin{aligned} R &= n(1 + D_\gamma + \gamma(\rho - 2))/rn \\ &\leq (1 + D_\gamma + \gamma(\rho - 2))/(1 - \gamma/\rho) \end{aligned}$$

III. ANALYSIS OF THE MEMORY SYSTEM

The storage capacity of a memory depends on the type of failures in the components. A logic gate is said to have failed if its output is flipped. A register is said to have failed if the bit stored in it is flipped. In this paper we consider the adversarial failure model also referred to as bit flipping channel model. In the adversarial model, the failures occur in the worst case fashion but no more than a fixed fraction of the components fail at any given time. In other words, the number of failures is bounded for a given number of components. As the number of components increases so does the number of failures. We denote the fraction of memory element failures in a time interval τ by α_m , fraction of two input XOR gate failures

for every use by α_{\oplus} and fraction of γ -input majority logic gate failures for every use by α_{γ} . As mentioned before, the component failures are transient and independent from one use to another.

A memory system is said to tolerate a constant fraction of errors in all components if at any time at most a constant fraction of components can fail and no memory failure occurs in the system at all times $t < \infty$. Recall that, from our definition, a memory failure occurs if the contents in the memory do not belong to the decoding class of the originally stored codeword. In this section, we prove that the memory architecture proposed in Section II can tolerate a constant fraction of failures in all the components. Our proof is based on the expansion property of the underlying Tanner graph, \mathcal{G} , of the code.

Definition 1: [9] A Tanner graph \mathcal{G} of a (n, γ, ρ) LDPC code is a $(\gamma, \rho, \alpha, \delta)$ expander if for every subset S of at most an αn variable nodes, at least $\delta|S|$ check nodes are incident to S .

The definition of expander is much more general but we restrict our attention to Tanner graphs of LDPC codes.

Sipser and Spielman in [9] proposed a class of asymptotically good error correcting codes based on expander graphs known as expander codes. They proposed two simple bit flipping algorithms, namely, serial and parallel and showed that when the underlying graph has sufficient expansion, these algorithms can correct a fixed fraction of errors. LDPC codes are a special case of expander codes in which the expander graph is the Tanner graph of the LDPC code.

We describe the parallel bit flipping algorithm and interested readers are referred to [9] for details about serial bit flipping. We say that a constraint is satisfied by a setting of variables if the sum of the variables in the constraint is even; otherwise, the constraint is unsatisfied. The set of variable nodes (bits) which differ from their original value are known as corrupt variables.

Parallel Bit Flipping Algorithm

- In parallel, flip each variable that is in more unsatisfied than satisfied constraints.
- Repeat until no such variable remains.

The following theorem from [9] gives the sufficient conditions for the parallel bit flipping algorithm to

correct a constant fraction of errors.

Lemma 1 ([9], Theorem 11): Let \mathcal{G} be a $(\gamma, \rho, \alpha, (3/4 + \epsilon)\gamma)$ expander over n variable nodes, for any $\epsilon > 0$. Then, the simple parallel decoding algorithm will correct any $\alpha_0 < \alpha(1 + 4\epsilon)/2$ fraction of error after $\log_{1-4\epsilon}(\alpha_0 n)$ decoding rounds. Also, if V denotes the set of corrupt variables in the input and $|V| < \alpha n(1 + 4\epsilon)/2$, then the parallel decoding algorithm produces a word with at most $|V|(1 - 4\epsilon)$ corrupt variables after one decoding round.

Proof: See [9] ■

From Lemma 1, it is clear that a word belongs to the decoding class of a codeword as long as the fraction of corrupt variables (bits) is less than $\alpha(1 + 4\epsilon)/2$. Note that Algorithm A is a slight modification of one iteration of the parallel bit flipping algorithm of [9]. In the parallel bit flipping algorithm, every check node indicates to its neighboring variable node if it is satisfied or not. In Algorithm A every check node gives an estimate of the variable node. Theoretically, both the algorithms are equivalent but we use the Algorithm A as it has lesser redundancy. We now state and prove our main theorem.

Theorem 1: Let \mathcal{G} be a $(\gamma, \rho, \alpha, (3/4 + \epsilon)\gamma)$ expander for any $\epsilon > 0$. The proposed memory architecture can tolerate constant fraction of errors in all the components if

$$\alpha_m + \gamma(\rho - 2)\alpha_{\oplus} + \alpha_{\gamma} < \alpha(1 + 4\epsilon)(4\epsilon)/2$$

Proof: At $t = 0$, a codeword from an (n, γ, ρ) LDPC code with Tanner graph \mathcal{G} is written into the memory. The contents are updated at times $\tau, 2\tau, \dots, L\tau$, $L \in \mathbb{N}$, by running Algorithm A. We bound the number of corrupt variables at time t . Let $\alpha_v(t)$ denote the fraction of corrupt variables at time t . We establish bounds on $\alpha_v(t)$ for all t . We first prove the following. Let $\delta > 0$, denote an infinitesimal duration of time. If

$$\alpha_v((l - 1)\tau - \delta) < \alpha(1 + 4\epsilon)/2,$$

then,

$$\alpha_v(l\tau - \delta) < \alpha(1 + 4\epsilon)/2$$

Let $V(t)$ denote the set of corrupt variables at time t .

$$|V((l-1)\tau - \delta)| = \alpha_v((l-1)\tau - \delta)n.$$

Since $\alpha_v((l-1)\tau - \delta) < \alpha(1+4\epsilon)/2$, a decoder built with reliable gates outputs a word with at most $|V((l-1)\tau - \delta)|(1-4\epsilon)$ corrupt variables (by Lemma 1). We now bound the number of errors introduced due to the faulty nature of the decoder. Each XOR gate failure can corrupt at most one variable and each majority logic gate failure can corrupt at most one variable. So,

$$\begin{aligned} |V((l-1)\tau)| &< |V((l-1)\tau - \delta)|(1-4\epsilon) \\ &+ \gamma(\rho-2)\alpha_{\oplus}n + \alpha_{\gamma}n \end{aligned} \quad (1)$$

Eq. 1 bounds the number of corrupt variables at the end of $(l-1)^{th}$ correcting cycle. However, in the time interval $[(l-1)\tau, l\tau)$, at most $\alpha_m n$ variables can get corrupted due to failures in memory elements. Therefore, the time at which there are maximum number of corrupt variables is just before the start of a correcting cycle, i.e.,

$$\alpha_v(l\tau - \delta) = \max\{\alpha_v(t) : (l-1)\tau \leq t < l\tau\}$$

Hence, it suffices to bound $\alpha_v(t)$ for $t = l\tau - \delta, l = 1, 2, \dots, L$.

$$\begin{aligned} |V((l\tau - \delta))| &< |V((l-1)\tau - \delta)|(1-4\epsilon) \\ &+ \gamma(\rho-2)\alpha_{\oplus}n + \alpha_{\gamma}n + \alpha_m n \end{aligned} \quad (2)$$

Dividing Eq. 2 by n gives

$$\begin{aligned} \alpha_v((l\tau - \delta)) &< \alpha_v((l-1)\tau - \delta)(1-4\epsilon) \\ &+ \gamma(\rho-2)\alpha_{\oplus} + \alpha_{\gamma} + \alpha_m \\ &< \alpha(1+4\epsilon)(1-4\epsilon)/2 + \alpha(1+4\epsilon)(4\epsilon)/2 \\ &= \alpha(1+4\epsilon)/2 \end{aligned}$$

Since

$$\alpha_v(\tau - \delta) \leq \alpha_m < \alpha(1 + 4\epsilon)/2,$$

it follows that

$$\alpha_v(l\tau - \delta) < \alpha(1 + 4\epsilon)/2 \quad \forall l \in \mathbb{N}.$$

Hence,

$$\alpha_v(t) < \alpha(1 + 4\epsilon)/2 \quad \forall t < \infty.$$

Since, the fraction of corrupt variables is less than $\alpha(1 + 4\epsilon)/2$, the contents of storage circuit correspond the decoding class of original codeword and hence, by our definition, no memory failure occurs. ■

It is instructive to see the behavior of the memory in the absence of the correcting circuit. In any time interval of τ seconds, at most $\alpha_m n$ fraction of the memories may fail. After sufficiently long time, the fraction of corrupt variables becomes more than $\alpha(1 + 4\epsilon)/2$ and a memory failure occurs. The presence of a correcting circuit ensures that at any time the number of corrupt variables remains less than the correcting capability of the code. However, for a given expander there is a loss in the tolerable memory failure due to the faulty nature of the gates as well as the iterative nature of the decoder. Consider the case of where decoder is reliable and failures occur only once. The tolerable fraction of errors for a given expander is close to $\alpha(1 + 4\epsilon)/2$. In the case of memories with unreliable memory elements but reliable logic gates, the tolerable fraction of memory errors is close to $\alpha(1 + 4\epsilon)(4\epsilon)/2$. The reduction by a factor of 4ϵ occurs due to the fact that decoder is iterative in nature and needs multiple rounds to converge to the codeword. One round of error correction decreases the errors by a factor of $(1 - 4\epsilon)$ and $\alpha_m n$ new errors might be introduced due to memory failures. In the extreme case of $\epsilon = 1/4$ we have a decoder which takes just one step to correct all the corrupt variables, in which case the tolerable failure rate is arbitrarily close to $\alpha(1 + 4\epsilon)/2$. The faulty nature of the decoder further reduces the tolerable memory failure rate. Given the values of $\alpha_m, \alpha_{\oplus}, \alpha_{\gamma}$, a code based on graph with sufficient expansion can be chosen to build a fault tolerant memory. It is well known that a random graph is a good expander with high probability (see [9] and references therein). In the next section, we illustrate this fact with a few examples.

IV. NUMERICAL RESULTS

In this section, we illustrate with specific numerical examples the redundancies and tolerable failure rates associated with different values of γ and ρ . We first make the following observations. The redundancy of a memory system depends on the parameters γ and ρ of the LDPC code used. Different values of γ and ρ can result in same redundancy. To compare across different values of γ and ρ , the values of D_γ and α_γ have to be chosen consistently. How D_γ and α_γ scale with γ depends on the technology and implementation. Assuming that all gates are built out of universal NAND gates also does not answer the question fully as different implementations can lead to different values. Hence for the sake of illustration we consider a specific implementation. It should be noted that the subsequent discussion is for illustration purpose only. Accurate analysis for a given case can be carried out along the lines of the method we present in this section. For a given implementation, we fix the the values of γ and ρ thereby fixing the redundancy as well as α_γ and $\gamma(\rho-2)\alpha_\oplus$. We then use the bounds on the achievable expansion of a (γ, ρ) regular bipartite graph to find bounds on the value of $\alpha_{total} = \alpha(1 + 4\epsilon)(4\epsilon)/2$. This in turn provides bounds on the value of α_m for fixed γ and ρ .

A. Redundancy

Recall that the redundancy of a memory system is given by

$$\begin{aligned} R &= n(1 + D_\gamma + \gamma(\rho - 2))/rn \\ &\leq (1 + D_\gamma + \gamma(\rho - 2))/(1 - \gamma/\rho) \end{aligned}$$

For a fixed γ , R is minimum for a certain ρ depending on the value of D_γ . For example, if $D_\gamma = 2\gamma - 1$, then it can be shown that $\rho = 2\gamma$ minimizes the redundancy. This implies that a rate $1/2$ code has the least redundancy for a given γ .

Fig. 1(a) and Fig. 1(b) show the dependence of the redundancy on ρ for a given value of γ .

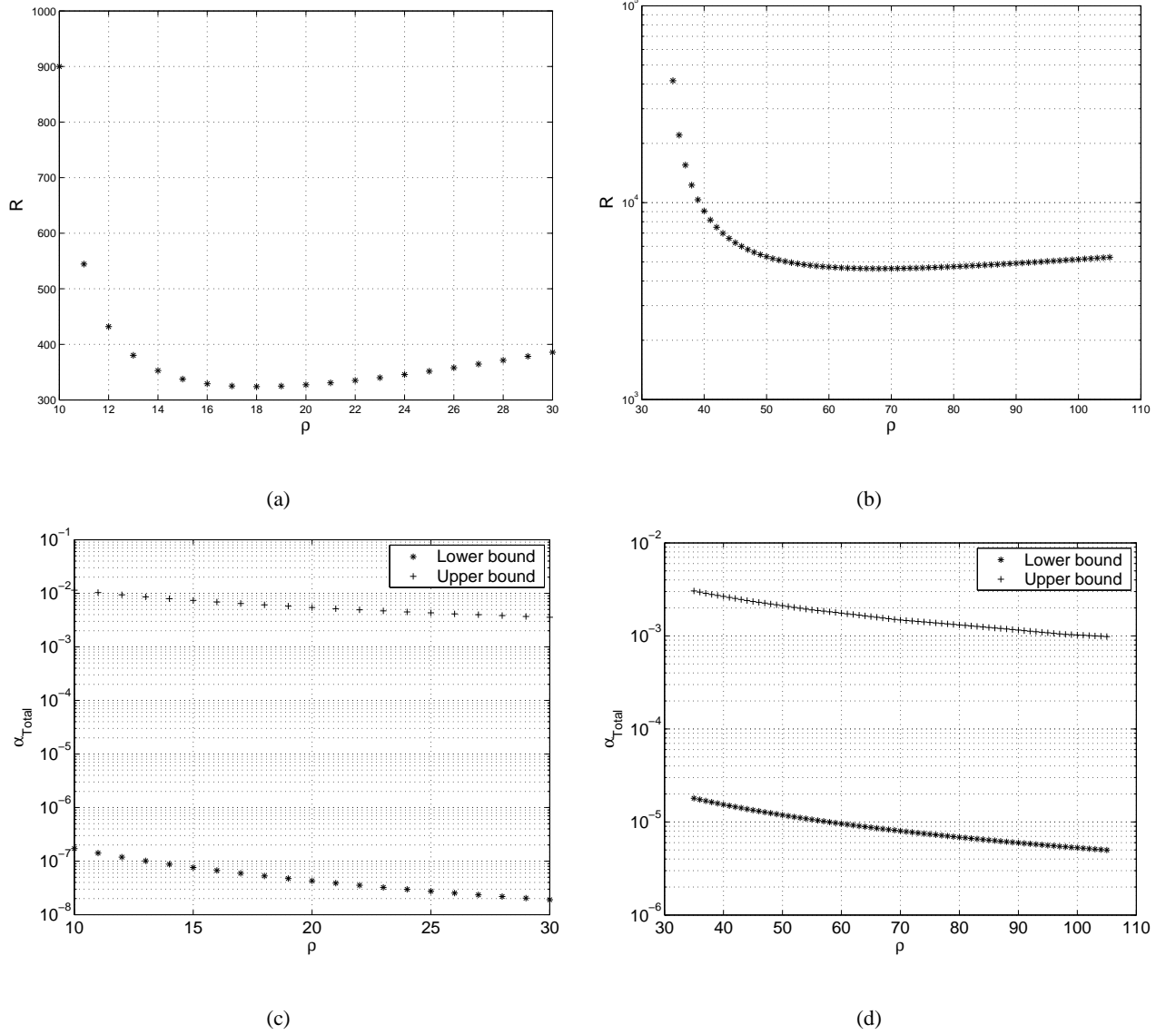


Fig. 1. Redundancies and bounds on expansion for different values of γ (a) redundancy for $\gamma = 9$ (b) redundancy for $\gamma = 34$ (c) bounds on α_{total} for $\gamma = 9$ (d) bounds on α_{total} for $\gamma = 34$

B. Bounds on Expansion

We make use of the following theorem from [9] to find an upper bound α_{total} for a given γ and ρ .

[Theorem 25, [9]]: Let B be a bipartite graph between n c -regular vertices and $(c/d)n$ d -regular vertices.

For all $0 < \alpha < 1$, there exists a set of αn c -regular vertices with at most

$$n \frac{c}{d} (1 - (1 - \alpha)^d) + O(1) \text{ neighbors}$$

It should be noted that the upper bound is tight for higher values of c .

Using this theorem, we can find an upper bound on α_{total} for a given γ and ρ . It should also be noted

that we look for graphs which expand by at least a factor of $(3/4 + \epsilon)$.

The following proposition from [11] addresses the issue of existence of expanders.

[Proposition 6, [11]]: Let $0 < r < 1$ and $0 < \delta < 1$ be any fixed constants, and let c be such that $(1 - \delta)c$ is an integer which is at least 2. Then for any n, m such that $r = 1 - m/n$ there is a Tanner graph with n variable nodes, m check nodes, and regular left degree c which is an $(\alpha n, \delta c)$ -expander, where

$$\alpha = (2e^{\delta c + 1}(\delta c / (1 - r))^{(1 - \delta)c})^{-\frac{1}{(1 - \delta)c - 1}}$$

It should be noted that the notation for expanders is different in [11]. Also, the proof does not guarantee that all the check nodes have same degree.

This proposition guarantees the existence of graphs with sufficient expansion and can be used to derive a lower bound on α_{total} for given γ and ρ . This in turn proves existence of memories which can tolerate $\alpha_m, \alpha_{\oplus}$ and α_{γ} fraction of failures in respective components as long as $\alpha_m + \gamma(\rho - 2)\alpha_{\oplus} + \alpha_{\gamma} < \alpha_{total}$. Figs. 1(c) and 1(d) illustrate the upper bounds and lower bounds on α_{total} for $\gamma = 9$ and $\gamma = 34$ respectively. We remark that the bounds have been derived numerically and we do not attempt to give closed form expressions for the bounds as the results are for illustration purpose only.

V. THE TAYLOR-KUZNETSOV SCHEME AND THE INDEPENDENT FAILURE MODEL

As mentioned in the introduction, Taylor [5], [18] was the first to investigate the capacity and fault-tolerant architectures of storage systems built entirely from unreliable components. His results were refined by Kuznetsov [6]. The aim of Taylor and Kuznetsov (TK) was to derive results analogous to the ones derived by Shannon on the capacity of communication systems. The spirit and methodology of Taylor and Kuznetsov's work [5], [6] is similar to Gallager's results [15] on LDPC codes. The bounds on probability of error are given for an ensemble of regular random LDPC codes of infinitely large length used in the correcting circuit. They are obtained under the assumptions that the bits in memory elements remain independent during the process of correction, i.e., under the assumption that the girth (the length of the

shortest cycle) of the Tanner graph corresponding to a code is infinitely large. Taylor and Kuznetsov considered a failure model in which a faulty component, generally a logic gate or a memory element, is subject to transient faults, i.e., faults that manifest themselves at particular time steps but do not necessarily persist for later times [1]. It is also assumed that gates fail independently of each other, and that the defects are not permanent, i.e., a gate that malfunctioned at some point in time may give correct output subsequently and that failure occurs by flipping the correct result with some probability p , i.e., if the correct result is “1”, the gate gives “0” and vice versa. Such failure mechanism is referred to as von Neumann type of error or as independent failure model. A faulty gate or memory element in this case can be modeled as as a binary symmetric channel (BSC) with crossover probability p .

A. The TK Scheme

The information to be stored is first encoded by a (n, γ, ρ) regular binary LDPC code. The stored codeword $\mathbf{v} = (v_1, v_2, \dots, v_n)$ consist of bits $v_i, 1 \leq i \leq n$ referred also as variables. Each variable bit $v_i, 1 \leq i \leq n$ is involved in γ parity-check equations by $\mathbf{x}H^T = \mathbf{c}$, where H is an $(m \times n)$ parity check matrix and all operations are in binary field. The degree of each check node is ρ . The vector $\mathbf{c} = (c_1, c_2, \dots, c_m)$ is called syndrome and c_j corresponds to the value of j^{th} parity-check sum for $1 \leq j \leq m$. Parity check c_j is said to be satisfied if $c_j = 0$ and unsatisfied if $c_j = 1$. A set of parity checks involving bit x_i is $\{c_i^{(1)}, c_i^{(2)}, \dots, c_i^{(\gamma)}\}$. After encoding, every coded bit x_i is replaced with γ bit-copies of itself $\{x_i^{(1)}, x_i^{(2)}, \dots, x_i^{(\gamma)}\}$ and stored in γ registers. All bit-copies initially have the same value. New estimates of each of these copies are obtained by using one combination of $\gamma - 1$ checks. Note that there are exactly $\binom{\gamma}{\gamma-1} = \gamma$ combinations. The estimates are obtained as follows.

- 1) Evaluate parity checks for each bit-copy (exclude one distinct parity check from the original set of checks for each bit-copy).
- 2) Flip the value of a particular bit-copy if half or more of the parity checks are unsatisfied.
- 3) Iterate (1) and (2).

The Tanner graph description of LDPC codes was unknown at the time of Taylor's paper. It is easy to see that each bit copy corresponds to an edge in the Tanner graph. The variable node corresponding to the edge is the corresponding bit and the check node is the parity check that is excluded in the estimation of that bit copy. If the update scheme is modified so that the check nodes indicate an estimate of the bit copy, then the update rule is an exact implementation of the hard decision message passing algorithm (known as Gallager B algorithm) for iterative decoding of LDPC codes (see [19] for a more detailed discussion). Such an equivalence is of great significance as expander graph arguments have been applied to message passing algorithms [10] and allows us to extend these results to the case of unreliable gates also.

The complexity and redundancy of the original TK scheme are given by

$$\mathcal{S} = (2 + D_{\gamma-1} + (\gamma - 1)(\rho - 1))\gamma n$$

$$R \leq (2 + D_{\gamma-1} + (\gamma - 1)(\rho - 1))\gamma / (1 - \gamma/\rho).$$

B. The Independent Failure Model

In this section, we extend our results to the independent failure model. By Chernoff bounds [8], it follows that a code which can correct a fraction of $p + \Delta$ errors achieves exponentially small probability of error on the BSC with crossover probability p [9]. In other words, if there are n components which can fail independently with probability p , then the probability that more than $p + \epsilon$ fraction of the components fail at any time is bounded by

$$P(\text{number of failures}/n > p + \Delta) \leq e^{-D(p+\Delta||p)n} \leq e^{-2\Delta^2 n}$$

where $D(x||y) = x \log(x/y) + (1 - x) \log((1 - x)/(1 - y))$ is the Kullback-Leibler divergence between Bernoulli random variables with parameters x and y respectively.

Now consider a memory architecture built from unreliable components subject to independent failures. Let p_m denote the probability of failure of memory element in time interval τ , p_{\oplus}, p_{γ} denote the probability of failure per use of an XOR gate and a γ -input majority logic gate respectively. Also, let $\epsilon_m, \epsilon_{\oplus}, \epsilon_{\gamma} > 0$ be such that $p_m + \epsilon_m = \alpha_m$, $p_{\oplus} + \epsilon_{\oplus} = \alpha_{\oplus}$ and $p_{\gamma} + \epsilon_{\gamma} = \alpha_{\gamma}$. Let $P_f(t)$ denote the probability of memory

failure at time t . For $\alpha_m, \alpha_\oplus, \alpha_\gamma$ and \mathcal{G} satisfying the conditions in Theorem 1, we now have the following theorem

Theorem 2: The proposed memory architecture has the following parameters for the independent failure model:

- 1) Information storage capability $\geq n(1 - \gamma/\rho)$
- 2) $R \leq (1 + D_\gamma + \gamma(\rho - 2))/(1 - \gamma/\rho)$
- 3) $P_f(L\tau) \leq Le^{(-\Omega(n))}$

Proof: (1) and (2) follow from our discussion in Section II. A memory failure may occur if the fraction of components which fail at a time is more than the tolerable fraction of errors. In L time steps, the correcting circuit is run for L times. The memory registers can fail L times. Hence, we have

$$P_f(L\tau) \leq L(e^{-2\epsilon_m^2 n} + e^{-2\epsilon_\oplus^2 n} + e^{-2\epsilon_\gamma^2 n})$$

■

The bound on the probability of memory failure given in Theorem 2 is a very weak bound and we do not try to improve it. Theorem 2 establishes the fact that in the proposed memory architecture, probability of memory failures decreases exponentially with the code length while the redundancy remains bounded. Theorem 2 has been stated in the same form as the main theorem in Kuznetsov's paper [6].

Hence, the proposed memory architecture has exponentially decreasing probability of memory failure in code length and redundancy which is roughly γ times less the TK scheme.

VI. DISCUSSION

Taylor in [5] remarks that memories have an associated non-zero storage capacity but an explicit calculation of the capacity is, in general, a difficult problem. For a given failure mechanism, finding storage capacity involves calculating the minimum redundancy to achieve arbitrarily low probability of error. The redundancy is a function of the coding scheme as well as the decoding algorithm. The TK scheme as well as the proposed memory architecture have finite redundancies and only give bounds on the

storage capacity. In this paper, we have shown that there exist reliable memories with redundancies less than that of the TK scheme. This implies that the proposed memory architecture improves the bound by a factor of γ at least in a few cases. The explicit calculation of the storage capacity still remains an unsolved problem. While the proposed architecture has less redundancy, the TK scheme may achieve better error exponents as it employs message passing algorithm which is in general more powerful than the parallel bit flipping algorithm. It is worth noting that Taylor in [5] describes the parallel bit flipping algorithm as a scheme for the update rule. He remarks that such an algorithm leads to complex interrelation between the errors as on successive iterations the values of the bits involved in the estimation of new value of each bit depend on previous value of the bit. We overcome this problem in this paper by using expander arguments. Also, extending the results from the adversarial model to the independent failure model using Chernoff bounds results in very weak bounds on the probability of memory failure. Using expander arguments directly for the independent failure model for both the proposed architecture and the TK scheme might result in better error exponents as well as lead to tighter bounds on the capacity.

Another problem which needs to be investigated is the bounds on the probabilities of failures of components, i.e., what are the upper bounds on the probability of failure of various components. Sipser and Spielman in [9] provided explicit construction of codes which can correct a certain fraction of errors. The fraction was later improved by Zemor in [20]. Barg and Zemor in [21] proved that expander codes achieve capacity on the BSC under iterative decoding. Guruswami and Indyk in [22] proposed linear time encodable and decodable codes which achieve optimal error correction performance. Study of fault-tolerant memory architectures based on these codes can provide the required bounds. However, these codes do not directly imply a specific implementation as is the case with parallel bit flipping algorithm. We noted earlier that Capalbo *et al.* [12] gave an explicit construction of expanders. However, the redundancies associated with such expanders are typically very high. This serves as another reason to consider expander codes and other linear time decodable codes based on expanders.

The proposed architecture as well as the TK scheme employ coding scheme based on regular LDPC

codes. The works of Richardson, Urbanke and Shokrollahi [23] and Luby, Mitzenmacher, Shokrollahi, and Spielman [24] show that well designed irregular codes perform close to capacity. Burshtein and Miller's work on expander graph arguments for message passing [10] is also based on irregular graphs. Investigating memory architectures based on irregular codes may serve as another avenue to study the storage capacity problem.

ACKNOWLEDGMENT

The authors would like to thank Milos Ivkovic for fruitful discussions.

REFERENCES

- [1] C. N. Hadjicostis and G. C. Verghese, "Coding approaches to fault tolerance in linear dynamic systems," *IEEE Trans. Inform. Theory*, vol. 51, no. 1, pp. 210–228, Jan. 2005.
- [2] J. V. Neumann, *Probabilistic Logics and the Synthesis of Reliable Organisms from Unreliable Components*, ser. Automata Studies. Princeton: Princeton University Press, 1956, pp. 43–98.
- [3] R. L. Dobrushin and S. I. Ortyukov, "Lower bound for the redundancy of self-correcting arrangements of unreliable functional elements," *Probl. Inform. Transm.*, vol. 13, pp. 59–65, 1977.
- [4] N. Pippenger, "Developments in 'the synthesis of reliable organisms from unreliable gates'," in *Symposia in Pure Mathematics*, 1990, pp. 311–324.
- [5] M. Taylor, "Reliable information storage in memories designed from unreliable components," *Bell System Technical Journal*, vol. 47, pp. 2299–2337, 1968.
- [6] A. Kuznetsov, "Information storage in a memory assembled from unreliable components," *Problems of Information Transmission*, vol. 9, pp. 254–264, 1973.
- [7] D. Spielman, "Highly fault-tolerant parallel computation," in *IEEE Conference on Foundations of Computer Science*, 1996, pp. 154–163.
- [8] H. Chernoff, "A measure of asymptotic efficiency for tests of a hypothesis based on the sum of observations," *Annals of Mathematical Statistics*, vol. 23, pp. 493–507, 1952.
- [9] M. Sipser and D. Spielman, "Expander codes," *IEEE Trans. Inform. Theory*, vol. 42, no. 6, pp. 1710–1722, Nov. 1996.
- [10] D. Burshtein and G. Miller, "Expander graph arguments for message-passing algorithms," *IEEE Trans. Inform. Theory*, vol. 47, no. 2, pp. 782–790, Feb. 2001.
- [11] J. Feldman, T. Malkin, R. A. Servedio, C. Stein, and M. J. Wainwright, "LP decoding corrects a constant fraction of errors," *IEEE Trans. Inform. Theory*, vol. 53, no. 1, pp. 82–89, Jan. 2007.

- [12] M. Capalbo, O. Reingold, S. Vadhan, and A. Wigderson, "Randomness conductors and constant-degree lossless expanders," in *STOC '02: Proceedings of the thirty-fourth annual ACM symposium on Theory of computing*. New York, NY, USA: ACM Press, 2002, pp. 659–668.
- [13] S. Lin and D. J. Costello, *Error Control Coding, Second Edition*. Upper Saddle River, NJ, USA: Prentice-Hall, Inc., 2004.
- [14] D. Terr, "Parity check matrix." [Online]. Available: <http://mathworld.wolfram.com/ParityCheckMatrix.html>
- [15] R. G. Gallager, *Low Density Parity Check Codes*. Cambridge, MA: M.I.T. Press, 1963.
- [16] A. Shokrollahi, "An introduction to low-density parity-check codes," in *Theoretical aspects of computer science: advanced lectures*. New York, NY, USA: Springer-Verlag New York, Inc., 2002, pp. 175–197.
- [17] R. M. Tanner, "A recursive approach to low complexity codes," *IEEE Trans. Inform. Theory*, vol. 27, pp. 533–547, Sept. 1981.
- [18] M. Taylor, "Reliable computation in computing systems designed from unreliable components," *Bell System Technical Journal*, vol. 47, pp. 2339–2266, Dec. 1968.
- [19] B. Vasic and S. K. Chilappagari, "An information theoretical framework for analysis and design of nano-scale fault-tolerant memories based on low-density parity-check codes," *IEEE Trans. Circuits Syst. I, Reg. Papers*, accepted for publication.
- [20] G. Zemor, "On expander codes," *IEEE Trans. Inform. Theory*, vol. 47, no. 2, pp. 835–837, Feb. 2001.
- [21] A. Barg and G. Zemor, "Error exponents of expander codes," *IEEE Trans. Inform. Theory*, vol. 48, no. 6, pp. 1725–1729, Jun. 2002.
- [22] V. Guruswami and P. Indyk, "Linear-time encodable/decodable codes with near-optimal rate," *IEEE Trans. Inform. Theory*, vol. 51, no. 10, pp. 3393–3400, Oct. 2005.
- [23] T. J. Richardson, M. Shokrollahi, and R. Urbanke, "Design of capacity-approaching irregular low-density parity-check codes," *IEEE Trans. Inform. Theory*, vol. 47, no. 2, pp. 638–656, Feb. 2001.
- [24] M. G. Luby, M. Mitzenmacher, M. A. Shokrollahi, and D. A. Spielman, "Improved low-density parity-check codes using irregular graphs," *IEEE Trans. Inform. Theory*, vol. 47, no. 2, pp. 585–598, Feb. 2001.